
ЕКОНО- МЕТРИКА

ПІДРУЧНИК

Ірина
ЛУК'ЯНЕНКО

Лариса
КРАСНІКОВА

Серія "ВИЩА ОСВІТА ХХІ СТОЛІТТЯ"

Ірина ЛУК'ЯНЕНКО
Лариса КРАСНІКОВА

ЕКОНО- МЕТРИКА

ПІДРУЧНИК

*Рекомендовано
Вченою радою Національного університету
"Киево-Могилянська академія"*



Київ
"Знання"
1998

Лук'яненко І. Г., Краснікова Л. І.
Л 84 Економетрика: Підручник. — К.: Товариство “Знання”, КОО, 1998.
— 494 с. (з табл., граф.). /
ISBN 966-7293-08-4

Даний підручник — результат багаторічного досвіду викладання авторами курсу економетрики студентам економічних спеціальностей Київського національного університету імені Тараса Шевченка та Національного університету “Києво-Могилянська академія”. Підручник побудовано за простою схемою. Матеріал викладено доступно, ясно і ґрунтовно. В кінці кожного розділу подається підсумок вивченого, а також наводяться необхідні навчальні завдання: контрольні запитання, задачі, вправи, тести, завдяки чому студенти зможуть не тільки перевірити свої знання, але й застосувати їх для аналізу реальних економічних процесів. Авторами підручника повною мірою враховано досвід викладання економетрики в провідних університетах США, Великобританії та Німеччини. Такого рівня підручник в Україні видається вперше.

Розраховано на студентів економічних спеціальностей, магістрантів, аспірантів, викладачів економічних дисциплін, широке коло фахівців.

Рецензенти:

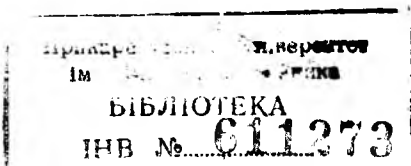
О. І. Ястремський, завідувач кафедри економічної теорії Національного університету “Києво-Могилянська академія”, доктор економічних наук, професор;

В. Л. Ревенко, професор Національного університету “Києво-Могилянська академія”, доктор економічних наук

ISBN 966-7293-08-4

© І. Г. Лук'яненко, Л. І. Краснікова, 1998

© Київська обласна організація
товариства “Знання” України



СТИСЛИЙ ЗМІСТ

ПЕРЕДМОВА	19
ВСТУП	20
Розділ 1. ЗАГАЛЬНІ ПРИНЦИПИ ПОБУДОВИ ЕКОНОМЕТРИЧНИХ МОДЕЛЕЙ. ЗВ'ЯЗОК ЕКОНОМЕТРИКИ З МАКРОЕКОНОМІКОЮ. ПРИКЛАДИ ПОБУДОВИ ЕКОНОМЕТРИЧНИХ МОДЕЛЕЙ	27
1.1. Зв'язок економетрики з макроекономікою. Приклади економетричних моделей	27
1.2. Роль економетрики в економічних дослідженнях ...	34
1.3. Інформаційна база економетричних моделей	36
Розділ 2. ПРОСТА ВИБІРКОВА ЛІНІЙНА РЕГРЕСІЯ	44
2.1. Загальне поняття про лінійну регресію	44
2.2. Оцінка параметрів лінійної регресії за допомогою методу найменших квадратів	44
2.3. Властивості простої вибіркової лінійної регресії ...	51
2.4. Коефіцієнти кореляції та детермінації	53
2.5. Поняття про ступені вільності	60
2.6. Простий ANOVA-аналіз у лінійній регресії: аналіз дисперсій	61
2.7. Перевірка простої регресійної моделі на адекватність. Поняття F-критерію Фішера	63
2.8. Інші критерії якості лінійної регресії	65
2.9. Імовірнісний зміст простої регресії	68
2.10. Закон розподілу параметрів. Математичне сподівання та дисперсія розподілу параметрів	78
2.11. Оцінка дисперсії випадкової величини ε	83
2.12. Побудова інтервалів довіри для параметрів β_0, β_1 ...	85
2.13. Прогнозування за моделями простої лінійної регресії	102
2.14. Властивості методу найменших квадратів	106
Розділ 3. КРИВІ ЗРОСТАННЯ	138
3.1. Поняття про криві зростання	138

3.2.	Найпростіші перетворення нелінійних моделей у лінійні. Експоненційна функція. Приклади застосування експоненційних функцій у бізнесі та фінансах. Зведення до лінійної регресії.....	139
3.3.	Степенева (мультиплікативна) функція. Зведення до лінійної регресії. Приклади застосування степених функцій у бізнесі та фінансах	142
3.4.	Зворотні перетворення. Приклади їх застосування на практиці	144
3.5.	Квадратичні функції	146
3.6.	Експоненційна модифікована крива.....	148
3.7.	Крива Гомперця	149
3.8.	Логістична крива	150
3.9.	Декілька простих методів обчислення невідомих параметрів нелінійних моделей. Метод трьох точок	151
3.10.	Зв'язок між коефіцієнтами еластичності і параметрами кривих зростання	153
Розділ 4.	БАГАТОФАКТОРНА РЕГРЕСІЯ	171
4.1.	Вступ. Приклади використання багатофакторного регресійного аналізу на практиці.....	171
4.2.	Класична лінійна багатофакторна модель. Основні припущення у багатофакторному регресійному аналізі	173
4.3.	Етапи побудови багатофакторної регресійної моделі.....	176
4.4.	Розрахунок невідомих параметрів багатофакторної регресії за методом найменших квадратів (МНК).....	178
4.5.	Властивості методу найменших квадратів	179
4.6.	Коефіцієнт множинної кореляції та детермінації. 180	
4.7.	Коефіцієнт детермінації R^2 та оцінений коефіцієнт детермінації \bar{R}^2	181
4.8.	ANOVA-дисперсійний аналіз	184
4.9.	Перевірка моделі на адекватність за F -критерієм Фішера	185
4.10.	Матричний підхід до лінійної багатофакторної регресії	187

4.11.	Методи побудови багатофакторної регресійної моделі	198
Розділ 5.	ОСОБЛИВІ ВИПАДКИ У БАГАТОФАКТОРНОМУ РЕГРЕСІЙНОМУ АНАЛІЗІ	228
5.1.	Мультиколінеарність та її наслідки. Знаходження мультиколінеарності та її вилучення	228
5.2.	Гетероскедастичність та її наслідки. Знаходження гетероскедастичності та її вилучення	249
5.3.	Автокореляція	273
5.4.	Авторегресивні і дистрибутивно-лагові моделі	278
5.5.	Димму-змінні	310
Розділ 6.	ЕКОНОМЕТРИЧНІ СИМУЛЬТАТИВНІ МОДЕЛІ	360
6.1.	Поняття про одночасну залежність економічних змінних. Приклади економетричних симульта- тивних моделей	360
6.2.	Порушення класичного припущення в симуль- тивних моделях	365
6.3.	Методи оцінювання невідомих параметрів у моде- лях симульта- тивних рівнянь	370
6.4.	Проблема ототожнення в симульта- тивних моде- лях. Загальні поняття	376
6.5.	Правильне, або точне ототожнення	379
6.6.	Проблема переототожнення	382
6.7.	Основні правила ототожнення	384
6.8.	Методи оцінювання невідомих параметрів симуль- тивних моделей	388
6.9.	Рекурсивні моделі	395
6.10.	Основні висновки про симульта- тивні моделі	396
Розділ 7.	ЕЛЕМЕНТИ ТЕОРІЇ МАТРИЦЬ ТА ОСНОВНІ ЗАКОНИ РОЗПОДІЛУ ВИПАДКОВИХ ВЕЛИЧИН .	418
7.1.	Елементи теорії матриць	418
7.2.	Матриці і визначники	427
7.3.	Геометрична інтерпретація векторів і матриць	439
7.4.	Системи лінійних рівнянь	443
7.5.	Характеристичні корені і вектори, характеристич- не рівняння	444
7.6.	Квадратичні форми та визначені матриці	454
7.7.	Обчислення і матрична алгебра	457
7.8.	Основні поняття теорії ймовірностей	470

ДОДАТКИ	482
Додаток 1. Графік і таблиця нормального закону розподілу	482
Додаток 2. Графік і таблиця F -розподілу Фішера	483
Додаток 3. Графік і таблиця t -розподілу Ст'юдента	484
Додаток 4. DW-статистика Дарбіна — Уотсона. Критичні точки d_L та d_U при рівні значимості $d = 0.05$	485
Додаток 5. DW-статистика Дарбіна — Уотсона. Критичні точки d_L та d_U при рівні значимості $d = 0.01$	487
Список літератури	489

ЗМІСТ

ПЕРЕДМОВА	19
ВСТУП	20
Що таке економетрика?	20
Чому економетрика є окремою дисципліною?	20
Виникнення, становлення та розвиток економетрики ..	21
Етапи проведення економетричного аналізу	22
Математичне та статистичне підґрунтя. Теоретична і прикладна економетрика	24
Роль комп'ютера в сучасній економетриці	25
Розділ 1. ЗАГАЛЬНІ ПРИНЦИПИ ПОБУДОВИ ЕКОНОМЕТРИЧНИХ МОДЕЛЕЙ. ЗВ'ЯЗОК ЕКОНОМЕТРИКИ З МАКРОЕКОНОМІКОЮ. ПРИКЛАДИ ПОБУДОВИ ЕКОНОМЕТРИЧНИХ МОДЕЛЕЙ	27
1.1. Зв'язок економетрики з макроекономікою. Приклади економетричних моделей	27
1.1.1. Модель валового національного продукту	27
1.1.2. Класична модель економіки	29
1.1.3. Повна кейнсіанська модель	33
1.2. Роль економетрики в економічних дослідженнях	34
1.3. Інформаційна база економетричних моделей	36
1.3.1. Динамічні ряди та їхні характеристики	36
1.3.2. Варіаційні ряди та їхні характеристики	40
Розділ 2. ПРОСТА ВИБІРКОВА ЛІНІЙНА РЕГРЕСІЯ	44
2.1. Загальне поняття про лінійну регресію	44
2.2. Оцінка параметрів лінійної регресії за допомогою методу найменших квадратів	44
2.3. Властивості простої вибіркової лінійної регресії	51
2.4. Коефіцієнти кореляції та детермінації	53
2.4.1. Поняття про коефіцієнт кореляції	53
2.4.2. Декомпозиція дисперсій. Поняття про коефіцієнт детермінації	53
2.4.3. Зв'язок між коефіцієнтом кореляції та нахилом b_1 ...	57
2.4.4. Зв'язок між коефіцієнтом кореляції (r) і коефіцієнтом детермінації (R^2)	58
2.5. Поняття про ступені вільності	60

2.6.	Простий ANOVA-аналіз у лінійній регресії: аналіз дисперсій	61
2.7.	Перевірка простої регресійної моделі на адекватність. Поняття F-критерію Фішера	63
2.8.	Інші критерії якості лінійної регресії	65
2.9.	Імовірнісний зміст простої регресії	68
2.9.1.	Узагальнена регресійна модель	68
2.9.2.	Класична модель лінійної регресії: основні припущення, що лежать в основі методу найменших квадратів ..	69
2.9.3.	Розподіл залежної змінної y	76
2.10.	Закон розподілу параметрів. Математичне сподівання та дисперсія розподілу параметрів	78
2.11.	Оцінка дисперсії випадкової величини ε	83
2.12.	Побудова інтервалів довіри для параметрів β_0, β_1	85
2.12.1.	Поняття про t -тест Ст'юдента. Перевірка нуль-гіпотези за допомогою t -тесту Ст'юдента	85
2.12.2.	t -тест Ст'юдента для перевірки на значимість параметрів b_0 і b_1 , визначених за методом найменших квадратів	89
2.12.3.	Знаходження інтервалів довіри для параметрів β_0 і β_1 за t -тестом Ст'юдента	93
2.12.4.	Тест Фішера для перевірки нуль-гіпотези $\beta_1=0$	95
2.12.5.	Порівняння F -критерію Фішера з t -критерієм Ст'юдента	97
2.12.6.	t -тест для оцінки значимості коефіцієнта кореляції	99
2.13.	Прогнозування за моделями простої лінійної регресії	102
2.14.	Властивості методу найменших квадратів	106
2.14.1.	Найпоширеніші критерії порівняння методів оцінювання	106
2.14.2.	Властивості оцінок, отриманих за методом найменших квадратів	108
Підсумуємо вивчене		109
	Після вивчення цього розділу ви зможете	109
	Короткий огляд розділу	109
	Основні формули	110
Вправи і коментарі		112
Виконайте самостійно		118
	Вправи. Проста лінійна регресія	118
Тести		121
	Тест 1	121
	Тест 2	124
Відповісти на запитання: так/ні; якщо ні, поясніть чому		126

	Частина 1	126
	Частина 2	127
Запитання і задачі		127
Відповіді		129
Застосування економічних знань		132
	Модель оцінки капітальних активів (МОКА-модель) ..	132
	Модель залежності зміни частки заробітної плати населення України у загальному доході	135
Розділ 3. КРИВІ ЗРОСТАННЯ		138
3.1.	Поняття про криві зростання	138
3.2.	Найпростіші перетворення нелінійних моделей у лінійні. Експоненційна функція. Приклади застосування експоненційних функцій у бізнесі та фінансах. Зведення до лінійної регресії	139
3.2.1.	Експоненційна функція. Приклади застосування експоненційних функцій у бізнесі та фінансах	139
3.2.2.	Зведення експоненційної кривої до простої лінійної функції	142
3.3.	Степенева (мультиплікативна) функція. Зведення до лінійної регресії. Приклади застосування степеневих функцій у бізнесі та фінансах	142
3.4.	Зворотні перетворення. Приклади їх застосування на практиці	144
3.5.	Квадратичні функції	146
3.6.	Експоненційна модифікована крива	148
3.7.	Крива Гомперця	149
3.8.	Логістична крива	150
3.9.	Декілька простих методів обчислення невідомих параметрів нелінійних моделей. Метод трьох точок	151
3.10.	Зв'язок між коефіцієнтами еластичності і параметрами кривих зростання	153
Підсумуємо вивчене		155
	Після вивчення цього розділу ви зможете	155
	Короткий огляд розділу	155
	Основні формули	156
Вправи і коментарі		157
Виконайте самостійно		158
	Вправи. Криві зростання	158
Тести		165
Відповісти на запитання: так/ні; якщо ні, поясніть чому		166
Відповіді		167
Застосування економічних знань		167

Зростання малих приватизованих підприємств в Україні	167
Прогнозування середнього темпу зростання обсягів продажу товару	169
Розділ 4. БАГАТОФАКТОРНА РЕГРЕСІЯ	171
4.1. Вступ. Приклади використання багатофакторного регресійного аналізу на практиці	171
4.2. Класична лінійна багатофакторна модель. Основні припущення у багатофакторному регресійному аналізі	173
4.3. Етапи побудови багатофакторної регресійної моделі	176
4.4. Розрахунок невідомих параметрів багатофакторної регресії за методом найменших квадратів (МНК)	178
4.5. Властивості методу найменших квадратів	179
4.6. Коефіцієнт множинної кореляції та детермінації	180
4.7. Коефіцієнт детермінації R^2 та оцінений коефіцієнт детермінації \bar{R}^2	181
4.8. ANOVA-дисперсійний аналіз	184
4.9. Перевірка моделі на адекватність за F -критерієм Фішера	185
4.9.1. Зв'язок між коефіцієнтом детермінації (R^2) та F -відношенням Фішера	185
4.9.2. Тестування адекватності багатофакторної регресійної моделі, виходячи із значення R^2	186
4.10. Матричний підхід до лінійної багатофакторної регресії	187
4.10.1. Запис лінійної багатофакторної регресії у матричному вигляді	187
4.10.2. Припущення класичної лінійної багатофакторної регресії у матричному вигляді	189
4.10.3. Оцінювання невідомих параметрів у багатофакторній регресії	190
4.10.4. Дисперсійно-коваріаційна матриця параметрів регресії	192
4.10.5. Оцінка дисперсії випадкової величини ε	193
4.10.6. Перевірка гіпотез щодо параметрів багатофакторної регресії у матричному вигляді	194
4.10.7. Знаходження інтервалів довіри для параметрів	196
4.10.8. Прогнозування за багатофакторною регресійною моделлю	196
4.11. Методи побудови багатофакторної регресійної моделі	198

4.11.1. Вибір “найкращого” рівняння регресії	198
4.11.2. Метод усіх можливих регресій	198
4.11.3. Метод виключень	200
4.11.4. Кроковий регресійний метод	200
4.11.5. Принцип “додаткової суми квадратів”. Обчислення додаткової суми квадратів	201
4.11.6. Часткові і послідовні F -критерії	203
Підсумуємо вивчене	204
Після вивчення цього розділу ви зможете	204
Короткий огляд розділу	204
Основні формули	205
Вправи і коментарі	206
Виконайте самостійно	209
Вправи. Багатофакторна лінійна регресія	209
Тести	212
Відповісти на запитання: так/ні; якщо ні, поясніть чому	214
Запитання і задачі	215
Відповіді	216
Застосування економічних знань	217
Приклади практичного застосування багатофакторної регресії	217
Моделювання процесу малої приватизації в Україні	221
Аналіз впливу соціальних факторів на рівень злочинності в Україні	225
Розділ 5. ОСОБЛИВІ ВИПАДКИ У БАГАТОФАКТОРНОМУ РЕГРЕСІЙНОМУ АНАЛІЗІ	228
5.1. Мультиколінеарність та її наслідки. Знаходження мультиколінеарності та її вилучення	228
5.1.1. Визначення мультиколінеарності та її природа	228
5.1.2. Теоретичні наслідки мультиколінеарності в загальному випадку	232
5.1.3. Практичні наслідки мультиколінеарності	233
5.1.4. Тестування наявності мультиколінеарності та засоби її вилучення	237
5.1.4.1. Тестування наявності мультиколінеарності	237
5.1.4.2. Визначення рівня мультиколінеарності	239
5.1.4.3. Засоби вилучення мультиколінеарності	239
5.1.5. Основні висновки щодо наявності мультиколінеарності в регресійній моделі	243
Додаток 5.1	244
5.2. Гетероскедастичність та її наслідки. Знаходження гетероскедастичності та її вилучення	249

5.2.1.	Визначення гетероскедастичності та її природа	249
5.2.2.	Правдоподібність припущення про гомоскедастичність	251
5.2.3.	Наслідки порушення припущення про гомоскедастичність	253
5.2.4.	Оцінювання параметрів методом узагальнених найменших квадратів (УНК) у разі гетероскедастичності	254
5.2.5.	Тестування наявності гетероскедастичності	257
5.2.6.	Вилучення гетероскедастичності	263
5.2.6.1.	Ефективність оцінок трансформованої моделі	267
5.2.7.	Ілюстративний приклад. Тест Голдфелда та Квондта	268
5.2.8.	Основні висновки щодо наявності гетероскедастичності в регресійній моделі	272
5.3.	Автокореляція	273
5.3.1.	Природа автокореляції. Основні поняття та означення	273
5.3.2.	Тестування автокореляції	274
5.3.3.	Оцінка параметрів регресійної моделі при наявності автокореляції	275
5.4.	Авторегресивні і дистрибутивно-лагові моделі	278
5.4.1.	Природа авторегресивних моделей. Приклади практичного застосування авторегресивних моделей ..	278
5.4.1.1.	Приклади використання лагових моделей в економіці. Роль “часу” або “часового лагу” в економіці	278
5.4.1.2.	Причини лагів	281
5.4.2.	Оцінка параметрів дистрибутивно-лагових моделей ..	283
5.4.2.1.	Послідовна оцінка дистрибутивно-лагових моделей ..	283
5.4.2.2.	Підхід Койка до дистрибутивно-лагових моделей	284
5.4.3.	Перша модифікація моделі Койка: модель адаптивних очікувань	287
5.4.4.	Друга модифікація моделі Койка: модель часткових пристосувань	289
5.4.5.	Комбінація моделей адаптивних очікувань і часткових пристосувань	291
5.4.6.	Оцінювання параметрів авторегресивних моделей	291
5.4.7.	Метод допоміжних змінних	293
5.4.8.	Виявлення автокореляції в авторегресивних моделях: <i>h</i> -тест Дарбіна	294
5.4.9.	Ілюстративний приклад: попит на гроші в Україні ..	296
5.4.10.	Ілюстративний приклад: короткострокова і довгострокова функції сукупного споживання у США	298
5.4.11.	Підхід Ш. Альмона до дистрибутивно-лагових моделей: поліноміальний лаг Альмона	300

5.4.12.	Причинність в економіці: тест Гренжера (Granger)	306
5.4.13.	Основні висновки про авторегресивні моделі	307
5.5.	Dummy-змінні	310
5.5.1.	Природа dummy-змінних	310
5.5.2.	Регресія однієї кількісної та однієї якісної змінної двох класів, або категорій	311
5.5.3.	Регресія кількісної змінної та однієї якісної змінної з більш ніж двома класами	314
5.5.4.	Регресія однієї кількісної і двох якісних змінних	315
5.5.5.	Порівняння двох регресійних моделей	317
5.5.5.1.	Shoу-тест для порівняння регресійних моделей	318
5.5.6.	Порівняння двох регресій: підхід з використанням dummy-змінної	320
5.5.7.	Особливий випадок одночасного впливу двох dummy-змінних	321
5.5.8.	Використання dummy-змінних у сезонному аналізі ...	323
5.5.9.	Основні висновки щодо моделей з dummy-змінними ..	325
	Підсумуємо вивчене	326
	Після вивчення цього розділу ви зможете	326
	Короткий огляд розділу	326
	Основні формули	328
	Вправи і коментарі	329
	Тести	332
	Відповіді на запитання: так/ні; якщо ні, поясніть чому	335
	Запитання і задачі	335
	Відповіді	337
	Виконайте самостійно	339
	Вправи. Мультиколінеарність	339
	Вправи. Гетероскедастичність	343
	Вправи. Авторегресивні моделі	348
	Вправи. Dummy-змінні	353
	Застосування економічних знань	357
	Визначення ціни на акцію	357
	Функція витрат. Тестування мультиколінеарності	358
	Розділ 6. ЕКОНОМЕТРИЧНІ СИМУЛЬТАТИВНІ МОДЕЛІ	360
6.1.	Поняття про одночасну залежність економічних змінних. Приклади економетричних симультаивних моделей	360
6.2.	Порушення класичного припущення в симультаивних моделях	365
6.2.1.	Отримання зміщених оцінок (оцінок з відхиленням) в моделях симультаивних рівнянь	367

6.2.2.	Приклад обчислення оцінок з відхиленнями в симультивних моделях	369
6.3.	Методи оцінювання невідомих параметрів у моделях симультивних рівнянь	370
6.3.1.	Загальні поняття про методи оцінювання	370
6.3.2.	Попередні відомості про структурні моделі. Ілюстративний приклад	371
6.3.3.	Структурні моделі скороченої форми	372
6.3.4.	Знаходження параметрів скороченої форми та оцінка їхнього впливу на змінні моделі	373
6.4.	Проблема ототожнення в симультивних моделях. Загальні поняття	376
6.5.	Правильне, або точне ототожнення	379
6.6.	Проблема переототожнення	382
6.7.	Основні правила ототожнення	384
6.8.	Методи оцінювання невідомих параметрів симультивних моделей	388
6.8.1.	Оцінювання точно ототожненого рівняння: метод непрямих найменших квадратів (ННК)	388
6.8.1.1.	Властивості ННК-оцінок	391
6.8.2.	Оцінка переототожненого рівняння: метод двокрокових найменших квадратів (2МНК)	391
6.8.3.	Характерні особливості методу 2МНК	393
6.8.4.	Ілюстративний приклад	394
6.9.	Рекурсивні моделі	395
6.10.	Основні висновки про симультивні моделі	396
Підсумуємо вивчене	398	
	Після вивчення цього розділу ви зможете	398
	Короткий огляд розділу	398
	Основні формули	399
Вправи і коментарі	399	
Тести	402	
Відповіді на запитання: так/ні; якщо ні, поясніть чому	404	
Відповіді	405	
Виконайте самостійно	406	
	Вправи	406
Додаток 6.1	415	
	Зміщеність оцінок, отриманих методом непрямих найменших квадратів (ННК)	415
Додаток 6.2	416	
	Оцінка стандартних відхилень методом 2МНК	416

Розділ 7. ЕЛЕМЕНТИ ТЕОРІЇ МАТРИЦЬ ТА ОСНОВНІ ЗАКони РОЗПОДІЛУ ВИПАДКОВИХ ВЕЛИЧИН .	418	
7.1. Елементи теорії матриць	418	
7.1.1.	Основні визначення	418
7.1.2.	Особливі види матриць	419
7.1.3.	Алгебраїчні дії над матрицями	420
7.2. Матриці і визначники	427	
7.2.1.	Теорія визначників	427
7.2.2.	Матричні рівняння і характеристичні поліноми	431
7.2.3.	Ранг матриці і ранг системи векторів	434
7.2.4.	Комбіновані матриці	436
7.3. Геометрична інтерпретація векторів і матриць	439	
7.3.1.	Вектор, операції над векторами, векторний простір	439
7.3.2.	Лінійна комбінація векторів. Базисні вектори. Лінійна залежність векторів	441
7.4. Системи лінійних рівнянь	443	
7.5. Характеристичні корені і вектори, характеристичне рівняння	444	
7.5.1.	Характеристичне рівняння	444
7.5.2.	Характеристичні вектори	445
7.5.3.	Загальні результати для характеристичних коренів і векторів	446
7.5.4.	Приведення матриці до діагональної форми	447
7.5.5.	Ранг характеристичної матриці	447
7.5.6.	Власні значення матриці	448
7.5.7.	Слід матриці	448
7.5.8.	Визначник матриці	449
7.5.9.	Спектральна декомпозиція матриць	450
7.5.10.	Степені матриці	450
7.5.11.	Рівнопотужні матриці	451
7.5.12.	Розкладання матриці на множники	452
7.5.13.	Узагальнені обернені матриці	453
7.6. Квадратичні форми та визначені матриці	454	
7.6.1.	Невід'ємно визначені матриці	455
7.6.2.	Рівнопотужні квадратичні форми	455
7.6.3.	Ранжування матриць	456
7.7. Обчислення і матрична алгебра	457	
7.7.1.	Диференціали та ряди Тейлора	457
7.7.2.	Оптимізація	461
7.7.3.	Умовна оптимізація	463
7.7.4.	Перетворення	466

Вправи	468
7.8. Основні поняття теорії ймовірностей	470
7.8.1. Інтегральна функція розподілу ймовірностей випадкової величини	470
7.8.2. Диференційна функція (функція густини) розподілу ймовірностей випадкової величини	471
7.8.3. Імовірнісний зміст диференційної функції	472
7.8.4. Закон рівномірного розподілу ймовірностей	472
7.8.5. Нормальний закон розподілу	473
7.8.5.1. Вплив параметрів нормального розподілу на форму нормальної кривої	475
7.8.6. Розподіл χ^2 (x_i — квадрат)	477
7.8.7. T -розподіл Ст'юдента	478
7.8.8. F -розподіл Фішера	479
ДОДАТКИ	482
Додаток 1. Графік і таблиця нормального закону розподілу	482
Додаток 2. Графік і таблиця F -розподілу Фішера	483
Додаток 3. Графік і таблиця t -розподілу Ст'юдента	484
Додаток 4. DW-статистика Дарбіна — Уотсона. Критичні точки d_L та d_U при рівні значимості $d = 0.05$	485
Додаток 5. DW-статистика Дарбіна — Уотсона. Критичні точки d_L та d_U при рівні значимості $d = 0.01$	487
Список літератури	489

ПЕРЕДМОВА

Підручник “Економетрика” — одна з перших спроб викласти цю нову для вищих навчальних закладів України дисципліну, яка є певним стандартом у підготовці сучасних економістів. Чим популярнішою стає ця дисципліна, тим частіше можна почути запитання: *“Що таке економетрика?”*

Економетрика — це наука, що вивчає кількісні закономірності та взаємозв'язки економічних об'єктів і процесів за допомогою математико-статистичних методів та моделей. Звичайно, таке визначення дещо обмежене, бо не дає уяви про економетрику як про синтетичну дисципліну, яка поєднує в собі економічну теорію, математичну економіку, теорію матриць, теорію ймовірностей, економічну та математичну статистику. Оволодіння економетрикою вимагає певної математичної та економічної підготовки. Ми намагалися дати всебічне уявлення про цей курс, не виходячи за межі початкового володіння елементами матричної алгебри, теорії ймовірностей, макро- та мікроекономіки. Головний наголос зроблено на принципах побудови економетричних моделей, оцінювання невідомих параметрів, перевірки моделей та всебічного аналізу отриманих результатів.

Крім викладу стандартних розділів, таких як проста лінійна регресія, криві зростання, багатофакторна регресія, в підручнику розглядаються симультивні моделі та особливі випадки в багатофакторному аналізі. Для тих, хто не дуже вільно володіє необхідним математичним апаратом, ми передбачили математичний додаток з основних понять теорії матриць та теорії ймовірностей.

Аналізуючи сучасні проблеми української та світової економіки, ми намагались прокласти місток між загальними теоретичними поняттями та розв'язанням практичних проблем сьогодення. Оскільки ця книжка перш за все — навчальний посібник, то кожний розділ супроводжується аналізом результатів розв'язку конкретної економічної проблеми, різноманітними вправами, тестами та контрольними запитаннями для самоперевірки.

Підручник написаний з використанням досвіду викладання курсів “Теорія економічного ризику”, “Соціально-економічне прогнозування”, “Економетрика-1” у Київському національному університеті імені Тараса Шевченка та Національному університеті “Києво-Могилянська академія”, а також із використанням досвіду виконання проекту “Модернізація економічної освіти в Україні”, спрямованого на перепідготовку викладачів економічних дисциплін українських вузів.

При цьому враховувався досвід викладання курсів економетрики в університетах США, зокрема в Техаському та Чиказькому.

Написання цього підручника було б неможливим без підтримки Фонду “Євразія”.

ВСТУП

Що таке економетрика?

Економічна теорія кінця ХХ століття — це потужний інструмент дослідження, аналізу, пояснення та прогнозування динаміки економічної реальності, яка невпинно розвивається та поповнюється розмаїттям подій і явищ, що безпосередньо впливають на життя як цілих суспільств та груп людей, так і окремих індивідів. Без перебільшення можна сказати, що сучасна економіка з різноманітністю її підходів та засобів спостереження, методів обробки інформації та моделювання економічних систем стала значною мірою міждисциплінарним утворенням, що акумулює результати багатьох дисциплін, таких як математика й інформатика, статистика і теорія імовірностей та інші. Серцевина цього утворення, що визначається терміном *економічна наука*, еволюціонувала так стрімко, що без спеціального путівника важко навіть охопити мережу окремих дисциплін, котрі виросли на стовбурі дерева економіки. І цей процес триває.

Особливе місце на дереві економіки посідає економетрика. **Що ж являє собою економетрика?**

У буквальному перекладі *економетрика* означає “вимірювання економіки”. Але, звичайно, такий переклад не відображає справжнього стану речей. Поняття економетрики є набагато ширшим, хоча вимірювання залишається однією з її складових частин.

За класичним визначенням, *економетрика* — це наука, що вивчає кількісні закономірності та взаємозв'язки економічних об'єктів і процесів за допомогою математико-статистичних методів та моделей.

Економетрика є інструментом, який дозволяє перейти від якісного рівня аналізу до рівня, що використовує кількісні статистичні значення досліджуваних величин.

Можливості економетрики залежать не тільки від якості тих моделей, що мають відображати об'єктивні закономірності економічних процесів, а й значною мірою від якості самої економетричної технології, яка сьогодні є достатньо розвинутою.

Чому економетрика є окремою дисципліною?

Економетрика є синтезною дисципліною, вона поєднує в собі економічну теорію, математичну економіку, економічну та математичну статистику. Що ж відрізняє економетрику від економічної теорії? Економічна теорія пропонує твердження чи гіпотези, що за своєю сутністю є переважно якіс-

ними. Наприклад, мікроекономічна теорія стверджує, що зниження ціни товару сприятиме зростанню попиту на цей товар. Але сама теорія не наводить жодного кількісного виміру взаємозв'язку цих двох показників, тобто вона не показує, наскільки зросте чи зменшиться кількість у результаті певної зміни ціни товару. Таким чином, економічна теорія приймає без доказів обернену залежність між ціною та попитом на товар. Завдання економетрики полягає в обчисленні відповідних кількісних оцінок. Інакше кажучи, економетрика забезпечує кількісну сторону економічної теорії.

На відміну від чистої математичної економіки, яка виражає економічну теорію в математичній формі без мети вимірювання чи емпіричного підтвердження теорії, економетрика, навпаки, зацікавлена в емпіричному підтвердженні економічної теорії. Економетрист часто використовує математичні рівняння, запропоновані математиком-економістом, але перетворює їх у форму, найбільш придатну для емпіричного тестування. Перетворення математичних рівнянь в економетричні вимагає великої винахідливості та практичних навичок.

Відмінність економетрики від економічної статистики також дуже наочна. Економічна статистика в основному стосується збору, обробки та зображення економічних даних у формі діаграм і таблиць. У цьому полягає робота економіста-статистика. Зібрані дані становлять основу для роботи економетриста. Але інколи економісти-статистики використовують зібрані дані тільки для того, щоб перевірити ту чи іншу економічну теорію, і мимоволі “перетворюються” на економетристів.

Спеціалісти-економетристи часто потребують спеціальних засобів для обробки даних, особливо отриманих у результаті неконтрольованого експерименту. Економетрист в основному залежить від інформації, що не може бути прямо проконтрольована. Така інформація часто містить помилки вимірювання, тому спеціаліст у галузі економетрики має розробляти спеціальні методи для аналізу подібних помилок.

Економетрика як самостійна наукова дисципліна використовує в той же час поняття та методи розв'язку задач з багатьох розділів математики, наприклад математичної статистики, теорії імовірностей, математичного програмування, лінійної алгебри, систем нелінійних рівнянь та ін.

Виникнення, становлення та розвиток економетрики

Як самостійна дисципліна *економетрика* сформувалась у 20—30-х роках нашого століття завдяки працям Г. Мура та Г. Шульца. До цього вже були відомі спроби математичної формалізації економіко-статистичних даних у працях В.Парето (рівняння гіперболи для опису розподілу прибутку населення (1897 р.)) та працях Р.Хукера й А.Чупрова з кореляційного аналізу економічних процесів. У перших працях у рамках економетрики розроблялись аналітико-статистичні моделі. Здебільшого, це були рівняння

лінійної регресії з параметрами, оцінюваними за методом найменших квадратів. Такі рівняння дозволяли описати як функції попиту й залежність їх від прибутків, обсягів випуску продукції, рівня цін, податків та ін., так і функції пропозиції, виробничі функції, що відображали технологічну залежність випуску продукції від затрат праці та засобів виробництва.

Одна з перших виробничих функцій була побудована Коббом та Дугласом у 1928 році, а потім узагальнена Р. Солоу. З часом регресійні моделі виявили свою обмеженість для опису узагальнених модельних економічних комплексів і нормативних моделей.

Починаючи з 30-х років відомі економісти Я. Танберген, Л. Клейн, Р. Стоун та інші розробили моделі економіки, які описували статистичні зв'язки виробництва, кінцевого індивідуального і державного попиту, цін, податків, зовнішньої торгівлі, пропозиції робочої сили, накопичення та зношування капіталу. Такі моделі склалися вже з багатьох рівнянь, у зв'язку з чим значно ускладнились проблеми оцінювання невідомих параметрів. А це в свою чергу привело до необхідності використання нового математичного апарату та розширило можливості практичного використання економетрики.

До числа типових економіко-математичних моделей, які на сьогоднішній день розробляє і вивчає економетрика, відносяться: виробничі функції, функції попиту різних груп споживачів та цільові функції переваги споживачів, статистичні та динамічні міжгалузеві моделі виробництва, розподілу і споживання продукції, моделі загальної економічної рівноваги. Це певним чином споріднює економетрику з макроекономікою, тому не дивно, що починаючи з 50-х років вони активно розвивалися поряд.

У практичних дослідженнях економетричні методи використовуються не тільки в економіці. Вони поширені у біології, історії, соціології та інших суспільних і природничих науках, де необхідно розробляти та оцінювати моделі, які формалізують зв'язки між великою кількістю змінних.

Крім того, сучасні економетричні методи широко використовуються для порівняння ефективності різноманітних економічних гіпотез та послідовного уточнення їх.

Етапи проведення економетричного аналізу

У стислому вигляді економетричний аналіз складається з таких етапів.

1. *Формулювання теорії чи гіпотези.*
2. *Розробка економетричної моделі для перевірки цієї теорії.*
3. *Оцінка параметрів обраної моделі.*
4. *Перевірка моделі, статистичні висновки.*
5. *Прогнозування на основі отриманої моделі.*
6. *Застосування моделі (для контролю тощо).*

Розглянемо більш детально деякі з цих етапів окремо на прикладі відомої кейнсіанської теорії споживання.

Формулювання теорії

Кейнс стверджував: "Фундаментальний психологічний закон... полягає в тому, що особа, як правило, збільшує споживання при зростанні доходів, але не на ту ж саму величину, на яку збільшується її доход". Коротше кажучи, Кейнс вважав, що гранична схильність до споживання (MPC) є величиною зміни доходу (від 0 до 1) при зміні споживання на одиницю (скажімо, на 1 долар).

Розробка моделі

Хоча Кейнс підкреслював позитивну залежність між споживанням та доходом, він не визначив чіткого функціонального зв'язку між цими двома параметрами. Для спрощення економетрист може запропонувати таку форму кейнсіанської функції споживання:

$$y = \beta_0 + \beta_1 x + \varepsilon \quad 0 < \beta_1 < 1, \quad (1)$$

де y — витрати на споживання;

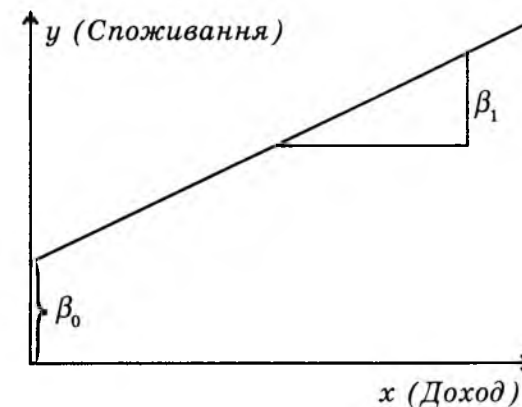
x — доход;

β_0, β_1 — параметри моделі;

ε — випадкова величина.

Параметр β_1 (тангенс кута нахилу) дорівнює граничній схильності до споживання MPC. Геометрично модель (1) подана на мал. 1.

Модель (1), що показує лінійну залежність споживання від доходу, є прикладом економетричної моделі. *Якщо модель складається лише з одного рівняння, як у попередньому прикладі, то вона називається простою, або моделлю з одним рівнянням, а якщо модель містить більше одного рівняння, то ми маємо справу зі складною моделлю, або моделлю багатьох рівнянь.*



Малюнок 1. Кейнсіанська функція споживання

Використання моделі для аналізу

Припустимо, що економіст-урядовець оцінює кейнсіанську функцію споживання і отримує такі результати:

$$\hat{y} = 5.0 + 0.7x, \quad (3)$$

де \hat{y} — витрати на споживання; x — доход у мільярдах гривень. Припустимо, уряд вважає, що витрати на суму 1060 мільярдів гривень утримують безробіття на низькому рівні, скажімо, 5%. Який рівень доходу гарантуватиме саме таку кількість витрат на споживання? Припускаємо, що модель (3) дійсна, тоді нескладні арифметичні підрахунки покажуть:

$$1060 = 5.0 + 0.7x \quad \text{чи} \quad x = 1055/0.7 = 1507.$$

Тобто доход на суму 1507 мільярдів гривень при даній MPC = 0.7 спонукає до витрат на суму 1060 мільярдів гривень.

Як показують обчислення на найпростішому умовному рівні, досліджувана модель може використовуватись для аналізу та досягнення різних політичних цілей. Завдяки відповідній податковій чи кредитно-грошовій політиці уряд може контролювати змінну x чи маніпулювати нею, щоб створити бажаний рівень доходу y .

Математичне та статистичне підґрунтя. Теоретична і прикладна економетрика

Економетрика поділяється на теоретичну та прикладну, кожна з яких у свою чергу ділиться на класичну та байєсівську. В цій книзі наголос робиться на класичному підході. Байєсівський підхід ми в цьому підручнику не розглядаємо, бо він, на жаль, не розрахований на початківців.

Теоретична економетрика стосується розвитку методів вимірювання економічних зв'язків, визначених економетричними моделями. У цьому аспекті економетрика базується на математичній статистиці. Наприклад, одним з найбільш використовуваних засобів у економетриці є метод найменших квадратів. Завдання теоретичної економетрики — детально записати припущення цього методу, його властивості та що відбувається з цими властивостями, коли одне чи більше припущень не виконуються.

У прикладній економетриці використовуються засоби теоретичної економетрики, наприклад, для вивчення функцій продуктивності, споживання, попиту та пропозиції тощо.

Ця книга в основному розглядає розвиток економетричних методів, їх припущень, застосувань та обмежень. Методи проілюстровано відповідними прикладами із різних сфер економіки та бізнесу.

Хоч матеріал цієї книги подано на елементарному рівні, ми вважаємо, що читач уже знайомий з основами статистики, особливо з фундаментальними концепціями оцінювання та перевірки гіпотез. Бажане також хоча б поверхове ознайомлення з методами обчислень, хоч це й не обов'язково. У підручнику є математичний додаток, тому при бажанні читач може поновити свої знання з необхідних розділів матричної алгебри та теорії ймовірностей.

Роль комп'ютера в сучасній економетриці

Сьогодні неможливо уявити собі економетричний аналіз без використання комп'ютера. Є декілька програм, що стосуються регресії та придатні для використання на мікрокомп'ютерах. З кожним днем кількість подібних програм збільшується.

У країнах СНД на сучасному ринку статистичних програм лідирують за якістю такі зарубіжні пакети, як STATGRAPHICS, SYSTAT, SPSS, SAS, BMDP, E.VIEWS та вітчизняні пакети МЕЗОЗАВР, САНИ, СИГАМД та ін.

Для загального уявлення про можливості деяких найвідоміших програм, які використовуються для розв'язку статистичних та економетричних задач, наведемо головні характеристики цих пакетів¹ (табл. 1).

Таблиця 1

№	Назва програми	Автор	Видавництво	Комп'ютер	Програми
I	II	III	IV	V	VI
1	BMDP Statistical Software	W.J. Dixon	University of California Press, Berkely, Calif.	IBM 360/370 (PC версія можлива)	1. Багатофакторна лінійна регресія 2. Крокова регресія 3. Всі можливі підмножини регресій 4. Поліноміальна регресія 5. Ступенева логістична регресія 6. Дисперсійний та коваріаційний аналіз
2	SAS (Statistical Analysis System)	Alice Allen Ray	SAS Institute, Inc. North Carolina	IBM 360/370 (PC версія можлива)	1. Лінійна регресія 2. Нелінійна регресія 3. RSQUARE-регресія 4. Ступенева регресія 5. Дисперсійний аналіз 6. Пробіт-модель
3	SPSS (Statistical Package for the Social Sciences)	Norman H.Nie	SPSS Inc, Chicago III	IBM 7090, 360 (PC версія можлива)	1. Багатофакторний регресійний аналіз 2. Процедура LOGLINEAR для категоричних ("фіктивних") змінних 3. Дисперсійний аналіз 4. Двовимірні графіки та діаграми розсіювання
4	RATS	Thomas Doan and Robert Litterman	VAR Econometrics, Inc., Minneapolis, Minn	PC, XT, AT computers	1. Звичайний метод найменших квадратів 2. Імовірнісні та логічні моделі 3. Двоступеневий метод найменших квадратів 4. Зважений метод найменших квадратів 5. Інструментальні змінні 6. Критерій Кокрена — Оркатта 7. Поліноміальні розподілені лаги

¹ Gujarati Damodar N. Basic Econometrics. — New York: McGraw-Hill Book Company, 1994.

I	II	III	IV	V	VI
5	Micro TSP	David M. Lilien	Mc Graw-Hill Book Company, New York	IBM PC, XT, AT, COMPAQ	<ol style="list-style-type: none"> 1. Проста регресія 2. Аналіз динамічних рядів 3. Прогнозування 4. Авторегресивні моделі та моделі ковзкого середнього 5. Симультаивні та імітаційні моделі 6. Використання електронних таблиць у поєднанні з мікро-TSP 7. Логічні та імовірнісні моделі
6	STAT GRAPHICS (Statistical Graphics System)	STSC corporation, USA		IBM PC, XT, AT	<ol style="list-style-type: none"> 1. Проста регресія 2. Аналіз динамічних рядів 3. Прогнозування 4. Авторегресивні моделі та моделі ковзкого середнього 5. Симультаивні та імітаційні моделі 6. Логічні та імовірнісні моделі

Вибір статистичного пакета для аналізу даних залежить від характеру завдань, що розглядаються, обсягу даних, які обробляються, наявного обладнання та кваліфікації користувача.

Додатком до цього підручника є "Економетрика: Практикум з використанням комп'ютера", в якому детально описано пакет прикладних програм STATGRAPHICS та інші пакети із статистики, що добре підходять до мікрокомп'ютерів. У книзі наведено приклади використання пакета STATGRAPHICS, оволодіння яким разом з іншими пакетами полегшить виконання обчислювальних завдань.

РОЗДІЛ 1. ЗАГАЛЬНІ ПРИНЦИПИ ПОБУДОВИ ЕКОНОМЕТРИЧНИХ МОДЕЛЕЙ. ЗВ'ЯЗОК ЕКОНОМЕТРИКИ З МАКРОЕКОНОМІКОЮ. ПРИКЛАДИ ПОБУДОВИ ЕКОНОМЕТРИЧНИХ МОДЕЛЕЙ

1.1. ЗВ'ЯЗОК ЕКОНОМЕТРИКИ З МАКРОЕКОНОМІКОЮ. ПРИКЛАДИ ЕКОНОМЕТРИЧНИХ МОДЕЛЕЙ

1.1.1. Модель валового національного продукту

Економетрика найтісніше пов'язана з аналізом та побудовою макроекономічних моделей економіки і макроекономічних агрегатів. Це й не дивно, бо аналіз найбільш агрегованих, спрощених моделей дає перше уявлення про стан або розвиток економічних процесів у суспільстві. Подальший розвиток моделей дозволяє глибше проникнути у сутність явища, ширше виявити взаємозв'язки та необхідні умови розвитку в тому чи іншому напрямку. Щоб зрозуміти основні принципи побудови економетричних моделей, почнемо з конкретних прикладів.

На початку розглянемо найпростішу модель валового національного продукту (ВНП), розрахованого за витратами для закритої економіки:

$$Y = C + I + G, \quad (1.1)$$

де Y — валовий національний продукт (ВНП); C — витрати споживачів; I — інвестиційні витрати; G — витрати уряду.

Для побудови економетричної моделі треба сформулювати гіпотези щодо визначення всіх типів витрат у рівнянні (1.1).

Найбільш поширеною гіпотезою щодо витрат споживачів є припущення залежності цього типу витрат від одержуваного прибутку $[(1 - \gamma)Y]$, норми сплачуваних податків (γ) та норми позичкового процента (r). Це можна умовно записати за допомогою функціонального зв'язку:

$$C = f((1 - \gamma)Y, r), \quad (1.2)$$

де γ — норма податків, яка для спрощеної моделі припускається однаковою для всіх економічних агентів; r — норма позичкового процента.

Виходячи з економічної теорії, на частинні похідні функції f можна накласти такі обмеження:

$$0 < f_Y < 1; f_r < 0, \quad (1.3)$$

де f_Y — частинна похідна функції f відносно Y , яка є не чим іншим, як граничною схильністю до споживання відносно одержуваного прибутку та додатною величиною, меншою за одиницю;

f_r — частинна похідна відносно норми процента. Збільшення норми позичкового процента негативно впливає на споживання, тому що збільшує накопичення, робить дорожчою купівлю товарів тривалого споживання у кредит, зменшує номінальну вартість облігацій, а все це в свою чергу зменшує реальну можливість споживання. Обмеження на частинні похідні допомагають при виборі конкретних функціональних залежностей, як це ми побачимо трохи далі.

Другу складову моделі (1.1) — інвестиції можна розглядати як функцію від зміни валового національного продукту та норми процента. Тобто маємо зв'язок:

$$I = \varphi(\Delta Y, r); \quad (1.4)$$

$$\varphi_Y > 0, \varphi_r < 0, \quad (1.5)$$

де ΔY — зміна ВВП.

У рівнянні (1.4) використано просте припущення, що спостережувані зміни ВВП певною мірою характеризують зміни прибутку, що залишається у населення, зростання якого в свою чергу позитивно впливає на інвестиції (перша нерівність (1.5)). Крім того, друга нерівність (1.5) відображає гіпотезу щодо оберненої залежності рівня інвестицій та норми позичкового процента.

Припущень, які відносяться до зміни урядових витрат, ми не розглядатимемо, тому отримаємо таку модель з трьома рівностями:

$$\begin{aligned} Y &= C + I + G; \\ C &= f((1 - \gamma)Y, r); \\ I &= \varphi(\Delta Y, r). \end{aligned} \quad (1.6)$$

Її слід доповнити гіпотезами щодо знака частинних похідних (1.3) та (1.5).

У моделі (1.6) треба одночасно оцінити три змінні C , I та Y . Їхні значення у свою чергу залежать від значень величин G , r та γ . Тому при конструюванні моделі (1.6) ми постаємо перед вибором, що є звичайним явищем в економетриці: як розглядати ці величини — як залежні змінні чи як незалежні.

Якщо ми будуватимемо функціональні залежності, що пов'язують також змінні G , r та γ з іншими величинами, то ми збільшимо початкову модель (1.6) ще на три рівності. Потім можна ще побудувати нові залежності, які у свою чергу описують нові зв'язки між величинами і т.д.

Де ж кінець? Цілком слушне запитання, яке, зрештою, ставить собі кожний, хто на практиці вирішує проблему побудови економетричних моделей економічних процесів та оцінки їхніх параметрів.

Усе залежить від ступеня спрощеності, від того, яку мету ми ставимо при розробці економетричних моделей.

У нашому простому випадку можна обмежитися трьома рівняннями, припустивши, що G , r та γ — незалежні величини.

Тепер розглянемо уважно модель (1.6) з припущеннями (1.3) та (1.5). Що нам дала економічна теорія? По-перше, можливість побудувати саму модель. По-друге, можливість розділити змінні на залежні та незалежні. По-третє, за чисто економічною логікою ми прийняли обмеження щодо знаків частинних похідних функцій зв'язку.

Постає запитання, чи є модель (1.6) достатньою для детального аналізу економічної реальності, яку вона відображає? *Звісно, ні.*

Оскільки функціональні зв'язки (1.2), (1.4) із гіпотезами (1.3), (1.5) мають лише “якісний вигляд”, то найбільше, що може дати така модель, це якісне пояснення деяких процесів, що відбуваються у реальності. Те, що дає класична макроекономічна теорія. Ні про детальний аналіз реальності, ні навіть про кількісні залежності між деякими величинами за моделлю (1.6) не йдеться. *І це природно: що у модель закладено, те й маємо. Модель (1.6) являє собою якісний рівень опису економічної реальності та є нестохастичним аналогом економетричної моделі.*

1.1.2. Класична модель економіки

Класична модель відображає головні доктрини класичної теорії. Розглянемо головні досягнення класичних економістів.

Ранні економісти розглядали таку економічну проблему: в ринковій економіці окремі особистості формують свої економічні плани незалежно один від одного, причому ці плани можна нормально виконати. Яким чином це можливо? Відповідь була така: *“Ціновий механізм координує індивідуальні економічні плани. Відповідно до цього механізму мо-*

дель виробництва має тенденцію регулювати модель попиту. Ранні економісти вважали, що зміна ціни є наслідком дії ринкового механізму, і якщо це справедливо для будь-якого окремого ринку, то це справедливо також і для всіх інших ринків.

Враховуючи наведений вище механізм, класичні та неокласичні економісти розглядали також **проблеми регулювання ринку**. Тезовий зміст їхнього аналізу: **ринкова економіка веде до загальної рівноваги країни**, у зв'язку з цим загальна рівновага є центром гравітації ринкової економіки.

Класичні та неокласичні економісти не відкидали можливості тимчасових криз, вони тільки стверджували, що перехід до загальної рівноваги відбувається за довгий період часу — від п'яти до тридцяти років. Далі вони розглядали загальну рівновагу як оптимальний стан для суспільства. **Ранні економісти рекомендували економічний порядок, який гарантує економічну свободу мешканцям: вони захищали ринкову директивну економіку.**

Крім аналізу цінового механізму, який ми називаємо ціною або ціннісною теорією, другим досягненням класичних економістів було **відкриття вуалі цін**. Вони стверджували: “Гроші і добробут — це зовсім різні речі”. Соціальний добробут складається із продукції за рік або існуючого потоку товарів, у той час, коли гроші є просто засобом обміну. Гроші потрібні не заради самих грошей, а тільки заради товарів, які можна купити на них. Тому **гроші є “вуаллю” реальних дій.**

Перебільшення значення цієї точки зору призвело до **макроекономічної дихотомії**, тобто припущення, що в ринковій економіці номінальна та реальна величини зовсім незалежні одна від однієї. Наприклад, збільшення кількості грошей не вносить будь-яких змін у реальний добробут, а дає тільки пропорційне зростання в цінах. Відповідно аналіз був поділений на ціннісну та монетарну теорії. **Ціннісна теорія** розглядає реальні величини, а також відносні ціни. **Монетарна ж теорія** має справу з чистими номінальними цінами.

Чітке відображення доктрини класичної теорії дає **класична модель**. У моделі робляться деякі припущення (для її спрощення).

Перше припущення. В економіці, що розглядається, є тільки два види дійових осіб — **фірми та домашні господарства**, об'єднані в два сектори. Кожен із секторів виконує три види діяльності:

фірми

- виробництво товарів;
- попит на робочу силу;
- здійснення інвестицій;

домашні господарства:

- споживання товарів;
- пропозиція здатності до праці;
- створення заощаджень.

Друге припущення. Розглядається тільки один однорідний товар,

кількість якого позначається через змінну Y . За рахунок попереднього припущення Y є реальним доходом економіки в цілому, який вимірюється в одиницях товару. В теорії очікувань (ex-ante аналізі) слід розрізняти три значення Y :

Y^S — сумарна кількість очікуваного виробництва;

Y^D — сумарна кількість очікуваного попиту на товари;

Y — сума трудових доходів, доходів від цінних паперів (або депозитів) і нетрудових доходів.

Модель розглядає чотири ринки разом:

— **ринок товарів**, де пропозиція товарів Y^S збігається з попитом споживачів C та попитом на інвестиції I ;

— **ринок праці**, де збігаються пропозиція праці N^S з попитом N^D ;

— **ринок капіталу**, де збігаються пропозиція капіталу S (заощадження) з попитом на капітал $I \equiv \Delta B^S / P$ (де B^S — номінальний потік облігацій, який вимірюється в грошових одиницях);

— **ринок грошей**.

Припускається, що три перші ринки є досконало конкурентними, тому на кожному з них встановлюється єдина ціна, тобто:

— **рівень цін (P)**, який означає грошову ціну однорідного товару;

— **номінальна заробітна плата (w)** як грошова ціна праці;

— **ставка процента (i)** як ціна капіталу.

Доки ми абстрагуємося від грошей, головними є не рівень цін P та номінальна заробітна плата w , а тільки їх відношення, тобто **реальна заробітна плата w/P** .

Реальна заробітна плата означає число одиниць товару, сплачених за одну робочу годину.

Класична модель включає:

— **функції попиту та пропозиції праці, тобто**

$$N^d\left(\frac{w}{P}\right); N^s\left(\frac{w}{P}\right);$$

— **виробничу функцію $Y = f(N)$;**

— **функції заощадження та інвестицій $S = S(i)$, $I = I(i)$;**

— **рівняння Кембріджа $M = k \cdot P \cdot Y$,**

де k (Cambridge k) — середня тривалість перебування готівкових коштів на руках у індивідумів — від отримання людьми грошових доходів до витрат цих коштів на товари, послуги чи заощадження, тобто певний період часу, протягом якого люди утримують кошти на руках).

Додаючи до цього тотожність $W = (W/P)P$, ми отримуємо одночасні рівняння для класичної моделі:

$$N^d\left(\frac{w}{P}\right) = N^* = N^s\left(\frac{w}{P}\right) \rightarrow N^*, \left(\frac{w}{P}\right)^* ; \quad (1.7)$$

$$Y = f(N) \rightarrow Y^* ; \quad (1.8)$$

$$S(i) = I(i) \rightarrow i^* ; \quad (1.9)$$

$$M = k \cdot P \cdot Y \rightarrow P^* ; \quad (1.10)$$

$$w = \left(\frac{w}{P}\right) \cdot P \rightarrow w^* . \quad (1.11)$$

Рівняння (1.7) — (1.11) — це 6 нестохастичних аналогів симультативних рівнянь. Слід відзначити, що (1.7) містить у собі два рівняння.

Рівняння (1.7) являє собою умову рівноваги на ринку праці, що задає повну зайнятість (N^*), а також реальну зарплату $(w/P)^*$ в рівновазі.

Умова (1.8) — це виробнича функція (з виключенням змінної K — капіталу і з урахуванням залежності виробництва лише від змінної зайнятості). Також можна було б записати функцію пропозиції товарів, але формула (1.8) повніше свідчить про те, що пропозиція товарів впливає безпосередньо з рівня зайнятості N^* .

Рівняння (1.9) — це умова, що стосується безпосередньо ринку капіталу, визначає природну ставку процента i^* . Звичайно, передбачається певна кількість заощаджень та інвестицій.

Умова (1.10) — рівняння Кембріджа. Оскільки кількість грошей та швидкість їх обігу припускаються заданими, рівень цін P^* у рівновазі визначається реальним випуском Y^* .

Рівняння (1.11) — це, нарешті, чисто формальна тотожність. Вона стверджує, що певна номінальна заробітна плата W^* визначається реальною зарплатою $(W/P)^*$ у рівновазі і рівнем цін P^* у рівновазі.

Крім того, модель пояснює класичну *дихотомію*: реальний сектор економіки зображується рівняннями (1.7) — (1.9), де всі реальні величини визначені. З рівнянь (1.10) і (1.11) рівень цін та номінальний рівень зарплати визначаються як чисто номінальні змінні, які не впливають на реальні змінні.

Така модель добре підходить для демонстрації сутності доктрини, але ці рівняння не слід сприймати як природні закони, вони тільки дають стисле пояснення деяких незаперечно важливих зв'язків.

1.1.3. Повна кейнсіанська модель

Так звана повна кейнсіанська модель являє собою модель загальної економіки. Вона включає в себе неокласичну модель ринку праці, виробничу функцію та IS-LM-модель як модель сектора споживання в економіці.

У позначеннях:

Y — валовий національний продукт;

N — кількість робочих місць;

W — заробітна плата;

P — індекс цін;

r — норма процента;

$\left(\frac{W}{P}\right) = w_q$ — норма реальної заробітної плати;

$N^d\left(\frac{W}{P}\right)$ — функція попиту на ринку праці;

$N^s\left(\frac{W}{P}\right)$ — функція пропозиції на ринку праці;

M — загальна грошова маса.

Повна кейнсіанська модель, як і класична модель, складається з шести рівнянь:

$$N^d\left(\frac{W}{P}\right) = N = N^s\left(\frac{W}{P}\right); \quad (1.12)$$

$$Y = f(N); \quad (1.13)$$

$$S(Y) = I(r); \quad (1.14)$$

$$L(Y, r) = \frac{M}{P}; \quad (1.15)$$

$$W = \left(\frac{W}{P}\right) \times P. \quad (1.16)$$

Рівняння (1.12) є подвійним, тобто це два рівняння ринку праці, за допомогою яких визначаються змінні зайнятості (кількість робочих місць) N^* та норма реальної заробітної плати.

Рівняння (1.13) містить виробничу функцію, яка за умов повної зайнятості на ринку праці (N^*) дає значення рівноважного валового національного продукту (Y^*).

Рівняння (1.14) та (1.15) представляють IS-LM-модель. В цій моделі рівень цін (P^*) є залежною змінною, і він, як і норма процента (r^*), визначається з поданих рівнянь.

Рівняння (1.16) визначає номінальну заробітну плату w^* .

Отже, послідовний розв'язок цих рівнянь дав би змогу послідовно визначити шість змінних N^* , r^* , w^* , $(W/P^*)^*$, Y^* , P^* .

Але для того, щоб знайти значення цих змінних, ми повинні спочатку встановити явний вигляд функцій зв'язку $N^d()$, $N^s()$, $f()$, $S()$, $I()$, $L()$ та оцінити їхні параметри, вводючи випадкові величини у праві частини рівнянь.

Для переходу до кількісного рівня опису розглянутих вище моделей і потрібен *економетричний підхід*, який оперує кількісними значеннями всіх величин, задіяних у моделі, та дозволяє не тільки побудувати явний вигляд функцій зв'язку, а й глибоко проаналізувати сутність самої моделі.

1.2. РОЛЬ ЕКОНОМЕТРИКИ В ЕКОНОМІЧНИХ ДОСЛІДЖЕННЯХ

Як показано вище, економічна теорія може дати тільки якісний рівень опису макро- та мікроекономічних моделей, допомогти виявити загальний характер зв'язку між факторами та зробити деякі припущення щодо знаків перших або других похідних.

Щоб встановити явний вигляд функції зв'язку, спочатку необхідно вибрати тип функції, що відповідає нашим припущенням. Пояснюючи це, знову повернемося до найпростішої моделі — моделі валового національного продукту (1.6). Для спрощення абстрагуємося від норми позичкового процента, тобто припустимо, що $r = const$. Розглянемо в моделі (1.6) сектор споживання (1.2;1.3). Як ми вже бачили, економічна теорія може допомогти виявити загальний характер зв'язку між факторами та зробити деякі припущення щодо знаків перших або других похідних. Але ж декілька типів функцій можуть одноразово відповідати таким припущенням. Яку ж із них вибрати? Наприклад, нехай

$$R = (1 - \tau)Y, \quad (1.17)$$

де R — прибуток, який залишається у розпорядженні держави. Тоді макроекономічний показник споживання може бути поданий за допомогою кількох функцій.

$$\hat{C} = a_0 + a_1R; \quad (1.18)$$

$$\hat{C} = AR^{a_1}; \quad (1.19)$$

$$\hat{C} = a_0 - a_1R^{-1}. \quad (1.20)$$

Усі функції (1.18) — (1.19) є зростаючими щодо R , до того ж їм відповідає додатна та нижча за одиницю гранична схильність до споживання (за умов, що параметри a_0 , a_1 та A відповідатимуть деяким вимогам (ми, звичайно, абстрагуємося від норми процента). Отже, як бачимо, у даному випадку три функції відповідають вимогам, розглянутим раніше, щодо опису сектора споживання (див.1.2). Але ж, маючи спільні властивості, ці функції описують різні якісні процеси споживання. Якщо взяти функцію (1.18), то вона описує ситуацію, за якої додатковий дохід у розмірі 100 гривень викликає завжди однакове зростання витрат споживання, тоді як у випадках (1.19) — (1.20) гранична схильність до споживання зменшується в міру зростання доходу. Крім того, у випадку (1.19) споживання нескінченно зростає разом з доходом, тоді як у (1.20) при дуже великому доході споживання прямує до рівня задоволення, поданого константою a_0 .

Як бачимо, економічна теорія не може точно виявити тип функціонального зв'язку. Крім того, до виявлення такого зв'язку ми маємо відповісти ще на деякі запитання. Наприклад, чи повинні ми враховувати сезонні коливання у попередньому випадку? Чи є кілька типів “норм процента”? Чи повинні ми розглядати “репрезентативну” норму процента або вводити комбінацію різних норм процентів? Чи буде норма процента однією і тією ж для функції споживання та інвестицій і т.ін.

Можна виділити 5 основних завдань, які розв'язує економетрика.

По-перше, модель має бути специфікована, тобто треба, щоб усі функціональні зв'язки входили до неї у явному вигляді. До цього економетрика може дійти шляхом від простого до складного: почавши з найпростіших функцій, вводити та перевіряти різні гіпотези і поступово ускладнювати характер функціональних зв'язків виходячи з реальних даних.

По-друге, завданням економетрики є вибір означення та виміру змінних, які входять до моделі.

По-третьє, необхідно оцінити всі невідомі параметри моделей та розрахувати інтервали довіри (інтервали, до яких із заданим ступенем імовірності потраплятиме обчислювана величина).

По-четверте, необхідно оцінити якість побудованих моделей за допомогою різних тестів та критеріїв. Це допомагає остаточно вирішити питання, чи треба змінювати початково обрану модель, та деякі теоретичні припущення. Якщо така зміна необхідна, то треба проводити нові розрахунки і нове тестування.

По-п'яте, маючи остаточну модель, необхідно провести глибокий аналіз результатів, які планується використовувати на практиці для прийняття рішень.

1.3. ІНФОРМАЦІЙНА БАЗА ЕКОНОМЕТРИЧНИХ МОДЕЛЕЙ

Економетрика як наука насамперед ставить собі за мету обґрунтований аналіз економічних явищ на базі математико-статистичних методів. Такий аналіз потребує наявності числових характеристик економічних явищ та можливості вимірювання їх. При вимірюванні кількісних ознак можуть бути отримані два типи рядів даних — динамічні та варіаційні. Ці ряди здебільшого і становлять інформаційну базу економетричних моделей.

1.3.1. Динамічні ряди та їхні характеристики

Динамічним рядом називається послідовність спостережень за процесом або явищем у рівновіддалені проміжки часу.

Позначимо через x_i значення деякої ознаки економічного процесу або явища в i -й проміжок часу. Тоді, вимірюючи значення цієї ознаки в рівновіддалені проміжки часу, отримуємо динамічний ряд: $x_1, x_2, \dots, x_p, \dots, x_n$. Окремі значення ознаки, які відносяться до певних проміжків часу, ще називають рівнями динамічного ряду. Наприклад, x_i — це значення ознаки, що вивчається у i -й проміжок часу або i -й рівень динамічного ряду.

Для того, щоб ці динамічні ряди можна було використовувати як інформаційну базу для побудови регресійних моделей, необхідно, щоб усі їхні рівні можна було порівняти. Наприклад, дані можуть стосуватися території, кордони якої з часом змінювалися. Непорівнянність такого роду усувається перерахунком даних із врахуванням змінених кордонів.

Непорівнянність може бути значною, коли показники, що розглядаються, підвладні сезонним або іншим періодичним коливанням (ціна на сільськогосподарську продукцію, тарифи перевезень та ін.). Такі дані необхідно проаналізувати за методами сезонної декомпозиції.

Непорівнянними є також дані, подані у різному масштабі виміру. Їх треба спочатку перевести в однакові одиниці.

При формуванні динамічних рядів можуть бути ускладнення, пов'язані з браком необхідних даних. Один з найбільш поширених засобів подолання цього — виявлення закономірностей, яким підпорядковується динамічний ряд, та екстраполювання або інтерполювання його недостатніх рівнів.

Після того, як динамічний ряд всебічно проаналізовано на достовірність та порівнянність даних, його можна використати як вхідну інформацію для побудови моделей.

Середні характеристики динамічного ряду

Узагальненими характеристиками динамічного ряду є середня хронологічна, середній абсолютний приріст, середній темп зростання та приросту.

Нехай ми маємо динамічний ряд з n спостережень: $x_1, x_2, \dots, x_p, \dots, x_n$.

Середня хронологічна показує, яким рівнем у середньому характеризується даний динамічний ряд і розраховується за формулою:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n},$$

де \bar{x} — середня хронологічна; x_i — i -й рівень динамічного ряду; n — кількість спостережень (рівнів).

Середні хронологічні корисні для порівняльного аналізу двох або кількох динамічних рядів (наприклад, порівняння середніх рівнів урожайності для різних регіонів країни, порівняння середньої заробітної плати у промисловості та в цілому по країні і т.ін.).

Характеристику зміни динамічного ряду та швидкість цієї зміни в середньому дають середній абсолютний приріст та середній темп зростання (приросту).

Середній абсолютний приріст показує, як швидко змінюється кінцевий рівень ряду відносно початкового:

$$\overline{\Delta x} = \frac{x_n - x_1}{n - 1},$$

де $\overline{\Delta x}$ — середній абсолютний приріст; x_n, x_1 — кінцевий та початковий рівень ряду.

Середній коефіцієнт зростання характеризує середню швидкість зміни економічного процесу або явища і розраховується за формулою:

$$\overline{k_p} = \sqrt[n-1]{\frac{x_n}{x_1}},$$

де $\overline{k_p}$ — середній коефіцієнт зростання.

Середній коефіцієнт приросту відрізняється від середнього коефіцієнта зростання на одиницю:

$$\overline{k_{np}} = \sqrt[n-1]{\frac{x_n}{x_1}} - 1,$$

де $\overline{k_{np}}$ — середній коефіцієнт приросту.

Середні коефіцієнти зростання та приросту, виражені у відсотках, називаються відповідно середнім темпом зростання ($\overline{T_p}$) та середнім темпом приросту ($\overline{T_{np}}$).

Приклад. Розглянемо поквартальну заробітну плату (ЗП) у державному секторі України за 1992 — 1994 рр. (табл. 1.1)*.

Таблиця 1.1

№ п/п	Квартал	ЗП у державному секторі (без премій), млрд.крб — (x_i)	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	1992 (1)	115	- 12771.5	163111212
2	1992 (2)	238	- 12648.5	159984552
3	1992 (3)	362	- 12524.5	156863100
4	1992 (4)	761	- 12125.5	147027750
5	1993 (1)	1223	- 11663.5	136037232
6	1993 (2)	2220	- 10666.5	113774222
7	1993 (3)	6366	- 6520.5	42516920
8	1993 (4)	21224	8335.5	69513906
9	1994 (1)	42043	29156.5	850101492
10	1994 (2)	54313	41426.5	1716154902

* Джерело: Тенденції української економіки // Щомісячний бюлетень / Європейський центр макроекономічного аналізу України. — 1994. — Серпень.

Проведемо підрахунки за цими даними й отримаємо такі середні характеристики.

Середня квартальна заробітна плата у державному секторі України в 1992 — 1994 рр. дорівнювала:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{128865}{10} = 12886.5.$$

Середній абсолютний приріст заробітної плати дорівнював:

$$\Delta \bar{x} = \frac{54313 - 115}{9} = 6022.$$

Середні темпи зростання та приросту відповідно становили:

$$\bar{T}_p = \sqrt[9]{\frac{54313}{115}} \times 100\% = 1.982 \times 100\% = 198.2\%;$$

$$\bar{T}_{np} = \left(\sqrt[9]{\frac{54313}{115}} - 1 \right) \times 100\% = 98.2\%.$$

Відхилення від середнього вимірює дещо штучна величина — *дисперсія*. Дисперсія показує середню суму квадратів відхилень членів ряду від свого середнього і позначається σ^2 , або $\text{var}(x)$:

$$\text{var}(x) = \sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2,$$

де x — середнє значення динамічного ряду; n — кількість спостережень.

У підрахунках дисперсії використовується середня сума квадратів відхилень, тому що середня сума відхилень дорівнює нулеві. Справді:

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}) = \bar{x} - \frac{n\bar{x}}{n} = 0.$$

Для того, щоб дисперсію можна було порівняти з середніми характеристиками, вводиться середнє квадратичне відхилення:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}.$$

Дисперсія цікава також тим, що яким би не був розподіл величини x , як мінімум 75% спостережень знаходяться між $(\bar{x} - 2\sigma)$ та $(\bar{x} + 2\sigma)$.

Для порівнювання ступеня коливання різнорідних показників у відсотках запроваджено інший показник — коефіцієнт варіації, який розраховується за формулою:

$$V = \frac{\sigma}{\bar{x}} \times 100\%,$$

де V — коефіцієнт варіації; \bar{x} — середнє значення ряду; σ — середнє квадратичне відхилення.

Для нашого прикладу дисперсія дорівнює: $\sigma^2 = 355508528.8$; середнє квадратичне відхилення $\sigma = 59624$; коефіцієнт варіації $V = 426.68\%$.

1.3.2. Варіаційні ряди та їхні характеристики

Іншим джерелом побудови економетричних моделей служать *варіаційні ряди*.

Варіаційні ряди — це ряди даних, які показують кількісну міру певної ознаки у всіх об'єктів однієї сукупності, наприклад, заробітна плата у викладачів однієї кафедри, вік студентів першого курсу і т.ін.

Нехай ми маємо варіаційний ряд:

$$x_1, x_2, \dots, x_n,$$

де x — числа, які показують зміну (варіацію) ознаки, що вивчається;
 i — номер варіанти ($i = 1, 2, \dots, n$).

Приклад. При обстеженні студентів 1-го курсу за віком було зафіксовано такі дані:

17,18,18,18,18,19,20,20,20,21,21,21,17,18,18,18,19,20,20,20,21,21,21,24.

Якщо впорядкувати ці дані у зростаючому або спадному порядку, то отримаємо ранжований ряд. Числа, які показують, скільки разів (як часто) зустрічаються окремі значення варіант, називаються частотами.

Позначимо частоту i -ї варіанти x_i через n_i , тоді ранжований дискретний варіаційний ряд запишеться у вигляді (табл. 1.2)

Таблиця 1.2

Варіаційний ряд у загальному вигляді			Варіаційний ряд для прикладу (вік студентів 1-го курсу)		
номер варіанти	значення варіанти	частота варіанти	номер варіанти	значення варіанти	частота варіанти
1	x_1	n_1	1	17	1
2	x_2	n_2	2	18	4
.	.	.	3	19	1
.	.	.	4	20	3
k	x_k	n_k	5	21	4

Для варіаційного ряду є також дві групи характеристик: *міри рівнів (середні) та міри розсіяння*.

Найбільш поширеними середніми характеристиками для варіаційних рядів є середня арифметична, медіана і мода. Вони розраховуються в залежності від того, який варіаційний ряд ми маємо — дискретний чи інтервальний (*).

Для дискретного варіаційного ряду *середня арифметична* розраховується за формулою:

$$\bar{x} = \frac{\sum_{i=1}^k x_i n_i}{\sum_{i=1}^k n_i},$$

де \bar{x} — середня арифметична; x_i — i -та варіанта; n_i — частота i -ї варіанти.

Якщо у ряді є нетипові дані, то це впливає на середню величину, яка вже не відображає середньої тенденції. У таких випадках краще використовувати медіану.

Медіаною називається таке значення ознаки, що вивчається, яке припадає на середину варіаційного ряду. При знаходженні медіани можливі два випадки: кількість членів ряду парна ($n = 2k$) та непарна ($n = 2k+1$).

У другому випадку знайти медіану дуже просто. За означенням вона дорівнює:

$$M_e = x_{k+1},$$

де M_e — медіана; x_{k+1} — значення $(k+1)$ -го члена варіаційного ряду.

У першому випадку:

$$M_e = \frac{(x_k + x_{k+1})}{2},$$

де M_e — медіана; x_k, x_{k+1} — відповідно k та $(k+1)$ варіанти.

Модою називається варіанта, яка найчастіше зустрічається в даному варіаційному ряду. Іншими словами, для дискретного ряду мода дорівнює варіанті з найбільшою частотою.

Середніх характеристик часом буває недостатньо для характеристики як варіаційного ряду, так і динамічного. Інколи два ряди можуть мати однакові середнє, медіану та моду, але по-різному можуть бути згруповані навколо середнього. Розглянемо це на дещо штучному прикладі (табл 1.3).

Таблиця 1.3

i	n_i	x_i	$n_i x_i$	i	n_i	x_i	$n_i x_i$
1	5	99	495	1	5	1	5
2	10	100	1000	2	10	100	1000
3	5	101	505	3	5	199	995
Σ	$N=20$	—	2000	Σ	$N=20$	—	2000
Σ/N	1	—	100	Σ/N	1	—	100

* Інтервальний варіаційний ряд для спрощення ми не розглядатимемо.

У наведеній таблиці два варіаційні ряди мають одні й ті самі середні характеристики, але по-різному згруповані навколо середнього.

Для того щоб охарактеризувати розсіяння навколо середнього, вводиться дисперсія (σ^2):

$$\sigma^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2 n_i}{\sum_{i=1}^k n_i}$$

Якщо дисперсію зручно подавати в тих самих одиницях виміру, що й варіанти, то використовують середнє квадратичне відхилення (σ):

$$\sigma = \sqrt{\frac{\sum_{i=1}^k (x_i - \bar{x})^2 n_i}{\sum_{i=1}^k n_i}}$$

Для варіаційних рядів розраховується також коефіцієнт варіації (V):

$$V = \frac{\sigma}{\bar{x}} \times 100\%$$

Коефіцієнт варіації дає змогу:

порівняти варіацію однієї і тієї самої ознаки у різних групах об'єктів; виявити ступінь відмінності однієї ознаки в одній групі об'єктів за різні проміжки часу; порівняти варіацію різних ознак в однакових групах об'єктів.

Повернемося до нашого умовного прикладу про обстеження студентів 1-го курсу за віком. Розрахуємо для отриманого варіаційного ряду всі характеристики. Дані для розрахунків наведені у табл.1.4.

Таблиця 1.4

Номер варіанти	Значення варіанти x_i	Частота варіанти n_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \cdot n_i$
1	17	1	- 2.38	5.66	5.66
2	18	4	- 1.38	1.90	7.61
3	19	1	- 0.38	0.14	0.14
4	20	3	0.62	0.38	1.15
5	21	4	1.62	2.62	10.49

Середній вік студентів 1-го курсу дорівнює: $\bar{x} = 19.38$.

Медіана відповідно дорівнює:

$$M_e = x_{20} = 20,$$

а мода:

$$M_o = x_2 = x_5 = (21;18).$$

Дисперсія варіаційного ряду $\sigma^2=1.927$, середнє квадратичне відхилення $\sigma = 1.38$, а варіація $V = 7.16\%$.

РОЗДІЛ 2. ПРОСТА ВИБІРКОВА ЛІНІЙНА РЕГРЕСІЯ

2.1. ЗАГАЛЬНЕ ПОНЯТТЯ ПРО ЛІНІЙНУ РЕГРЕСІЮ

Прості лінійні регресійні моделі встановлюють лінійну залежність між двома змінними, наприклад витратами на відпустку та складом родини; витратами на рекламу та обсягом продукції, що випускається, витратами на споживання та валовим національним продуктом (ВНП); зміною ВНП залежно від часу і т.ін.

При цьому одна із змінних вважається залежною змінною (y) та розглядається як функція від незалежної змінної (x).

У загальному вигляді проста вибіркова регресійна модель запишеться так:

$$y = b_0 + b_1 x + e, \quad (2.1)$$

де y — вектор спостережень за залежною змінною; $y = \{y_1, y_2, \dots, y_n\}$;

x — вектор спостережень за незалежною змінною; $x = \{x_1, x_2, \dots, x_n\}$;

b_0, b_1 — невідомі параметри регресійної моделі;

e — вектор випадкових величин (помилки); $e = \{e_1, e_2, \dots, e_n\}$.

Регресійна модель називається лінійною, якщо вона лінійна за своїми параметрами. Отже, модель (2.1) є лінійною регресійною моделлю. Її ще можна трактувати і як пряму на площині, де b_0 — перетин з віссю ординат, а b_1 — нахил (звичайно, якщо абстрагуватися від випадкової величини e).

2.2. ОЦІНКА ПАРАМЕТРІВ ЛІНІЙНОЇ РЕГРЕСІЇ ЗА ДОПОМОГОЮ МЕТОДУ НАЙМЕНШИХ КВАДРАТІВ

Щоб мати явний вид залежності, необхідно знайти (оцінити) невідомі параметри b_0, b_1 цієї моделі. Як це зробити? Яким критерієм краще користуватися? Щоб відповісти на ці запитання, розглянемо спочатку приклад.

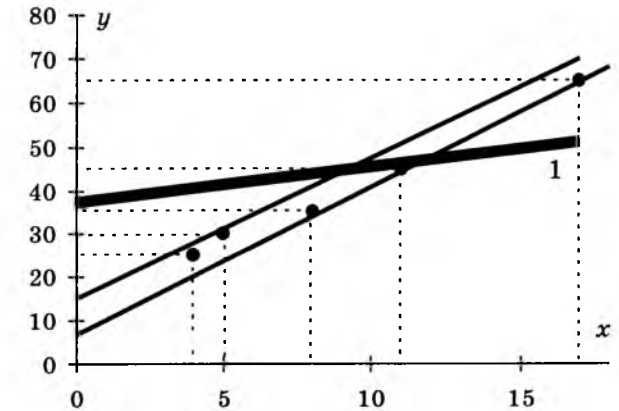
Приклад. Бюро економічного аналізу фабрики “Світоч” оцінює ефективність відділу маркетингу з продажу цукерок. Для такої оцінки вони мають досвід праці у 5 географічних зонах з майже однаковими умовами (потенційні клієнти, ставлення до товарного знака і т. ін.). У цих зонах вони зафіксували протягом однакового періоду обсяги продажів (млн. коробок), витрати (млн. грн.) фірми та просування товару на ринку. Дані наведені в табл.2.1.

Реальні спостереження y_i зобразимо точками у системі координат (X, Y) (мал.2.1).

Візуально можна припустити, що між даними є лінійна залежність, тобто їх можна апроксимувати прямою лінією.

Таблиця 2.1

I	y_i	x_i
1	25	5
2	30	6
3	35	9
4	45	12
5	65	18



Малюнок 2.1. Залежність між обсягами продажу продукції та витратами на рекламу

Взагалі, існує необмежена кількість прямих $y = b_0 + b_1 x$, які можна провести через множину спостережуваних точок. Яку ж із них вибрати?

Щоб це визначити, потрібно мати у розпорядженні певний критерій, що дозволяв би вибрати з множини можливих прямих “найкращу” з точки зору даного критерію. Найпоширенішим є критерій мінімізації суми квадратів відхилень. На мал. 2.1, наприклад, пряма (1), як і інші, розташована таким чином, що деякі точки знаходяться вище, деякі нижче цієї прямої, на основі чого можна встановити відхилення (помилки) відносно цієї прямої:

$$e_i = y_i - \hat{y}_i = y_i - b_0 - b_1 x_i, \quad i = \overline{1, n}, \quad (2.2)$$

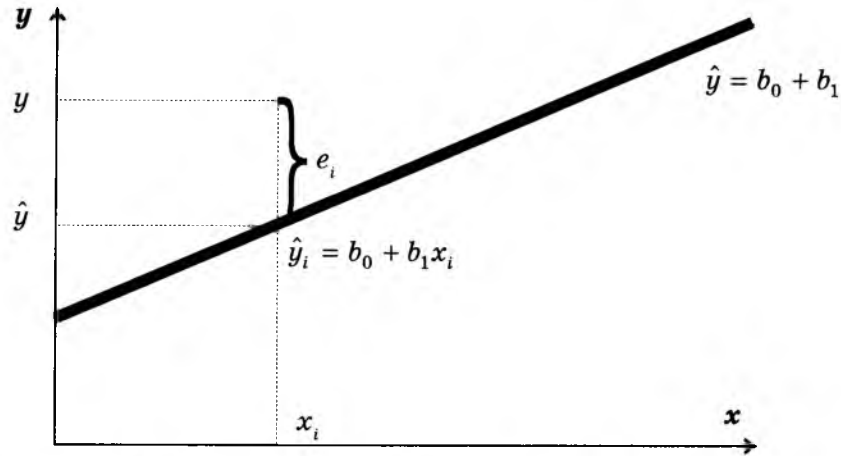
де \hat{y}_i — i -та точка на прямій, яка відповідає значенню x_i (див. мал. 2.2).

Відхилення, або помилки, ще інколи називають залишками. Логічно, що треба проводити пряму таким чином, щоб сума квадратів помилок була мінімальною. В цьому і полягає *критерій найменших квадратів*: невідомі параметри b_0 та b_1 визначаються таким чином, щоб мінімізувати $\sum_{i=1}^n e_i^2$.

Справді, за критерієм маємо

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2 = f(b_0, b_1) \rightarrow \min. \quad (2.3)$$

Визначимо значення b_0 та b_1 , які мінімізують вираз (2.3). Мінімум функції (2.3) досягається за необхідних умов, коли перші похідні дорівнюють нулеві, тобто



Малюнок 2.2. Відхилення теоретичних значень від фактичних

$$\frac{\partial \left(\sum_{i=1}^n e_i^2 \right)}{\partial b_0} = \frac{\partial f(b_0, b_1)}{\partial b_0} = 0; \quad \frac{\partial \left(\sum_{i=1}^n e_i^2 \right)}{\partial b_1} = \frac{\partial f(b_0, b_1)}{\partial b_1} = 0; \quad (2.4)$$

$$\frac{\partial \left(\sum_{i=1}^n e_i^2 \right)}{\partial b_0} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) = 0; \quad (2.5)$$

$$\frac{\partial \left(\sum_{i=1}^n e_i^2 \right)}{\partial b_1} = -2 \sum_{i=1}^n x_i (y_i - b_0 - b_1 x_i),$$

звідки отримуємо систему лінійних рівнянь:

$$\begin{cases} \sum_{i=1}^n y_i = n b_0 + b_1 \sum_{i=1}^n x_i; \\ \sum_{i=1}^n y_i x_i = b_0 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n x_i^2, \end{cases} \quad (2.6)$$

яка називається нормальною.

Розв'язок (2.6) відносно нахилу прямої (невідомо b_1) дає

$$b_1 = \frac{\sum_{i=1}^n x_i y_i - \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n}}{\sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2}. \quad (2.7)$$

З метою спрощення виразу для b_1 чисельник та знаменник виразу 2.7 помножимо на $1/n$.

Отримуємо:

$$b_1 = \frac{\frac{1}{n} \left(\sum_{i=1}^n x_i y_i \right) - \bar{x} \bar{y}}{\frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2}, \quad (2.8)$$

$$\text{де } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i; \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i.$$

Вираз (2.8) можна записати ще таким чином:

$$b_1 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\text{cov}(x, y)}{\text{var}(x)}. \quad (2.9)$$

Справді,

$$\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - \bar{x} \sum_{i=1}^n y_i - \bar{y} \sum_{i=1}^n x_i + n \bar{x} \bar{y} = \sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}; \quad (2.10)$$

$$\sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 = \sum_{i=1}^n x_i^2 - n \bar{x}^2. \quad (2.11)$$

Чисельник (2.9) є не що інше, як коефіцієнт коваріації між x та y . За означенням, коефіцієнт коваріації між двома змінними x та y визначається за формулою:

$$\text{cov}(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}). \quad (2.12)$$

Знаменник (2.9) є дисперсією величини x , тобто

$$\text{var}(x) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (2.13)$$

Отже, кут нахилу прямої регресії можна встановити як за формулою (2.7), так і за формулами (2.8) та (2.9).

Для визначення параметра b_0 повернемося до (2.5). Маємо:

$$\frac{\partial \left(\sum_{i=1}^n e_i^2 \right)}{\partial b_0} = \sum_{i=1}^n \frac{\partial}{\partial b_0} \left[(y_i - b_0 - b_1 x_i)^2 \right] = \sum_{i=1}^n (y_i - b_0 - b_1 x_i) = 0. \quad (2.14)$$

Вираз (2.14) дає нам, по-перше, підтвердження того, що сума помилок дорівнює нулеві. Справді,

$$y_i - b_0 - b_1 x_i = e_i \Rightarrow \sum_{i=1}^n e_i = 0; \quad (2.15)$$

по-друге, розділивши (2.14) на n , маємо вираз для визначення b_0 :

$$\frac{1}{n} \sum_{i=1}^n y_i - b_0 - b_1 \frac{1}{n} \sum_{i=1}^n x_i = 0 \Rightarrow b_0 = \bar{y} - b_1 \bar{x}. \quad (2.16)$$

Таким чином, ми знайшли формули для визначення невідомих параметрів b_0 та b_1 , і можемо записати у явному вигляді регресію y від x , у якій параметри обчислені за методом найменших квадратів. Її інколи називають регресією найменших квадратів y від x . Маємо:

$$\hat{y} = b_0 + b_1 x, \quad (2.17)$$

або

$$y = \hat{y} + e = b_0 + b_1 x + e. \quad (2.18)$$

Для ілюстрації цих викладок повернемося до нашого прикладу про дослідження ефективності витрат на рекламу. Проведені попередні розрахунки подамо у вигляді табл. 2.2.

Для обчислення невідомих параметрів b_0 , b_1 необхідно послідовно здійснити такі розрахунки:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{50}{5} = 10; \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{200}{5} = 40;$$

Таблиця 2.2

I	y_i	x_i	x_i^2	$x_i y_i$
1	25	5	25	125
2	30	6	36	180
3	35	9	81	315
4	45	12	144	540
5	65	18	324	1170
Σ	200	50	610	2330
Σ/n	40	10	122	466

$$\text{var}(x) = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 = \frac{610}{5} - 100 = 22;$$

$$\text{cov}(x, y) = \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y} = \frac{2330}{5} - 400 = 66;$$

$$b_1 = \frac{\text{cov}(x, y)}{\text{var}(x)} = \frac{66}{22} = 3; \quad b_0 = \bar{y} - b_1 \bar{x} = 40 - 3 \times 10 = 10.$$

Знаючи параметри b_0 , b_1 , отриману пряму запишемо у вигляді:

$$\hat{y} = 3x + 10.$$

Приклад. У таблиці 2.3 наведено умовні дані спостережень витрат на відпустку залежно від кількості членів родини.

Таблиця 2.3

Кількість членів родини x	Витрати на відпустку, ум. од. y
1	16
2	12
2	23
4	19
6	30
$\bar{x} = 3$	$\bar{y} = 20$

Для того, щоб встановити залежність витрат на відпустку від розмірів родини, припустимо, що ця залежність описується лінійною функцією (2.17), (табл. 2.3), тобто її можна розглядати як просту лінійну регресію (2.18).

Встановимо її невідомі параметри за формулами (2.16) та (2.7). Незважаючи на громіздкість цієї формули з першого погляду, вона найчастіше використовується на практиці.

Для підрахунку b_1 нам потрібно визначити

$$n, \sum_{i=1}^n x_i, \sum_{i=1}^n y_i, \sum_{i=1}^n x_i y_i, \sum_{i=1}^n x_i^2.$$

Відобразимо ці дані за допомогою табл. 2.4.

Таблиця 2.4

	x_i	y_i	$x_i y_i$	x_i^2	\hat{y}_i	$e_i = y_i - \hat{y}_i$
	1	16	16	1	14.74	1.26
	2	12	24	4	17.37	-5.37
	2	23	46	4	17.37	5.63
	4	19	76	16	22.63	-3.63
	6	30	180	36	27.89	2.11
Всього	15	100	342	61	100	0

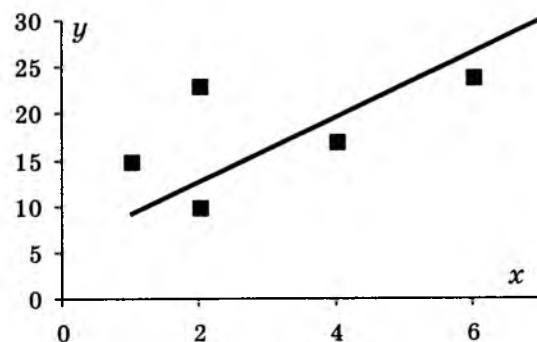
$$b_1 = \frac{342 - 15 \times 100 / 5}{61 - 225 / 5} = \frac{42}{61 - 45} \approx 2.63,$$

$$b_0 = \bar{y} - b_1 \bar{x} = 20 - 2.63 \times 3 = 12.11.$$

Отже, маємо:

$$\hat{y}_i = 12.11 + 2.63x_i. \quad (2.19)$$

Рівняння (2.19) дає для кожного спостережуваного значення x_i значення \hat{y}_i та e_i (дві останні колонки табл. 2.4). Підкреслимо, що сума оцінених значень дорівнює сумі фактичних значень y_i , а сума помилок дорівнює нулеві.



Малюнок 2.3. Залежність витрат на відпустку від кількості членів родини

2.3. ВЛАСТИВОСТІ ПРОСТОЇ ВИБІРКОВОЇ ЛІНІЙНОЇ РЕГРЕСІЇ

Проста вибіркова лінійна регресія, в якій невідомі параметри b_0 та b_1 визначені за методом найменших квадратів, має багато корисних властивостей.

1. Регресійна пряма проходить через середню точку (це аналогічно тому, що сума помилок дорівнює нулю). Ця властивість була виведена у (2.16).

$$\bar{y} = b_0 + b_1 \bar{x}. \quad (2.20)$$

2. Залишки мають нульову коваріацію зі спостережуваними значеннями x та оціненими значеннями \hat{y}_i .

Повернемося ще раз до (2.6), з якого випливає:

$$\frac{\partial \left(\sum_{i=1}^n e_i^2 \right)}{\partial b_1} = -2 \sum_{i=1}^n x_i (y_i - a - bx_i) = -2 \sum_{i=1}^n x_i e_i = 0;$$

$$\text{cov}(x, e) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(e_i - \bar{e}) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})e_i, \text{ тому що } \bar{e} = 0;$$

$$= \frac{1}{n} \sum_{i=1}^n x_i e_i - \frac{1}{n} \bar{x} \sum_{i=1}^n e_i = \frac{1}{n} \sum_{i=1}^n x_i e_i, \text{ тому що } \sum_{i=1}^n e_i = 0;$$

$$= 0 \text{ за (2.15).}$$

Змінна \hat{y} є лінійною функцією від x , звідки випливає, що $\text{cov}(\hat{y}, e) = 0$.

3. Сума квадратів залишків є функцією від кута нахилу.

Як уже було виведено у (2.16) та (2.9) (1.6.10), параметри b_0 та b_1 мають вигляд:

$$b_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (2.21)$$

$$b_0 = \bar{y} - b_1 \bar{x}. \quad (2.22)$$

$$\text{Введемо позначення } \tilde{x} = x - \bar{x}, \tilde{y} = y - \bar{y}, \quad (2.23)$$

тобто \tilde{x} та \tilde{y} є відхиленнями від середніх значень. Проста лінійна ре-

гресія за першою властивістю проходить через середню точку (\bar{x}, \bar{y}) , тому її можна взяти за основу (див. мал. 2.4).

Розглянемо точку А з координатами (x_i, y_i) . Відносно нової осі перша координата дорівнюватиме: $\tilde{x} = x_i - \bar{x}$, друга — $\tilde{y}_i = y_i - \bar{y}$.

Друга рівність може бути розкладена таким чином:

$$\tilde{y}_i = \hat{y}_i + e_i, \quad (2.24)$$

де $\hat{y}_i = \hat{y}_i - \bar{y}$, яка також є відхиленням оціненого значення \hat{y}_i від середнього \bar{y} ,

$$e_i = \tilde{y}_i - \hat{y}_i = y_i - \bar{y} - \hat{y}_i + \bar{y} = y_i - \hat{y}_i.$$

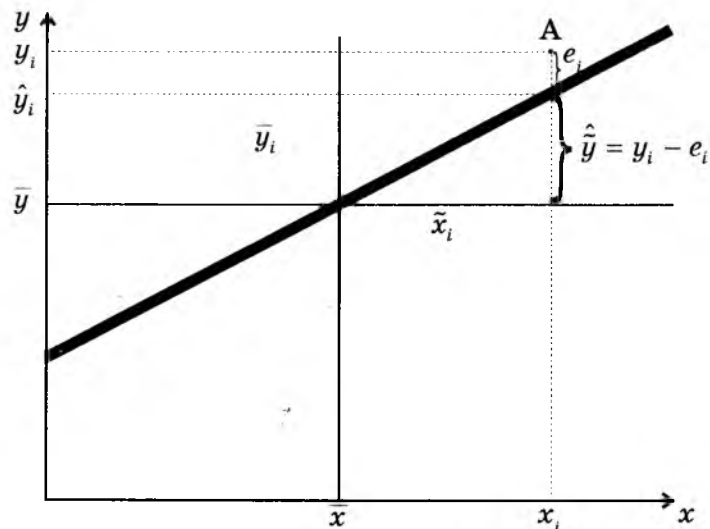
Лінійну регресію та суму квадратів залишків відповідно можна записати у вигляді:

$$\hat{y} = b_1 \tilde{x}$$

та

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (\tilde{y}_i - \hat{y}_i)^2 = \sum_{i=1}^n (\tilde{y}_i - b_1 \tilde{x}_i)^2 = \sum_{i=1}^n \tilde{y}_i^2 - 2b_1 \sum_{i=1}^n \tilde{x}_i \tilde{y}_i + b_1^2 \sum_{i=1}^n \tilde{x}_i^2. \quad (2.25)$$

Вираз (2.25) показує, що сума квадратів залишків є функцією від кута нахилу b_1 .



Малюнок 2.4. Перенесення осей координат у простій лінійній регресії

2.4. КОЕФІЦІЄНТИ КОРЕЛЯЦІЇ ТА ДЕТЕРМІНАЦІЇ

2.4.1. Поняття про коефіцієнт кореляції

Після того, як визначені невідомі параметри регресійної моделі, спробуємо оцінити щільність зв'язку між залежною величиною y і незалежною x . Тобто спробуємо відповісти на запитання, наскільки значним є вплив змінної x на y . Чи є якийсь критерій, який допомагає кількісно оцінити цей вплив? Найпростішим критерієм, який дає кількісну оцінку зв'язку між двома показниками, є коефіцієнт кореляції. Він розраховується за такою формулою:

$$r_{yx} = \frac{\text{cov}(x, y)}{\sqrt{\text{var}(x) \text{var}(y)}} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (2.26)$$

де $\text{cov}(x, y)$ — коефіцієнт коваріації між x та y ; $\text{var}(x)$ — дисперсія змінної x ; $\text{var}(y)$ — відповідно дисперсія змінної y .

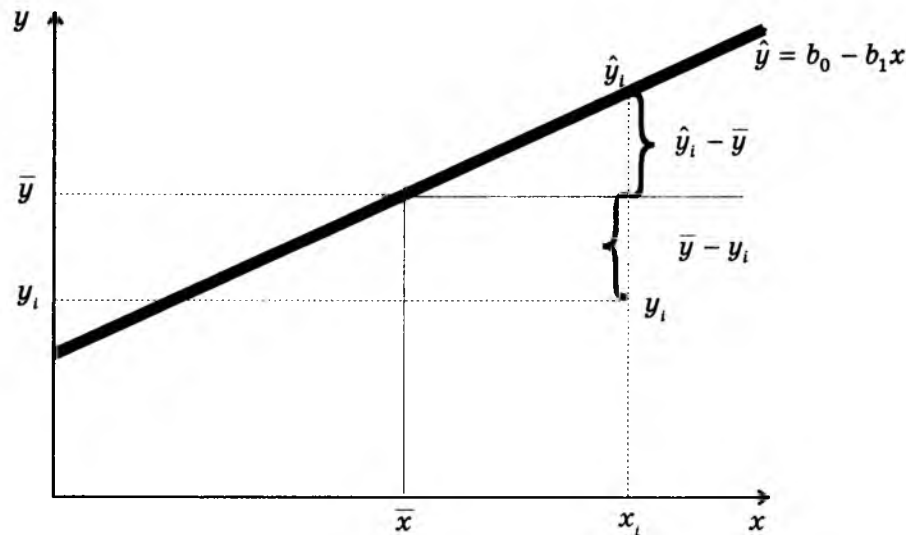
Як бачимо з виразу (2.26), коефіцієнт кореляції дорівнює відношенню коефіцієнта коваріації до кореня з добутку двох дисперсій. Коефіцієнт кореляції, на відміну від коефіцієнта коваріації, є вже не абсолютною, а відносною мірою зв'язку між двома факторами. Тому значення коефіцієнта кореляції завжди розташовані, як можна побачити з виразу (2.26), між -1 та $+1$ ($-1 \leq r_{xy} \leq 1$). Позитивне значення коефіцієнта кореляції свідчить про прямий зв'язок між показниками, а негативне — про зворотний зв'язок. Коли коефіцієнт кореляції прямує за абсолютною величиною до 1 , це свідчить про наявність сильного зв'язку ($r_{xy} \rightarrow \pm 1$ — щільність зв'язку велика); у протилежному випадку, коли коефіцієнт кореляції прямує до нуля ($r_{xy} \rightarrow 0$), зв'язку немає.

2.4.2. Декомпозиція дисперсій. Поняття про коефіцієнт детермінації

Поряд з коефіцієнтом кореляції використовується ще один критерій, за допомогою якого також вимірюється щільність зв'язку між двома або більше показниками та перевіряється адекватність (відповідність) побудованої регресійної моделі реальній дійсності. Тобто дається відповідь на запитання, чи справді зміна значення y лінійно залежить саме від зміни значення x , а не відбувається під впливом різних випадкових факторів. Таким критерієм є *коефіцієнт детермінації*. Перед тим, як розглянути,

що саме являє собою коефіцієнт детермінації та як він пов'язаний з коефіцієнтом кореляції, розглянемо питання про декомпозицію дисперсій, яке є одним з центральних у статистиці.

Спочатку спробуємо за допомогою мал. 2.5 уявити, як можна розбити на дві частини відхилення фактичних значень залежної змінної y від значень, що знаходяться на побудованій регресійній прямій (теоретичних значень).



Малюнок 2.5. Декомпозиція відхилень фактичних значень від теоретичних

Як бачимо з мал. 2.5, такі відхилення можна записати у вигляді:

$$\begin{aligned} (\hat{y}_i - y_i) &= (\hat{y}_i - \bar{y}) + (\bar{y} - y_i) \\ -(\hat{y}_i - y_i) &= -(\hat{y}_i - \bar{y}) + (y_i - \bar{y}) \Rightarrow \\ (y_i - \hat{y}_i) &= (y_i - \bar{y}) - (\hat{y}_i - \bar{y}). \end{aligned} \quad (2.27)$$

Вираз (2.27) перепишемо таким чином:

$$(y_i - \bar{y}) = (\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i). \quad (2.28)$$

У статистиці різницю $(y_i - \bar{y})$ прийнято називати **загальним відхиленням**. Різницю $(\hat{y}_i - \bar{y})$ називають **відхиленням, яке можна пояснити, виходячи з регресійної прямої**. Справді, якщо x_i змінюється, то можна завжди знайти значення цього відхилення, маючи тільки регресійну пряму, бо \bar{y} завжди залишається незмінною величиною. Різницю $(y_i - \hat{y}_i)$ називають **відхиленням, яке не можна пояснити, виходячи з регресійної прямої, або непояснюваним відхиленням**. Справді, якщо x_i змінюється, то змінюються обидві величини y_i і \hat{y}_i , тому, виходячи тільки з регресійної прямої, неможливо пояснити це відхилення.

Таким чином, якщо уважно розглянути вираз (2.28), то виявиться, що ми розклали загальне відхилення $(y_i - \bar{y})$ на відхилення $(y_i - \hat{y}_i)$, яке не можна пояснити з регресійної лінії, так зване непояснюване відхилення, та на відхилення $(\hat{y}_i - \bar{y})$, яке можна пояснити, виходячи з регресійної лінії.

Піднесемо обидві частини (2.28) до квадрата та підсумуємо за всіма індексами.

Отримаємо:

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + 2 \sum_{i=1}^n (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) + \sum_{i=1}^n (\hat{y}_i - \bar{y})^2.$$

Перепишемо суму добутку у вигляді:

$$\begin{aligned} \sum_{i=1}^n (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) &= \sum_{i=1}^n (b_0 + b_1 x_i - (b_0 + b_1 \bar{x}))(y_i - \hat{y}_i) = \\ &= -b_1 \bar{x} \sum_{i=1}^n (y_i - \hat{y}_i) + b_1 \sum_{i=1}^n (y_i - \hat{y}_i) x_i = \\ &= -b_1 \bar{x} \sum_{i=1}^n e_i + b_1 \sum_{i=1}^n e_i x_i = 0. \end{aligned}$$

Після виконання всіх дій отримаємо остаточний вираз:

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \sum_{i=1}^n (\hat{y}_i - \bar{y})^2, \quad (2.29)$$

де $\sum_{i=1}^n (y_i - \bar{y})^2$ — загальна сума квадратів, яка позначається, як правило, через SST ; $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ — сума квадратів помилок, яка позначається через

SSE ; $\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ — сума квадратів, що пояснює регресію та позначається через SSR .

Таким чином, вираз (2.29) у скороченому вигляді можна записати:

$$SST = SSE + SSR.$$

Поділивши (2.29) на n , отримуємо вираз для дисперсій:

$$\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} + \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{n}, \quad (2.30)$$

де $\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}$ — загальна дисперсія, яку позначимо $\sigma_{заг.}^2$;

$\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}$ — дисперсія помилок, яку позначимо $\sigma_{пом.}^2$;

$\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{n}$ — дисперсія, яку прийнято називати дисперсією, що пояснює регресію, позначимо її $\sigma_{регр.}^2$.

Таким чином, ми розклали загальну дисперсію на дві частини: дисперсію, що пояснює регресію, та дисперсію помилок (або дисперсію випадкової величини). Умовно це можна записати у вигляді:

$$\sigma_{заг.}^2 = \sigma_{пом.}^2 + \sigma_{регр.}^2 \quad (2.31)$$

Поділимо обидві частини (2.31) на $\sigma_{заг.}^2$ і отримаємо:

$$1 = \frac{\sigma_{пом.}^2}{\sigma_{заг.}^2} + \frac{\sigma_{регр.}^2}{\sigma_{заг.}^2}. \quad (2.32)$$

Як можна побачити з виразу (2.32), перша частина ($\sigma_{пом.}^2 / \sigma_{заг.}^2$) є пропорцією дисперсії помилок у загальній дисперсії, тобто являє собою частину дисперсії, яку не можна пояснити через регресійний зв'язок. Друга частина ($\sigma_{регр.}^2 / \sigma_{заг.}^2$) є складовою дисперсії, яку можна пояснити через регресійну лінію.

Частина дисперсії, що пояснює регресію, називається коефіцієнтом детермінації і позначається R^2 . Коефіцієнт детермінації використовується як критерій адекватності моделі, бо є мірою пояснювальної сили незалежної змінної x .

Таким чином, коефіцієнт детермінації можна записати у вигляді двох еквівалентних виразів:

$$R^2 = \sigma_{регр.}^2 / \sigma_{заг.}^2, \quad (2.33)$$

або

$$R^2 = \frac{SSR}{SST}. \quad (2.34)$$

З (2.32) випливає, що коефіцієнт детермінації завжди позитивний і перебуває у межах від нуля до одиниці ($0 \leq R^2 \leq 1$).

2.4.3. Зв'язок між коефіцієнтом кореляції та нахилом b_1

Звичайно, нас цікавить, чи є зв'язок між коефіцієнтом кореляції та детермінації, і якщо є, то який? Перш ніж відповісти на це запитання, розглянемо зв'язок між коефіцієнтом кореляції та нахилом регресійної лінії, тобто параметром b_1 . Нагадаємо формули для розрахунків коефіцієнта кореляції та нахилу:

$$r = \frac{\text{cov}(x, y)}{\sqrt{\text{var}(x) \cdot \text{var}(y)}} = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y}; \quad (2.35)$$

$$b_1 = \frac{\text{cov}(x, y)}{\sqrt{\text{var}(x)}} = \frac{\text{cov}(x, y)}{\sigma_x^2}. \quad (2.36)$$

Вираз (2.35) може бути переписаний у вигляді:

$$r = \left[\frac{\text{cov}(x, y)}{(\sigma_x)^2} \right] \times \left(\frac{\sigma_x}{\sigma_y} \right) \Rightarrow b_1 \frac{\sigma_x}{\sigma_y} \quad (2.37)$$

З того, що обидва значення σ_x і σ_y додатні, випливає, що знак коефіцієнта кореляції r завжди збігається із знаком параметра b_1 .

Крім того, з (2.37) випливає, що значення коефіцієнта кореляції (r) пов'язане із значеннями нахилу b_1 та середніх квадратичних відхилень σ_x і σ_y .

2.4.4. Зв'язок між коефіцієнтом кореляції (r) і коефіцієнтом детермінації (R^2)

Знаючи зв'язок між коефіцієнтом кореляції та нахилом регресійної лінії, розглянемо зв'язок між коефіцієнтом кореляції та детермінації. Нагадаємо формулу для розрахунку коефіцієнта детермінації:

$$R^2 = \frac{\sigma_{\text{регр.}}^2}{\sigma_{\text{заг.}}^2} = \frac{SSR}{SST}. \quad (2.38)$$

Нагадаємо також, що:

$$SST = \sum_{i=1}^n (y_i - \bar{y})^2; \quad (2.39)$$

$$SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2. \quad (2.40)$$

Перепишемо (2.40) у такому вигляді:

$$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \sum_{i=1}^n (b_0 + b_1 x_i - (b_0 + b_1 \bar{x}))^2 = b_1^2 \sum_{i=1}^n (x_i - \bar{x})^2. \quad (2.41)$$

Внесемо зміни до (2.38), враховуючи (2.39) і (2.41). Отримаємо:

$$R^2 = \frac{SSR}{SST} = \frac{b_1^2 \sum_{i=1}^n (x_i - \bar{x})^2 \cdot \left(\frac{1}{n}\right)}{\sum_{i=1}^n (y_i - \bar{y})^2 \cdot \left(\frac{1}{n}\right)} = \frac{b_1^2 \sigma_x^2}{\sigma_y^2} = b_1^2 \cdot \frac{\sigma_x^2}{\sigma_y^2}. \quad (2.42)$$

З (2.37) маємо:

$$r = b_1 \frac{\sigma_x}{\sigma_y}.$$

Отже, порівнюючи вирази (2.37) та (2.42), встановлюємо, що коефіцієнт детермінації дорівнює квадрату коефіцієнта кореляції:

$$R^2 = r^2. \quad (2.43)$$

Для ілюстрації наведених викладок повернемося до нашого прикладу. За наведеними раніше даними, розрахуємо коефіцієнт кореляції та детермінації, скориставшись табл. 2.5.

Таблиця 2.5

i	y_i	x_i	x_i^2	$x_i y_i$	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$	\hat{y}_i	$(\hat{y}_i - y_i)^2$	$(\hat{y}_i - \bar{y})^2$
1	25	5	25	125	-15	225	25	0	225
2	30	6	36	180	-10	100	28	4	144
3	35	9	81	315	-5	25	37	4	9
4	45	12	144	540	5	25	46	1	36
5	65	18	324	1170	25	625	64	1	576
Σ	200	50	610	2330	0	1000		10	990
Σ/n	40	10	122	466	0	200			

$$\bar{x} = 10; \quad \bar{y} = 40;$$

$$\text{var}(y) = \frac{1}{n} \sum (y_i - \bar{y})^2 = 200;$$

$$\text{var}(x) = 22; \quad \text{cov}(x, y) = 66$$

$$\sqrt{\text{var}(y)} = \delta_y = 14;$$

$$b_1 = 3; \quad b_0 = 10.$$

$$\sqrt{\text{var}(x)} = 4.7.$$

$$y = 3x + 10;$$

$$r = \frac{\text{cov}(x, y)}{\sqrt{\text{var}(x)} \cdot \sqrt{\text{var}(y)}} = \frac{\text{cov}(x, y)}{\sqrt{\text{var}(x)} \cdot \text{var}(y)} = \frac{66}{66.3} = 0.995;$$

$$r^2 = \frac{[\text{cov}(x, y)]^2}{\text{var}(x) \text{var}(y)} = \frac{66 \cdot 66}{22 \cdot 200} = 0.99.$$

2.5. ПОНЯТТЯ ПРО СТУПЕНІ ВІЛЬНОСТІ

Повернемося до тотожності (2.29), яка пов'язує загальну суму квадратів із сумою квадратів залишків та сумою квадратів, що пояснює регресію:

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \sum_{i=1}^n (\hat{y}_i - \bar{y})^2, \quad (2.44)$$

$$SST = SSE + SSR$$

Кожна сума квадратів пов'язана з числом, яке називають її “**ступенем вільності**”. Це число показує, скільки незалежних елементів інформації, що утворилися з елементів y_1, y_2, \dots, y_n , потрібно для розрахунку даної суми квадратів.

У статистиці кількістю ступенів вільності певної величини часто називають різницю між кількістю різних дослідів і кількістю констант, встановлених в результаті цих дослідів, незалежно один від одного. Окреме застосування цього поняття відноситься до суми квадратів.

Розглянемо, скільки ступенів вільності має кожна вивчена нами сума квадратів.

Почнемо із загальної суми квадратів (SST). Для утворення SST потрібно $(n-1)$ незалежних чисел, тому що з чисел $\{(y_1 - \bar{y}), (y_2 - \bar{y}), \dots, (y_n - \bar{y})\}$ незалежні тільки $(n-1)$ завдяки властивості:

$$\sum_{i=1}^n (y_i - \bar{y}) = 0.$$

Суму квадратів, що пояснює регресію (SSR), отримують, використовуючи тільки єдину незалежну одиницю інформації, яка утворюється з y_1, y_2, \dots, y_n , а саме b_1 .

Покажемо що, справді, нахил b_1 можна передати як функцію від y_1, y_2, \dots, y_n . Запишемо відхилення, що пояснює регресію, у вигляді:

$$\hat{y}_i - \bar{y} = b_1(x_i - \bar{x}).$$

З (2.41) маємо:

$$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = b_1^2 \sum_{i=1}^n (x_i - \bar{x})^2.$$

Отже, суму квадратів, що пояснює просту лінійну регресію, можна утворити, використовуючи тільки одну одиницю незалежної інформації, а саме

b_1 (для випадку багатофакторної регресії маємо іншу ситуацію, яку розглянемо пізніше). Звідси SSR має один ступінь вільності. Звертаємо увагу читача на те, що ступінь вільності в даному випадку збігається з кількістю незалежних змінних, що входять до регресійної моделі.

Сума квадратів помилок (SSE) має $(n - 2)$ ступенів вільності. Ця сума базується на кількості ступенів вільності, яка дорівнює різниці між кількістю спостережень і кількістю параметрів, що оцінюються. У разі простої лінійної регресії оцінюються два параметри b_0 та b_1 . Якщо позначити кількість спостережень через n , то для SSE маємо $(n - 2)$ ступенів вільності.

Ступені вільності прийнято позначати через DF , або dF , або df .

У разі простої лінійної регресії ступені вільності, як і суми квадратів, можна розкласти таким чином:

$$n - 1 = 1 + (n - 2). \quad (2.45)$$

2.6. ПРОСТИЙ ANOVA-АНАЛІЗ У ЛІНІЙНІЙ РЕГРЕСІЇ: АНАЛІЗ ДИСПЕРСІЙ

У цьому параграфі ми зупинимося лише на загальному уявленні про дисперсійний аналіз та розглянемо лише ті поняття, які надалі будуть потрібні для аналізу побудованих економетричних моделей.

Використовуючи суми квадратів та відповідні їм ступені вільності, введемо поняття про середні квадрати. **Середнім квадратом називається сума квадратів, поділена на відповідний їй ступінь вільності.** Таким чином, **середнім квадратом помилок називається сума квадратів помилок, поділена на відповідний ступінь вільності, який позначається через MSE .** У разі простої лінійної регресії середній квадрат помилок має вигляд:

$$MSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n - 2).$$

Середній квадрат, що пояснює регресію, позначається через MSR та відповідно дорівнює сумі квадратів, що пояснює регресію, поділену на її ступінь вільності. У разі простої лінійної регресії сума квадратів, що пояснює регресію (SSR), має лише один ступінь вільності, тобто середній квадрат збігається з сумою квадратів, а саме:

$$MRS = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 / 1 = SSR.$$

Слід зазначити, що для загальної суми квадратів середній квадрат не розраховується.

Суми квадратів пов'язані з певним джерелом варіації, а також із ступенями вільності і середніми квадратами. Зведемо їх усіх у таблиці 2.6, яка і називається базовою таблицею дисперсійного аналізу (ANOVA-таблиця).

Таблиця 2.6

ANOVA-таблиця

Джерело варіації	Кількість ступенів вільності	Сума квадратів	Середні квадрати
Зумовлене регресією (модель)	1	$SSR = \sum_{i=1}^n (y_i - \bar{y})^2$	$MRS = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 / 1$
Непояснюване за допомогою регресії (помилка)	$n-2$	$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$	$MSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 / n - 2$
Загальне	$n-1$	$\sum_{i=1}^n (y_i - \bar{y})^2$	

Побудуємо ANOVA-таблицю для розглянутого нами прикладу про залежність між обсягами реалізації продукції та витратами на рекламу. ANOVA-таблиця в умовах нашого прикладу матиме вигляд (табл. 2.7):

Таблиця 2.7

Джерело варіації	Кількість ступенів вільності	Сума квадратів	Середні квадрати
Модель	1	$\sum_{i=1}^n (y_i - \bar{y})^2 = 990$	$MRS = 990$
Помилка	$5-2=3$	$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = 10$	$MSE = 10/3=3.33$
Загальне	$5-1=4$	$\sum_{i=1}^n (y_i - \bar{y})^2 = 1000$	

2.7. ПЕРЕВІРКА ПРОСТОЇ РЕГРЕСІЙНОЇ МОДЕЛІ НА АДЕКВАТНІСТЬ. ПОНЯТТЯ F-КРИТЕРІЮ ФІШЕРА

Раніше ми показали, що адекватність простої лінійної регресійної моделі можна перевірити за допомогою коефіцієнта детермінації. Якщо його значення близьке до одиниці, то можна вважати, що модель адекватна. Якщо його значення близьке до нуля, то модель неадекватна, тобто не має лінійного зв'язку між залежною та незалежною змінними. Але який висновок можна зробити, якщо значення коефіцієнта кореляції має нечітко виражене граничне значення, наприклад 0.5, 0.45, 0.44 і т. ін. Зрозуміло, що в таких випадках важко зробити однозначний висновок про наявність зв'язку, тобто про адекватність моделі. Потрібен інший критерій, який би однозначно відповідав на питання про адекватність побудованої моделі. Найпоширенішим з таких критеріїв є критерій Фішера. Розглянемо, як він утворюється. Для цього повернемося до простої регресійної моделі:

$$y_i = b_0 + b_1 x_i + e_i. \quad (2.46)$$

На підставі тільки того, що до правої частини (2.46) входить випадкова величина e_i , уже можна зробити висновок, що величини y_i будуть також випадковими. Будь-яка функція від них буде також випадковою. Запам'ятаємо цей факт і повернемося ще раз до таблиці ANOVA-дисперсійного аналізу. Розглянемо:

$$MSR = \sum_{i=1}^n (\hat{y} - \bar{y})^2 / 1 = SSR/1; \quad (2.47)$$

$$MSE = \frac{SSE}{n-2} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-2}. \quad (2.48)$$

Як бачимо, середні квадрати MSR і MSE є функціями від залежних змінних, тому також будуть випадковими величинами, тобто матимуть свій розподіл, математичне сподівання, дисперсію та моменти.

З теорії імовірностей відомо (ми цей факт детально не розглядатимемо, а відсилаємо читача до математичного додатка цього підручника), що величина

$$F = \frac{MSR}{MSE} \quad (2.49)$$

має функцію розподілу F з $\{(1 \text{ та } (n-2))\}$ ступенями вільності у разі простої лінійної регресії за умови, що нахил узагальненої моделі дорівнює нулеві, тобто $\beta_1 = 0$. (Що таке узагальнена модель і чому ми вводимо β_0 та β_1 — пояснимо трохи пізніше). На цьому базується F -критерій Фішера, який дозволяє оцінити, чи значно нахил b_1 відрізняється від нуля, тобто перевірити побудовану модель на адекватність. Пояснимо цей факт. Справді, якщо оцінка нахилу b_1 незначно відрізняється від нуля, тоді:

$$\hat{y}_i = \beta_0 + \beta_1 x_i = \bar{y} + \beta_1 (x_i - \bar{x}) = \bar{y}, \quad (2.50)$$

за умови, що $\beta_1 = 0$.

Отже, вираз (2.50) дає змогу по-іншому інтерпретувати критерій Фішера. Він дозволяє перевірити базову гіпотезу (в статистиці вона називається нульовою гіпотезою (H_0)), що краще апроксимувати дані середнім значенням ($\hat{y}_i = \bar{y}$), ніж регресійною прямою ($\hat{y}_i = \beta_0 + \beta_1 x_i$). Це в свою чергу і дає змогу перевірити наявність або відсутність лінійного зв'язку між змінними, іншими словами, адекватність побудованої регресійної моделі реальній дійсності.

Перевірка моделі на адекватність за F -критерієм Фішера передбачає здійснення певних етапів:

1. На першому етапі розраховуємо величину так званого F -відношення:

$$F_{(1, n-2)} = \frac{MSR}{MSE},$$

де MSR — середній квадрат, який можна пояснити з регресійної моделі; MSE — середній квадрат помилок; $1, (n-2)$ — ступені вільності, відповідно пов'язані з MSR і MSE .

2. На другому етапі задаємо рівень значимості α або $\alpha \cdot 100\%$. Наприклад, якщо ми вважаємо, що можлива помилка α для нас становить 0.05 (або 5%), це означає, що ми можемо помилитися не більше ніж у 5% випадків, а в 95% випадків ($100 \cdot (1 - \alpha)\%$) наші висновки будуть правильними.

3. На третьому етапі за статистичними таблицями F -розподілу Фішера з $(1, n-2)$ ступенями вільності і рівнем значимості $100 \cdot (1 - \alpha)\%$ обчислимо критичне значення ($F_{кр}$).

4. Якщо розраховане нами значення $F > F_{кр}$, то ми відкидаємо гіпотезу H_0 , що $\beta_1 = 0$ (або що $(\hat{y}_i = \bar{y})$ з ризиком помилитися не більше ніж у 5% випадків).

Отже, якщо $F > F_{кр}$, то побудована нами регресійна модель адекватна реальній дійсності.

Повернемося до нашого прикладу. Перевіримо розраховану раніше модель на адекватність за F -критерієм Фішера. Використаємо для цього дані табл. 2.8.

Таблиця 2.8

i	y_i	\hat{y}_i	$(y_i - \hat{y}_i)^2$	$(\hat{y}_i - \bar{y})^2$
1	25	25	0	225
2	30	28	4	144
3	35	37	4	9
4	45	46	1	36
5	65	64	1	576
Σ	200		10	990

За табл. 2.8 і за таблицями ANOVA-дисперсійного аналізу (2.6) і (2.7), знайдемо середній квадрат, що пояснює регресію, та середній квадрат помилок:

$$SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = 990; \quad MSR = \frac{990}{1} = 990;$$

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = 10; \quad MSE = \frac{10}{3} = 3.33.$$

Використовуючи значення середніх квадратів, обчислимо F -відношення Фішера:

$$F_{(1, n-2)} = F_{(1, n-2)} = \frac{MSR}{MSE} = \frac{990}{3.33} = 300.$$

За таблицею F -розподілу знаходимо критичне значення $F_{кр}$ з 1 та 3 ступенями вільності, задавши попередньо рівень довіри 95% або рівень значимості (помилки) 5% . Це буде точка $F_{(1;3;0.95)} = 10.13$.

Розраховане значення $F_{(1,3)} = 300$; а табличне значення $F_{(1;3;0.95)кр} = 10.13$.

Отже, $F > F_{кр}$, що дозволяє зробити висновок про адекватність побудованої моделі реальній дійсності.

2.8. ІНШІ КРИТЕРІЇ ЯКОСТІ ЛІНІЙНОЇ РЕГРЕСІЇ

Припустимо, що нам відомі n прогнозних даних $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n$, які відповідають n реальним даним y_1, y_2, \dots, y_n , тобто ми маємо відповідно n помилок прогнозу e_1, e_2, \dots, e_n . Для визначення якості прогнозу на практиці дуже широко використовуються такі прості критерії.

1. **Середня помилка прогнозу ME (mean error)**, яка розраховується за формулою

$$ME = \bar{e} = \frac{1}{n} \sum_{i=1}^n e_i. \quad (2.51)$$

Критерій ME характеризує ступінь зміщення прогнозу і для правильних прогнозів повинен прямувати до 0 за умови великої кількості спостережень, тобто

$$ME \rightarrow 0, \text{ при } n \rightarrow \infty.$$

2. Дисперсія помилок (variation)

$$\text{var}(e) = \frac{1}{n} \sum_{i=1}^n (e_i - \bar{e})^2 = \sigma^2 \quad (2.52)$$

та стандартне відхилення (standart deviation)

$$\sigma = \sqrt{\text{var}(e)} = \sqrt{\frac{1}{n} \sum_{i=1}^n (e_i - \bar{e})^2}. \quad (2.53)$$

Цей критерій, який взагалі є класичним у статистиці, вимірює ступінь розкиду значень змінної навколо свого середнього значення.

Для простої лінійної регресії, як нам уже відомо, середнє значення помилок дорівнює нулеві. Тому

$$\text{var}(e) = \frac{1}{n} \sum_{i=1}^n e_i^2, \bar{e} = 0. \quad (2.54)$$

3. Абсолютне середнє відхилення (mean absolute deviation)

$$MAD(e) = \frac{1}{n} \sum_{i=1}^n |e_i - \bar{e}|. \quad (2.55)$$

У деяких машинних процедурах цей критерій розраховується за дещо іншою формулою, а саме:

$$MAD_t = \sum_{j=0}^{t-1} \alpha(1-\alpha)^j |e_{t-j}|, t = \overline{1, n}, \quad (2.56)$$

де α — довільно задана змінна; $0 < \alpha < 1$.

З формули (2.56) випливає рекурентна формула для визначення цього критерію:

$$MAD_t = \alpha |e_t| + (1-\alpha)MAD_{t-1}. \quad (2.57)$$

Початкове значення MAD_0 найчастіше приймається за таке, що дорівнює e_1 . Для достатньо великого класу статистичних розподілів значення стандартного відхилення дещо більше, ніж значення середнього абсолютного відхилення та суворо пропорційне йому. Константа пропорційності лежить у межах від 1,2 до 1,3. Здебільшого вибирається середнє значення 1,25, що дозволяє записати:

$$\sigma_t = 1.25 \cdot MAD_t. \quad (2.58)$$

4. Середній квадрат помилки MSE (mean square error)

$$MSE(e) = \frac{1}{n} \sum_{i=1}^n e_i^2. \quad (2.59)$$

Цей критерій для лінійної регресії збігається з дисперсією помилок (2.54).

Замість середнього квадрата помилок дуже часто використовується просто сума квадратів помилок:

$$SSE = \sum_{i=0}^n e_i^2. \quad (2.60)$$

Цей критерій особливо поширений при виборі оптимальних моделей прогнозування. З декількох моделей вибирається та, яка дає меншу суму квадратів помилок.

5. Абсолютна середня процентна помилка MAPE (mean absolute percentage error)

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|e_i|}{y_i} 100\%. \quad (2.61)$$

Цей критерій використовується при порівнянні точності прогнозів різних об'єктів, бо характеризує відносну точність прогнозу. При цьому

вважається, що значення $MAPE$ менше 10% дає високу точність прогнозу, а отже, і якість моделі; від 10 до 20% — добру точність, від 20 до 50% — задовільну точність; понад 50% — незадовільну точність.

6. Середня процентна помилка MPE (mean percentage error)

$$MPE = \frac{1}{n} \sum_{i=1}^n \frac{e_i}{y_i} 100\%. \quad (2.62)$$

Це показник незміщеності прогнозу. З точки зору практики для якісних моделей цей показник має бути малим, загалом не перевищувати 5%. Зазначимо, що як і показник $MAPE$, він не визначений для нульових значень y .

7. Середня абсолютна помилка MAE (mean absolute error)

$$MAE(e) = \frac{1}{n} \sum_{i=1}^n |e_i|. \quad (2.63)$$

Це дає змогу визначити середнє значення помилки, без врахування знака.

Описані вище критерії якості використовуються як додаткова інформація при виборі найкращої моделі з можливих. Ми розглянемо це питання в наступному розділі.

2.9. ІМОВІРНІСНИЙ ЗМІСТ ПРОСТОЇ РЕГРЕСІЇ

2.9.1. Узагальнена регресійна модель

Повернемося до нашого прикладу: як обсяги реалізації продукції фірми залежать від витрат на рекламу. Дані для виявлення форми зв'язку між цими двома факторами ми отримали в п'яти однорідних (за вибраними властивостями) географічних зонах. За цими даними будувалась лінійна регресійна модель та розраховувались невідомі параметри. Тепер припустимо, що такі спостереження ми проводили кожного місяця протягом 3 років, залишаючи незмінними витрати на рекламу (x_i) у кожній географічній зоні. Безперечно, обсяги реалізації продукції (y_i) будуть різними навіть при незмінних витратах на рекламу. Єдиним джерелом зміни є випадкова величина, яка знаходить відображення у різних значеннях y .

Припустимо, що узагальненою регресійною моделлю є:

$$y = \beta_0 + \beta_1 x + \varepsilon, \quad (2.64)$$

де β_0, β_1 — правильні параметри всієї генеральної сукупності, ε — неспостережувана випадкова величина.

Проводячи спостереження кожного місяця протягом 3 років, ми отримуємо множину, яка складається з 36 вибірок, кожна з яких має 5 пар значень x та y і кожна з яких дає пару значень невідомих параметрів b_0 та b_1 , знайдених за допомогою методу найменших квадратів.

Постає питання, як за значеннями параметрів b_0, b_1 зробити висновки про параметри усїєї сукупності. І взагалі, чи можна зробити якісь висновки? Далі ми покажемо, що коли параметри вибіркової лінійної моделі розраховані за методом найменших квадратів, то при певних класичних припущеннях математичне сподівання параметрів b_0 та b_1 дорівнює значенням параметрів узагальненої моделі (моделі, яка є дійсною для всієї генеральної сукупності) β_0, β_1 , тобто $E(b_0) = \beta_0, E(b_1) = \beta_1$.

Розглянемо спочатку основні припущення для простої лінійної регресії.

2.9.2. Класична модель лінійної регресії: основні припущення, що лежать в основі методу найменших квадратів

Мета регресійного аналізу полягає не тільки у визначенні невідомих параметрів вибіркової лінійної моделі b_0 та b_1 , а, насамперед, у висновках, які ми можемо зробити щодо дійсних значень параметрів узагальненої моделі β_0 і β_1 . Для того, щоб відповісти на запитання, наскільки наближаються знайдені оцінки b_0 і b_1 до відповідних значень параметрів узагальненої моделі, або, що те ж саме, наскільки наближається теоретичне значення Y_i до дійсного значення свого математичного сподівання $E(y/x_i)$, ми повинні не тільки точно визначити функціональну форму моделі, а й зробити певні припущення щодо випадкової величини ε та зв'язку між випадковою величиною і незалежною змінною x_i . Щоб з'ясувати, чому це так, повернемося ще раз до узагальненої лінійної регресійної моделі $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$. Як можна побачити, y_i залежить від x_i і ε_i . Тому, поки ми не зробимо певних припущень щодо випадкової величини ε та незалежної змінної x_i , ми не зможемо зробити ніякого статистичного висновку про y_i , а також, як це буде показано далі, про значення дійсних параметрів β_0 і β_1 . Отже, припущення щодо змінної x_i та випадкової величини ε_i є головними для інтерпретації регресійних оцінок.

Розглянемо припущення, які становлять основу класичного регресійного аналізу. Для простої лінійної регресії вони мають такий вигляд.

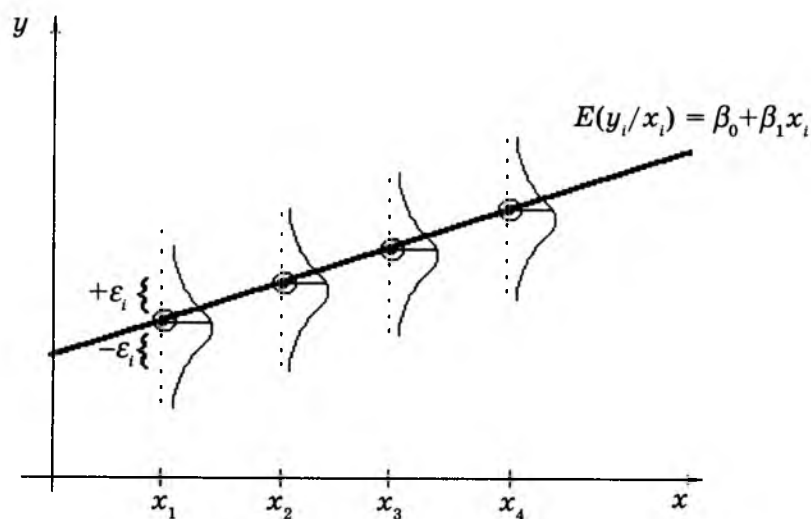
Припущення 1. Математичне сподівання випадкової величини ε дорівнює нулеві. У скороченому вигляді це припущення можна записати:

$$E(\varepsilon_i/x_i) = 0. \quad (2.65)$$

Припущення 1 констатує, що значення математичного сподівання ε_i , зумовлене даним x_i , дорівнює нулеві. Геометрично це припущення зображено на мал. 2.5, де показано кілька значень змінної x і набір y , що відповідають кожному з них. Кожен набір y , який відповідає даному x , розподілено навколо значення його математичного сподівання (обведені точки на регресійній прямій) з деякими значеннями y над математичним сподіванням і деякими під ним. Відстані над і під математичними сподіваннями і є випадковими величинами ε_i .

Припущення 1 вимагає, щоб математичне сподівання цих відхилень, відносно будь-якого даного x , дорівнювало нулеві.

Припущення 1 реально стверджує, що фактори, які не враховано в моделі і тому віднесено до ε_i , не впливають систематично на математичне сподівання y , тобто додатні значення ε_i нейтралізують від'ємні ε_i , тому їхній усереднений чи очікуваний вплив на y дорівнює нулю.



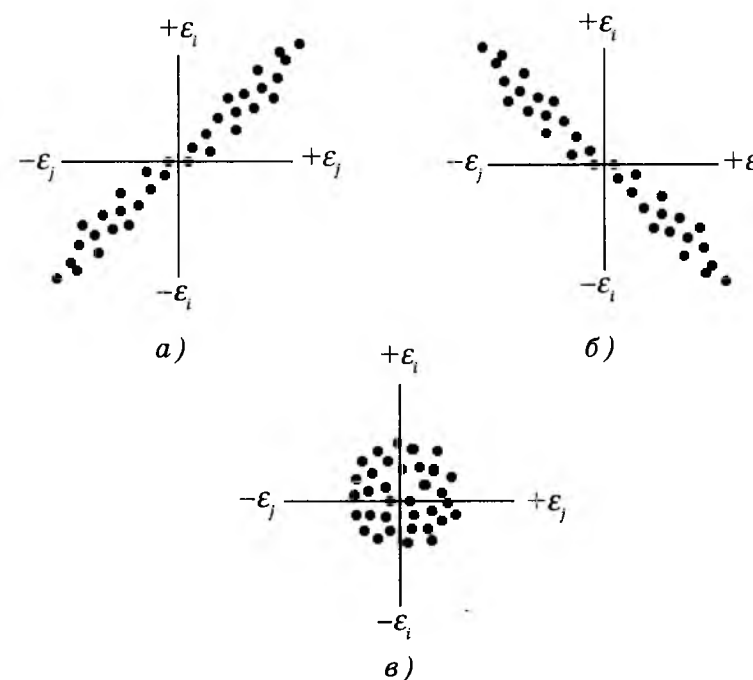
Малюнок 2.5. Умовний розподіл випадкової величини ε_i .

Зазначимо також, що припущення $E(\varepsilon_i/x_i) = 0$ передбачає $E(y_i/x_i) = \beta_0 + \beta_1 x_i$. Отже, ці два припущення еквівалентні.

Припущення 2. Відсутність автокореляції між випадковими величинами ε . Це припущення означає, що випадкові величини повинні бути незалежними між собою, тобто коефіцієнт коваріації між випадковими величинами повинен дорівнювати нулеві, що можна записати таким чином:

$$\text{cov}(\varepsilon_i, \varepsilon_j) = (E([\varepsilon_i - E(\varepsilon_i)][\varepsilon_j - E(\varepsilon_j)])) = E(\varepsilon_i \varepsilon_j) = 0, \quad i \neq j, \quad (2.66)$$

Припущення 2 стверджує, що випадкові величини незалежні одна від одної, тобто будь-яке i -те значення випадкової величини ε не впливає на будь-яке j -те значення, інакше кажучи, кореляції між ε_i і ε_j немає. У математичній статистиці та економетриці ця властивість формулюється через нульову коваріацію. На мал. 2.6, а зображена наявність додатної кореляції між випадковими величинами ε : додатне значення ε_i супроводжується додатним ε_j , а від'ємне значення ε_i супроводжується від'ємним ε_j . На мал. 2.6, б зображена наявність від'ємної кореляції між випадковими величинами ε — від'ємне значення ε_i супроводжується додатним ε_j , і навпаки. На мал. 2.6, в показано класичний випадок відсутності кореляції між випадковими величинами, тобто немає систематичності у розміщенні випадкових значень ε , тому коваріація між ними дорівнює нулю.



Малюнок 2.6. Зв'язок між випадковими величинами: а — наявність додатного зв'язку; б — наявність від'ємного зв'язку; в — відсутність зв'язку, кореляція та коваріація дорівнюють нулеві.

Припущення 2 дає змогу розглянути найпростіший випадок, коли вивчається систематичний вплив (якщо він є) x_i на y_i без врахування впливу інших факторів, виражених випадковою величиною ε .

Якщо це не так, то ми матимемо складнішу залежність, розгляд якої виходить за рамки класичної регресії. Такі ситуації ми будемо розглядати пізніше, а поки що проілюструємо на простому прикладі, що відбувається в разі порушення припущення 2.

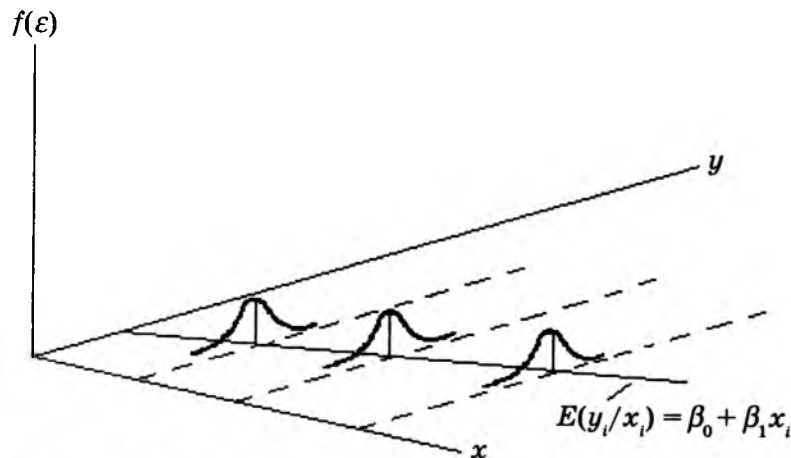
Припустимо, що у регресійній моделі $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ випадкові величини ε_i і ε_{i-1} мають додатну кореляцію. Тоді y_i залежатиме не тільки від x_i , а й від ε_{i-1} , оскільки значення ε_{i-1} певним чином визначає величину ε_i . Пізніше ми розглянемо, як можна тестувати наявність зв'язку між випадковими величинами та яким чином відсутність незалежності впливає на регресійну модель.

Припущення 3. Гомоскедастичність, або однакова дисперсія випадкових величин ε_i . Це припущення вимагає, щоб усі випадкові величини, незалежно від номера спостереження, мали однакову дисперсію.

Математично це припущення можна записати таким чином:

$$\text{var}(\varepsilon_i/x_i) = E[\varepsilon_i - E(\varepsilon_i)] = E(\varepsilon_i^2) = \sigma^2. \quad (2.67)$$

Вираз (2.67) означає, що дисперсія ε_i для кожного x_i (тобто умовна дисперсія ε_i) є константою, що дорівнює σ^2 . А це в свою чергу свідчить про те, що умовна дисперсія розподілу y є також сталою величиною. Цю ситуацію показано на мал. 2.7.

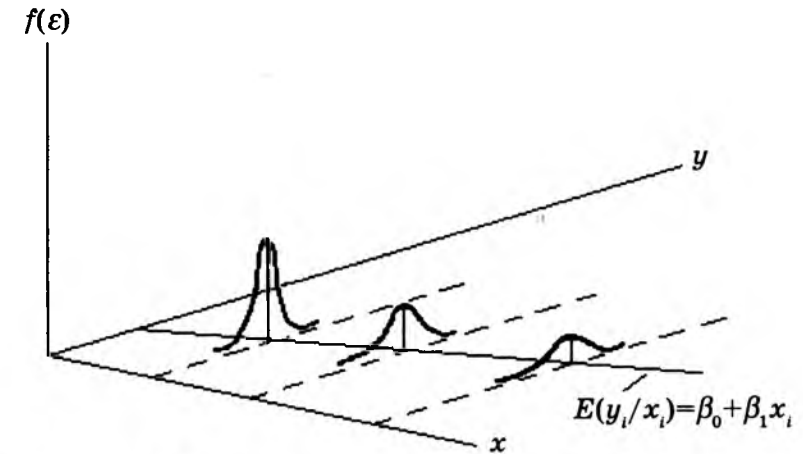


Малюнок 2.7. Гомоскедастичність, або постійна дисперсія

Порівняйте мал. 2.7 з мал. 2.8, де умовна дисперсія розподілу y зростає із збільшенням значень x . Ця ситуація відома як гетероскедастичність, або нерівна дисперсія. У цьому разі дисперсія випадкових величин уже не буде константою, а вираз (2.67) відповідно трансформується, тобто його можна розписати як:

$$\text{var}(\varepsilon_i/x_i) = \sigma_i^2. \quad (2.68)$$

Зверніть увагу, що індекс біля σ^2 у рівнянні (2.68) показує, що тепер дисперсія розподілу випадкової величини вже не є сталою.



Малюнок 2.8. Гетероскедастичність, або нерівна дисперсія

Щоб зрозуміти необхідність припущення про постійну дисперсію, або як часто кажуть про гомоскедастичність, звернемося ще раз до мал. 2.8. Як показано на малюнку, $\text{var}(\varepsilon/x_1) < \text{var}(\varepsilon/x_2), \dots < \text{var}(\varepsilon/x_i)$. Тому зразу ж постає запитання, які з розподілів залежних змінних y вибирати для опису реальної ситуації — ті, що щільніше наближені до своїх математичних сподівань, чи ті, що мають великий розкид?

Вводячи *припущення 3*, обмежимося випадком, коли всі значення y , які відносяться до різних значень x , є однаково важливими. У розділі 5 ми покажемо, що відбувається, коли припущення про однакову дисперсію порушується, тобто ми розглянемо складніший випадок — випадок гетероскедастичності.

Зверніть увагу на те, що *припущення 3* означає, що $\text{var}(y_i/x_i) = \sigma^2$, тобто умовна дисперсія y , теж є гомоскедастичною. Спробуйте аргументувати це самостійно.

Припущення 4. Незалежність між значеннями випадкової величини ε_i і значеннями змінної x_i , або нульова коваріація між ε_i та x_i .

Формально це припущення можна записати так:

$$\begin{aligned} \text{cov}(\varepsilon_i, x_i) &= E(\varepsilon_i - E(\varepsilon_i))(x_i - E(x_i)) = \\ &= E(\varepsilon_i(x_i - E(x_i))) = \quad \text{оскільки } E(\varepsilon_i) = 0 \\ &= E(\varepsilon_i x_i) - E(x_i)E(\varepsilon_i) = \text{оскільки } E(x_i) = \text{const} \\ &= E(\varepsilon_i x_i) =, \quad \text{за припущенням.} \\ &= 0, \end{aligned}$$

Припущення 4 передбачає відсутність кореляції між випадковою величиною ε і незалежною змінною x . Якщо припустити протилежне, то коли x і ε мають додатну кореляцію, x зростає із зростанням ε , і зменшується при зменшенні ε . Аналогічно, якщо x і ε мають від'ємну кореляцію, x зростає у випадку зменшення ε і зменшується, коли ε збільшується. В обох випадках важко простежити вплив x на y .

Припущення 4 виконується автоматично, якщо змінна x є не випадковою, або нестохастичною, і зберігається припущення 1, з якого випливає, що $\text{cov}(\varepsilon_i, x_i) = E((\varepsilon_i - E(\varepsilon_i))(x_i - E(x_i))) = 0$.

Але, оскільки ми припустили, що змінна x є не тільки нестохастичною, а має також фіксовані значення у повторюваних вибірках, то припущення 4 для нас не є вирішальним. Його застосовуємо для того, аби показати, що регресійна теорія зберігатиметься навіть тоді, коли значення x будуть випадковими за умови, що вони незалежні чи хоча б не мають кореляції з відхиленнями ε_i . Наслідки відкидання припущення 4 розглянемо у розділі 5.

Припущення 5. Регресійну модель визначено (специфіковано) правильно (відсутність похибки). Це найвагомніше і, напевне, найрідше застосовуване припущення. Повну значущість цього припущення буде розглянуто у розділі 5. Зараз ми спробуємо пояснити його якнайпростіше.

Як було зазначено у вступі, економетричне дослідження починається із специфікації економетричної моделі, яка має бути адекватною економічному процесу, що вивчається. При специфікації моделі, яка описує досліджувану ситуацію, виникає кілька важливих запитань, а саме:

які змінні потрібно включати в модель; якою є функціональна форма моделі; чи є вона лінійною за параметрами та змінними, чи ні; які можливі припущення щодо y_i , x_i і ε_i можна зробити у моделі?

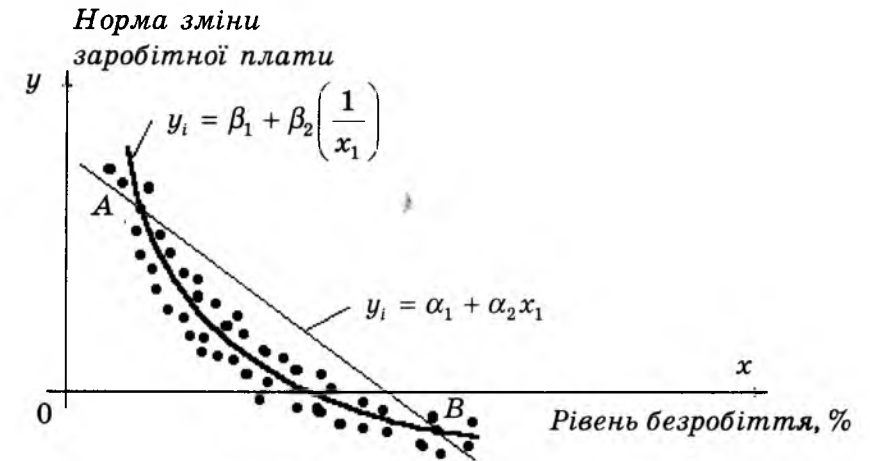
Це надзвичайно важливі питання, бо, наприклад, вилучаючи з моделі важливі змінні, чи вибираючи неправильну функціональну форму зв'язку, чи вдаючись до неправильних припущень щодо змінних моделі, ми ставимо під сумнів правильність інтерпретації оцінюваної регресії. Щоб дати інтуїтивне уявлення про важливість **припущення 5**, звернемося до кривої Філіпса, зображеної на мал. 2.9. Припустимо, що ми обираємо дві різні моделі, які відображають зв'язок між рівнем зміни номінальної заробітної плати і рівнем безробіття:

$$y_i = \alpha_1 + \alpha_2 x_i + \varepsilon_i; \quad (2.69)$$

$$y_i = \beta_1 + \beta_2 \left(\frac{1}{x_i} \right) + \varepsilon_i, \quad (2.70)$$

де y_i — дорівнює рівневі зміни номінальної заробітної плати; x_i — рівень безробіття.

Регресійна модель (2.69) є лінійною і за параметрами, і за змінними, тоді як (2.70) є лінійною за параметрами (тобто моделлю лінійної регресії, за нашим означенням), але нелінійною за змінною x .



Малюнок 2.9. Лінійна і нелінійна криві Філіпса

Якщо в дійсності є “правильною” модель (2.70), а ми вибираємо модель (2.69), то, як показано на мал. 2.9, вона забезпечить неправильний прогноз: між точками A і B для будь-якого даного x_i модель (2.69) переоцінюватиме справжнє математичне сподівання y_i , і в той же час недооцінюватиме (чи переоцінюватиме — в абсолютних термінах) математичне сподівання y зліва від A (справа від B).

Цей умовний приклад є зразком того, що називається неправильною специфікацією, або помилкою специфікації, і полягає у виборі неправильної функціональної форми¹.

На жаль, на практиці ми не завжди відразу можемо встановити правильні змінні, які потрібно залучити в модель, або не відразу можемо знайти правильну форму моделі чи правильні припущення щодо змінних. Отже, при побудові моделі спеціалісти повинні проводити копітку роботу щодо вибору змінних для моделі, робити певні припущення щодо стохастичної природи змінних, послідовно вибирати найкращу функціональну форму моделі.

Ще раз задамо собі запитання, а для чого все ж таки потрібне припущення 5? Не вникаючи в деталі, достатньо сказати, що дане припущення нагадує нам про залежність регресійного аналізу і, отже, його результатів від обраної моделі, і застерігає нас про необхідність обережно формулювати економетричні моделі, особливо, коли одразу є кілька теорій щодо пояс-

¹ Gujarati D.N. Basic Econometrics. — P. 58.

нення економічного явища, наприклад процентної ставки, попиту на гроші, визначення рівноважної вартості акцій і облігацій тощо. Тому економетрична побудова моделей — це, як ми побачимо далі, більшою мірою мистецтво, ніж наука.

Припущення 6. Випадкова величина розподілена нормально з математичним сподіванням нуль та сталою дисперсією σ^2 .

Формально це припущення можна записати у вигляді:

$$\varepsilon \sim N(0, \sigma^2).$$

Припущення 6, як ми побачимо далі, є необхідним при побудові інтервалів довіри для параметрів, залежної змінної y та інших характеристик лінійної регресійної моделі.

Слід зауважити, що дані припущення стосуються тільки узагальненої регресійної моделі (УРМ) і не стосуються вибіркової моделі (ВРМ). Проте цікаво спостерігати, що метод найменших квадратів, описаний раніше, має деякі властивості, схожі на припущення, які ми щойно зробили. Наприклад, висновок, що $\sum e_i = 0$ і, отже, $\bar{e} = 0$, схожий на припущення, що $E(\varepsilon_i / x_i) = 0$. Аналогічно висновок, що $\sum e_i x_i = 0$, схожий на припущення, що $\text{cov}(\varepsilon_i, x_i) = 0$.

2.9.3. Розподіл залежної змінної y

Як пояснювалося вище, при переході від однієї вибірки до множини вибірок, коли x залишається незмінним, єдиним джерелом зміни y є випадкова величина. Отже, змінна y також є випадковою величиною. Звичайно, розподіл залежної змінної y певною мірою залежить від припущень, прийнятих для випадкової величини ε .

Покажемо, що залежна змінна має нормальний розподіл з математичним сподіванням

$$E(y_i) = \beta_0 + \beta_1 x_i \quad (2.71)$$

та дисперсією

$$\text{var}(y_i) = E[y_i - E(y_i)]^2 = E(\varepsilon_i^2) = \sigma_\varepsilon^2. \quad (2.72)$$

Доведення 1

Математичне сподівання залежної змінної $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ дорівнює:

$$E(y_i) = E(\beta_0 + \beta_1 x_i + \varepsilon_i) = E(\beta_0 + \beta_1 x_i) + E(\varepsilon_i).$$

Враховуючи, що β_0 та β_1 є параметрами узагальненої моделі, а отже, константами, та що x_i має фіксоване значення, отримуємо першу складову частину:

$$E(\beta_0 + \beta_1 x_i) = \beta_0 + \beta_1 x_i.$$

За припущенням 1 простої лінійної регресії маємо: друга складова частина дорівнює нулю — $E(\varepsilon_i) = 0$.

Звідки остаточно $E(y_i) = \beta_0 + \beta_1 x_i$.

Дисперсія залежної змінної дорівнює дисперсії випадкової величини:

$$D(y_i) = E[y_i - E(y_i)]^2 = \sigma_\varepsilon^2.$$

Введемо заміну змінних: $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ та $E(y_i) = \beta_0 + \beta_1 x_i$. Після відповідної заміни у виразі для дисперсії y отримуємо:

$$E[y_i - E(y_i)]^2 = E[\beta_0 + \beta_1 x_i + \varepsilon_i - \beta_0 - \beta_1 x_i]^2 = E(\varepsilon_i)^2.$$

Отже, дисперсія змінної y дорівнює дисперсії випадкової величини.

Тепер залишилось тільки показати, що випадкова величина y розподілена нормальним законом розподілу з відповідним знайденим математичним сподіванням та дисперсією.

Як відомо, тип розподілу y визначається типом розподілу випадкової величини ε , який є нормальним за припущенням 6. Очевидно, що β_0 та β_1 , які є константами, не впливають на розподіл y . Крім того, значення змінної x_i — це за властивістю 5 ряд констант, а тому воно також не впливає на розподіл y . Отже, розподіл випадкової величини y визначається тільки розподілом випадкової величини ε , тобто є також нормальним.

2.10. ЗАКОН РОЗПОДІЛУ ПАРАМЕТРІВ. МАТЕМАТИЧНЕ СПОДІВАННЯ ТА ДИСПЕРСІЯ РОЗПОДІЛУ ПАРАМЕТРІВ

Математичне сподівання та дисперсія розподілу параметрів b_0 та b_1

У цьому параграфі покажемо, що:

1. Математичне сподівання параметра b_0 дорівнює: $E(b_0) = \beta_0$.
2. Дисперсія b_0 в свою чергу матиме вигляд:

$$\text{var}(b_0) = E[b_0 - \beta_0]^2 = \sigma_\varepsilon^2 \frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2}.$$

3. Математичне сподівання параметра b_1 дорівнює: $E(b_1) = \beta_1$.
4. Дисперсія параметра b_1 визначається за формулою:

$$\text{var}(b_1) = E[b_1 - \beta_1]^2 = \sigma_\varepsilon^2 \frac{1}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

5. Оцінка дисперсії випадкової величини ε набуде виразу:

$$\hat{\sigma}_\varepsilon^2 = S^2 = \frac{\sum_{i=1}^n e_i^2}{n - k},$$

де k — кількість оцінюваних параметрів у регресійній моделі. У разі простої лінійної регресії, коли ми оцінюємо тільки два параметри, $k=2$, отже, оцінка дисперсії випадкової величини має такий вигляд:

$$\hat{\sigma}_\varepsilon^2 = S^2 = \frac{\sum_{i=1}^n e_i^2}{n - 2}.$$

Знаходження математичного сподівання параметра b_1

Згадаємо, що параметр b_1 у простій лінійній регресії визначається за відомою формулою:

$$b_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n \tilde{x}_i (y_i - \bar{y})}{\sum_{i=1}^n \tilde{x}_i^2} = \frac{\sum_{i=1}^n \tilde{x}_i y_i - \bar{y} \sum_{i=1}^n \tilde{x}_i}{\sum_{i=1}^n \tilde{x}_i^2} =$$

$$= \frac{\sum_{i=1}^n \tilde{x}_i y_i}{\sum_{i=1}^n \tilde{x}_i^2} = \sum_{i=1}^n \alpha_i y_i, \quad (2.73)$$

де введено такі додаткові позначення:

$$\alpha_i = \frac{\tilde{x}_i}{\sum_{i=1}^n \tilde{x}_i^2};$$

$$\sum_{i=1}^n \tilde{x}_i = \sum_{i=1}^n (x_i - \bar{x}) = 0. \quad (2.74)$$

З (2.74) випливає, що:

$$\sum_{i=1}^n \alpha_i = 0; \quad \sum_{i=1}^n \alpha_i \tilde{x}_i = 1; \quad (2.75)$$

$$\sum_{i=1}^n \alpha_i^2 = \frac{1}{\sum_{i=1}^n \tilde{x}_i^2}. \quad (2.76)$$

Покажемо, що це справді так:

$$\sum_{i=1}^n \alpha_i = \frac{\sum_{i=1}^n \tilde{x}_i}{\sum_{i=1}^n \tilde{x}_i^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})}{\sum_{i=1}^n \tilde{x}_i^2} = \frac{0}{\sum_{i=1}^n \tilde{x}_i^2} = 0;$$

$$\sum_{i=1}^n \alpha_i \tilde{x}_i = \frac{\sum_{i=1}^n \tilde{x}_i \tilde{x}_i}{\sum_{i=1}^n \tilde{x}_i^2} = \frac{\sum_{i=1}^n (x_i - \bar{x}) \tilde{x}_i}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n \tilde{x}_i^2 - \bar{x} \sum_{i=1}^n \tilde{x}_i}{\sum_{i=1}^n (x_i - \bar{x})^2} = 1,$$

тому що

$$\sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n (x_i^2 - 2x_i \bar{x} + \bar{x}^2) = \sum_{i=1}^n x_i^2 - \bar{x} \sum_{i=1}^n x_i;$$

Для того, щоб знайти *математичне сподівання параметра* b_1 , передамо його через випадкову величину генеральної сукупності i , підставивши в (2.73), отримуємо:

$$b_1 = \sum_{i=1}^n \alpha_i (\beta_0 + \beta_1 x_i + \varepsilon_i) = \beta_1 + \sum_{i=1}^n \alpha_i \varepsilon_i. \quad (2.77)$$

Далі розрахуємо математичне сподівання членів у правій та лівій частині залежності:

$$E(b_1) = E(\beta_1) + E\left(\sum_{i=1}^n \alpha_i \varepsilon_i\right) = E(\beta_1) + \frac{1}{\sum_{i=1}^n \tilde{x}_i^2} E\left(\sum_{i=1}^n \tilde{x}_i \varepsilon_i\right).$$

Виходячи з того, що β_1 — правильний параметр генеральної сукупності є константою, маємо $E(\beta_1) = \beta_1$.

Чисельник другого додатка є не що інше, як коефіцієнт коваріації випадкової величини ε та незалежної змінної x_i . За припущенням 4, він дорівнює нулеві:

$$\text{cov}(x, \varepsilon) = E((x - E(x))(\varepsilon - E(\varepsilon))) = E((x - \bar{x})(\varepsilon - \bar{\varepsilon})) = 0.$$

Тобто другий додаток дорівнює нулеві. Звідси *математичне сподівання параметра* b_1 :

$$E(b_1) = \beta_1. \quad (2.78)$$

Знаходження дисперсії параметра b_1

Дисперсія параметра b_1 дорівнює:

$$\text{var}(b_1) = E[b_1 - E(b_1)]^2 = E(b_1 - \beta_1)^2 = \sigma_\varepsilon^2 \frac{1}{\sum_{i=1}^n \tilde{x}_i^2} = \sigma_\varepsilon^2 \frac{1}{\sum_{i=1}^n (x_i - \bar{x})^2}. \quad (2.79)$$

Покажемо, що рівняння (2.79) правильне. З (2.73) маємо:

$$b_1 = \sum_{i=1}^n \alpha_i y_i. \quad (2.80)$$

З (2.80) виразимо дисперсію параметра:

$$\text{var}(b_1) = \text{var}\left(\sum_{i=1}^n \alpha_i y_i\right) = \sum_{i=1}^n \alpha_i^2 \text{var}(y_i),$$

вважаючи, що α_i — константа, яка не залежить від значень y_i (див. властивість 5).

Але ж, як було показано у попередньому параграфі:

$$\text{var}(y_i) = \sigma_\varepsilon^2.$$

Звідси остаточно формула дисперсії параметра набуде вигляду:

$$\begin{aligned} \text{var}(b_1) &= \sum_{i=1}^n \alpha_i^2 \sigma_\varepsilon^2 = \sigma_\varepsilon^2 \sum_{i=1}^n \alpha_i^2 = \sigma_\varepsilon^2 \sum_{i=1}^n \left(\frac{\tilde{x}_i^2}{\sum_{i=1}^n \tilde{x}_i^2} \right)^2 = \\ &= \sigma_\varepsilon^2 \frac{\sum_{i=1}^n \tilde{x}_i^2}{\left(\sum_{i=1}^n \tilde{x}_i^2 \right)^2} = \sigma_\varepsilon^2 \frac{1}{\sum_{i=1}^n \tilde{x}_i^2} = \sigma_\varepsilon^2 \frac{1}{\sum_{i=1}^n (x_i - \bar{x})^2}. \end{aligned} \quad (2.81)$$

Знаходження математичного сподівання та дисперсії параметра b_0

Обчислення математичного сподівання параметра b_0 .

Тепер знайдемо математичне сподівання параметра b_0 . Для цього повернемося до відомої формули розрахунку параметра b_0 :

$$b_0 = \bar{y} - b_1 \bar{x} = \beta_0 + \beta_1 \bar{x} + \bar{\varepsilon} - b_1 \bar{x} = \beta_0 - (b_1 - \beta_1) \bar{x} + \bar{\varepsilon}. \quad (2.82)$$

Розрахуємо математичне сподівання правої та лівої частин цієї рівності. Раніше ми вже показували, що:

$$E(b_1) = \beta_1 \text{ та } E(\bar{\varepsilon}) = 0.$$

Звідси:

$$E(b_0) = \beta_0.$$

Знаходження дисперсії параметра b_0

Значення дисперсії параметра b_0 обчислюється аналогічно дисперсії параметра b_1 . Покажемо, що:

$$\begin{aligned} \text{var}(b_0) &= E[b_0 - E(b_0)]^2 = E(b_0 - \beta_0)^2 = \sigma_\varepsilon^2 \frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n \tilde{x}_i^2} = \\ &= \sigma_\varepsilon^2 \frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2}. \end{aligned}$$

Повернемося ще раз до відомої формули розрахунку параметра b_0 :

$$b_0 = \bar{y} - b_1 \bar{x}. \quad (2.83)$$

Замінімо в цій формулі

$$b_1 = \sum_{i=1}^n \alpha_i y_i.$$

Виходячи з цієї заміни з (2.83) отримаємо:

$$b_0 = \bar{y} - \bar{x} \sum_{i=1}^n \alpha_i y_i = \frac{\sum_{i=1}^n y_i}{n} - \bar{x} \sum_{i=1}^n \alpha_i y_i = \sum_{i=1}^n \left(\frac{1}{n} - \bar{x} \alpha_i \right) y_i. \quad (2.84)$$

З (2.84) маємо:

$$\text{var}(b_0) = \text{var} \left[\sum_{i=1}^n \left(\frac{1}{n} - \bar{x} \alpha_i \right) y_i \right] = \sum_{i=1}^n \left(\frac{1}{n} - \bar{x} \alpha_i \right)^2 \text{var}(y_i). \quad (2.85)$$

Ми вже показали у попередньому параграфі, що:

$$\text{var}(y_i) = \sigma_\varepsilon^2.$$

Виходячи з цього (2.85) можна записати у вигляді:

$$\text{var}(b_0) = \sigma_\varepsilon^2 \sum_{i=1}^n \left(\frac{1}{n} - \frac{2\bar{x}\alpha_i}{n} + \bar{x}^2 \alpha_i^2 \right).$$

Беручи до уваги (2.75 та 2.76), що

$$\sum_{i=1}^n \alpha_i = 0; \sum_{i=1}^n \alpha_i^2 = \frac{1}{\sum_{i=1}^n \tilde{x}_i^2},$$

отримаємо:

$$\text{var}(b_0) = \sigma_\varepsilon^2 \sum_{i=1}^n \left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n \tilde{x}_i^2} \right) = \sigma_\varepsilon^2 \left[\frac{\sum_{i=1}^n \tilde{x}_i^2 + n\bar{x}^2}{\sum_{i=1}^n \tilde{x}_i^2} \right].$$

Враховуючи, що

$$\sum_{i=1}^n \tilde{x}_i^2 = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - n\bar{x}^2,$$

остаточне значення *дисперсії параметра* матиме вигляд:

$$\text{var}(b_0) = \sigma_\varepsilon^2 \frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2}. \quad (2.86)$$

Отже, ми встановили значення математичного сподівання та дисперсії параметрів лінійної регресії.

Як було уже показано, параметри b_0 та b_1 можна виразити як лінійні функції від випадкової величини ε . З припущення 6 випливає, що випадкові величини розподілені нормально, тому і параметри лінійної регресії є випадковими величинами, розподіленими нормально.

2.11. ОЦІНКА ДИСПЕРСІЇ ВИПАДКОВОЇ ВЕЛИЧИНИ ε

У формулах для розрахунку дисперсії параметрів b_0 та b_1 наявна дисперсія випадкової величини ε генеральної сукупності — σ_ε^2 . Взагалі, саме значення дисперсії σ_ε^2 не може бути визначене, тому що випадкову величину ε ми не можемо спостерігати. Покажемо, що цю дисперсію можна замінити на оцінку:

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_{i=1}^n e_i}{n-2}.$$

Для цього нам необхідно довести, що $E(\hat{\sigma}_\varepsilon^2) = \sigma_\varepsilon^2$.

Якщо брати до уваги множину вибірок, то формула для довільної i -ої помилки має вигляд:

$$e_i = y_i - \hat{y}_i = \beta_0 + \beta_1 x_i + \varepsilon_i - b_0 - b_1 x_i = \varepsilon_i - (b_0 - \beta_0) - (b_1 - \beta_1) x_i. \quad (2.87)$$

З огляду на (2.82) з рівняння (2.87) випливає:

$$\begin{aligned} b_0 - \beta_0 &= \bar{\varepsilon} - (b_1 - \beta_1) x_i \Rightarrow \\ e_i &= (\varepsilon_i - \bar{\varepsilon}) - (b_1 - \beta_1)(x_i - \bar{x}). \end{aligned} \quad (2.88)$$

Зуважимо, що $\bar{\varepsilon}$ — це середнє значення помилки у генеральній сукупності; воно не може дорівнювати нулю, навіть коли $E(\varepsilon)=0$.

Піднесемо обидві частини (2.88) до квадрата та підсумуємо за всіма i :

$$\begin{aligned} \sum_{i=1}^n e_i^2 &= \sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^2 + (b_1 - \beta_1)^2 \sum_{i=1}^n (x_i - \bar{x})^2 - 2(b_1 - \beta_1) \sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})(x_i - \bar{x}) = \\ &= \sum_{i=1}^n \varepsilon_i^2 - n\bar{\varepsilon}^2 + (b_1 - \beta_1)^2 \sum_{i=1}^n (x_i - \bar{x})^2 - 2(b_1 - \beta_1) \sum_{i=1}^n (\varepsilon_i \tilde{x}_i), \end{aligned} \quad (2.89)$$

де $\tilde{x}_i = (x_i - \bar{x})$,

тому що

$$E\left(\sum_{i=1}^n \varepsilon_i^2\right) = n\sigma_\varepsilon^2 \Rightarrow E(\bar{\varepsilon}^2) = \text{var}(\bar{\varepsilon}) = \frac{\sigma_\varepsilon^2}{n}. \quad (2.90)$$

У попередньому параграфі ми встановили вираз для дисперсії нахилу:

$$E(b_1 - \beta_1)^2 = \text{var}(b_1) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^n \tilde{x}_i^2}. \quad (2.91)$$

Крім того, з (2.77):

$$E\left((b_1 - \beta_1) \sum_{i=1}^n (\varepsilon_i \tilde{x}_i)\right) = E\left(\left(\sum_{i=1}^n \alpha_i \varepsilon_i\right) \left(\sum_{i=1}^n \varepsilon_i \tilde{x}_i\right)\right) = \sigma_\varepsilon^2 \sum_{i=1}^n \alpha_i \tilde{x}_i = \sigma_\varepsilon^2. \quad (2.92)$$

Враховуючи (2.90 — 2.92), маємо:

$$E\left(\sum_{i=1}^n e_i^2\right) = (n-2)\sigma_\varepsilon^2 \Rightarrow E\left(\frac{\sum_{i=1}^n e_i^2}{n-2}\right) = \sigma_\varepsilon^2. \quad (2.93)$$

Отже, ми отримали оцінку дисперсії випадкової величини ε , яку часто позначають кількома позначеннями, а саме:

$$\hat{\sigma}_\varepsilon^2 = S^2 = \hat{\sigma}^2 = \frac{\sum_{i=1}^n e_i^2}{n-2}.$$

Середнє квадратичне відхилення оцінки дисперсії відповідно має вигляд:

$$\sqrt{\hat{\sigma}_\varepsilon^2} = \sqrt{S^2} = \sqrt{\hat{\sigma}^2} = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n-2}}.$$

Оскільки випадкова величина ε — неспостежувана, то і її дисперсію неможливо обчислити, тому на практиці дисперсія випадкової величини ε замінюється на свою оцінку.

2.12. ПОБУДОВА ІНТЕРВАЛІВ ДОВІРИ ДЛЯ ПАРАМЕТРІВ β_0, β_1

2.12.1. Поняття про t -тест Ст'юдента. Перевірка нуль-гіпотези за допомогою t -тесту Ст'юдента

Як було показано в параграфі 2.10, випадкові параметри b_0 і b_1 розподілені за нормальним законом розподілу з відповідним математичним сподіванням та дисперсією, що формально можна записати таким чином:

$$b_0 \sim N\left\{\beta_0, \frac{\sigma_\varepsilon^2 \sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2}\right\}; \quad (2.94)$$

$$b_1 \sim N\left\{\beta_1, \frac{\sigma_\varepsilon^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right\}. \quad (2.95)$$

У виразі (2.94), (2.95) дисперсія параметрів b_0 і b_1 , загалом, невідома, тому що вона залежить від дисперсії помилок σ_ε^2 випадкової величини ε , яку не можна спостерігати. Як уже говорилося вище (параграф 2.11), невідома дисперсія σ_ε^2 замінюється на свою оцінку $\hat{\sigma}_\varepsilon^2$. Таким чином, і для параметрів b_0 і b_1 дійсна дисперсія замінюється на свою оцінку:

$$\hat{\sigma}_{b_0}^2 = \hat{\sigma}_\varepsilon^2 \left(\frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2} \right); \quad (2.96)$$

$$\hat{\sigma}_{b_1}^2 = \frac{\hat{\sigma}_\varepsilon^2}{\sum_{i=1}^n (x_i - \bar{x})^2}. \quad (2.97)$$

Встановлені значення параметрів вибіркової моделі є оцінками параметрів узагальненої регресійної моделі. Постає питання: чи можна знайти інтервали довіри для параметрів узагальненої моделі, тобто інтервали, у які із заданою імовірністю потрапляють їхні значення? Справді, така можливість є і широко використовується у практичних дослідженнях. Інтервали довіри можна будувати за допомогою кількох тестів.

Щоб зрозуміти суть побудови інтервалів довіри для параметрів регресійної моделі, розглянемо спочатку один з найпоширеніших тестів — *t-тест Ст'юдента у загальному випадку*.

Розглянемо довільну випадкову величину, розподілену за нормальним законом розподілу з математичним сподіванням a та дисперсією σ^2 . Як ми вже знаємо, нормально розподілену випадкову величину з математичним сподіванням a і дисперсією σ^2 можна звести до нормально розподіленої величини з математичним сподіванням 0 і дисперсією 1 шляхом перетворення:

$$z_i = \frac{x_i - a}{\sigma_x}, \quad (2.98)$$

де x_i — випадкова нормально розподілена величина з математичним сподіванням a і дисперсією σ^2 , а z_i — нормально розподілена випадкова величина з математичним сподіванням 0 і дисперсією 1 .

Ще раз підкреслимо, що перетворення (2.98) дає нам нормально розподілені величини. Наприклад, для параметра b_1 таке перетворення має вигляд:

$$z_i = \frac{b_1 - \beta_1}{\sigma_{b_1}},$$

де b_1 — нормально розподілена випадкова величина з математичним сподіванням β_1 і дисперсією $\sigma_{b_1}^2$.

Якщо у формулі (2.98) ми використаємо замість невідомої дійсної дисперсії σ_ε^2 її оцінку $\hat{\sigma}_\varepsilon^2$, то при невеликій кількості вибірових даних

($n < 30$) ми перейдемо до іншого *t-перетворення*, тобто у загальному випадку матимемо:

$$t_i = \frac{x_i - a}{\hat{\sigma}_x}, \quad (2.99)$$

де x_i — нормально розподілена величина з математичним сподіванням a і дисперсією σ_x^2 ; t_i — випадкова величина, розподілена за *t-законом розподілу Ст'юдента* з $(n - 1)$ ступенем вільності, де ступені вільності розраховуються за виразом оціненої дисперсії, а саме:

$$\hat{\sigma}_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}.$$

Перетворення (2.99) можна записати і для випадкової величини \bar{x} , якщо мати на увазі, що вона розподіляється за нормальним законом розподілу з математичним сподіванням a і дисперсією σ_x^2 , тобто $\bar{x} \sim N(a, \sigma_x^2)$. Тоді *t-перетворення* матиме вигляд:

$$t = \frac{\bar{x} - a}{\sqrt{\frac{\hat{\sigma}_x^2}{n}}}, \quad (2.100)$$

де t — випадкова величина, яка розподіляється за *t-законом розподілу Ст'юдента* з $(n - 1)$ ступенями вільності.

t-розподіл — це симетричний розподіл із середнім нуль і дисперсією $\frac{n-1}{n-3}$, яка наближається до 1 , коли n — велике. Зрозуміло, що *t-розподіл* наближається до нормального закону розподілу, якщо $n \rightarrow \infty$.

Для використання t-тесту Ст'юдента необхідно:

- **обрати бажаний рівень значимості (від 1 до 10%);**
- **визначити кількість ступенів вільності.**

Володіючи цією інформацією, ми можемо визначити критичне значення t , яке поділяє усю множину значень на дві підмножини: множину, яку ми відкидаємо, і множину, яку ми приймаємо при заданому рівні значимості.

Приклад. Ми тестуємо нуль-гіпотезу, а саме:

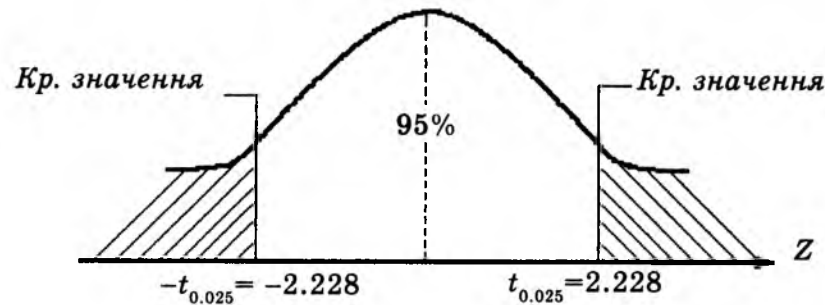
$$H_0: a = a_0$$

проти альтернативної гіпотези

$$H_1: a \neq a_0$$

для вибірки з невідомою дисперсією.

Оберемо рівень значимості 5%. Критичне значення t знаходимо за таблицями t -розподілу Ст'юдента. Рядки цієї таблиці мають різні ступені вільності, а колонки відповідають різним рівням значимості. Наприклад, для 10 ступенів вільності і 5% рівня значимості відповідні критичні значення (мал. 2.11) дорівнюють $t_1 = -2.228$; $t_2 = 2.228$.



Малюнок 2.11. Критична зона для випадкової величини за t -розподілом Ст'юдента при 10 ступенях вільності і обраному рівні значимості 5%.

Зона, яку ми приймаємо, є такою:

$$-2.228 \leq \left\{ t = \frac{(\bar{x} - a_0) / \sqrt{\hat{\sigma}_x^2 / n}} \right\} \leq 2.228.$$

За даними вибірки знаходимо \bar{x} і обчислюємо значення t -статистики:

$$t^* = \frac{\bar{x} - a_0}{\sqrt{\hat{\sigma}_x^2 / n}}.$$

Якщо значення t^* потрапляє в критичну зону (в одну із заштрихованих), то можливі 2 випадки.

1. Нуль-гіпотеза правильна, але сталася малоймовірна подія.
2. Нуль-гіпотеза неправильна.

Ми беремо до уваги найпростіший випадок (випадок 2), тобто відкидаємо нуль-гіпотезу.

2.12.2. t -тест Ст'юдента для перевірки на значимість параметрів b_0 і b_1 , визначених за методом найменших квадратів

Як зазначалося вище, параметри, визначені за методом найменших квадратів, розподіляються за нормальним законом розподілу, який формалізовано можна записати таким чином:

$$b_0 \sim N(\beta_0, \text{невідома дисперсія } \sigma_{b_0}^2 = \sigma_\varepsilon^2 \frac{\sum_{i=1}^n x_i^2}{n \cdot \sum_{i=1}^n (x_i - \bar{x})^2});$$

$$b_1 \sim N(\beta_1, \text{невідома дисперсія } \sigma_{b_1}^2 = \sigma_\varepsilon^2 \frac{1}{\sum_{i=1}^n (x_i - \bar{x})^2}).$$

Нагадаємо, що у загальному випадку $\sigma_{b_0}^2$ і $\sigma_{b_1}^2$ невідомі, тому що не можна обчислити σ_ε^2 , адже випадкові величини ε взагалі є неспостережуваними. Але ми можемо обчислити оцінку дисперсій $\sigma_{b_0}^2$ і $\sigma_{b_1}^2$, тобто знайти:

$$\hat{\sigma}_{b_0}^2 = \hat{\sigma}_\varepsilon^2 \cdot \frac{\sum_{i=1}^n x_i^2}{n \cdot \sum_{i=1}^n (x_i - \bar{x})^2}; \quad \hat{\sigma}_{b_1}^2 = \hat{\sigma}_\varepsilon^2 \cdot \frac{1}{\sum_{i=1}^n (x_i - \bar{x})^2},$$

де $\hat{\sigma}_\varepsilon^2 = \frac{\sum_{i=1}^n e_i^2}{n - k}$; k — кількість оцінених параметрів (у разі простої регресії $k = 2$).

Далі будемо t -статистику для кожного параметра:

$$t = \frac{b_i - \beta_i^*}{\sqrt{\hat{\sigma}_{b_i}^2}} \text{ з } (n - k) \text{ ступенями вільності,} \quad (2.101)$$

де b_i — оцінка параметра β_i , отримана за методом найменших квадратів;
 β_i^* — гіпотетичне значення, якого має набути параметр $\beta_i (H_0: \beta_i = \beta_i^*)$;
 $\hat{\sigma}_{b_i}^2$ — оцінка дисперсії параметра b_i (з регресії);
 n — розмір вибірки (кількість спостережень);
 k — загальна кількість оцінених параметрів ($k=2$ у нашій моделі, бо ми використовуємо 2 ступені вільності, щоб оцінити 2 параметри b_0 і b_1).

У економетриці поширеною формою нуль-гіпотези є така:

$$H_0: \beta_i^* = 0$$

проти альтернативної

$$H_1: \beta_i^* \neq 0$$

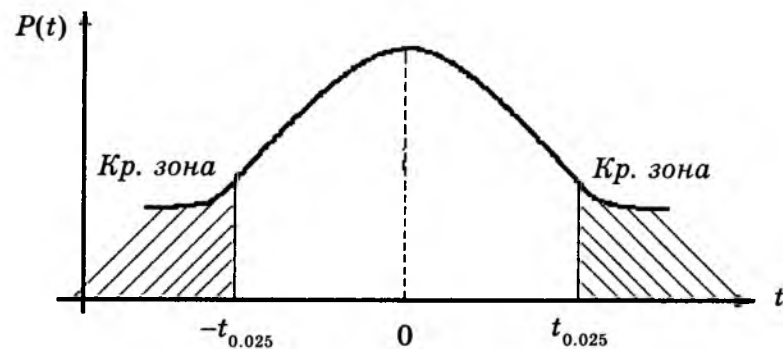
У цьому разі t -статистика для параметрів має вигляд:

$$t^* = \frac{b_i}{\hat{\sigma}_{b_i}^2}. \quad (2.102)$$

Її значення порівнюємо з табличним значенням, яке дає змогу знайти критичну зону з $(n - k)$ ступенями вільності.

Якщо значення t^* потрапляє не в критичну зону, тобто $-t_{\alpha/2} < t^* < t_{\alpha/2}$ з $(n - k)$ ступенями вільності і при $\alpha \cdot 100\%$ -ному рівні значимості (у загальному випадку), ми можна стверджувати, що з імовірністю $(1 - \alpha)$ оцінка b_i є статистично незначимою (тобто ми приймаємо нуль-гіпотезу).

Двовимірний тест для нуль-гіпотези $H_0: \beta_i^* = 0$ показано на мал. 2.12 ($\alpha = 0.05$).



Малюнок 2.12. Двовимірний тест для нуль-гіпотези з 5%-ним рівнем значимості

З мал.2.12 видно, що коли значення t^* потрапляє в критичну (заштриховану) зону, то ми відкидаємо нуль-гіпотезу.

Використовувана t -статистика (2.102) є ніщо інше, як відношення b_i до оцінки свого стандартного відхилення, або, інакше кажучи, до свого середньоквадратичного відхилення.

У багатьох програмах поряд із значенням b_i видають значення оцінки стандартного відхилення та відношення між b_i і цією оцінкою. Це відношення називають t -значенням для відповідного параметра. Якщо воно перевищує критичне значення, яке ми знаходимо за таблицею, то приймаємо гіпотезу $H_1: \beta_i \neq 0$ і оцінюємо відповідний параметр як статистично значимий. У разі простої лінійної регресії це також означає, що x має значимо впливає на зміну y .

T -тест може бути спрощеним. Якщо уважно подивитись на t -таблицю Ст'юдента, то можна помітити, що значення t змінюється дуже повільно, коли кількість ступенів вільності $(n - k)$ більша, ніж 8. Наприклад, $t_{0.025}$ для 8 ступенів вільності дорівнює 2.3, і прямує до 1.96, коли $(n - k) \rightarrow \infty$. Справді, різниця між 2.3 і 1.96 є незначною. Тому для будь-якого значення $(n - k) > 8$ можна вважати, що критичне значення t приблизно дорівнює 2.

Тому за спрощеним тестом ми відкидаємо нуль-гіпотезу, якщо $t^* > 2$. Іншими словами:

$$t^* > 2, \text{ якщо } b_i > 2 \hat{\sigma}_{b_i}^2 \text{ або } \hat{\sigma}_{b_i}^2 < b_i/2.$$

Звідси ми можемо зробити висновок:

нуль-гіпотезу відкидаємо, якщо $t^* > 2$

або

нуль-гіпотезу відкидаємо, якщо $\hat{\sigma}_{b_i}^2 < b_i/2$.

Приклад. Нехай ми маємо вибірку з 20 спостережень за значеннями ВВП (x) і витрат на споживання (y). Припустимо, що наявність лінійного зв'язку, отже, і лінійна регресія після оцінки параметрів матиме вигляд:

$$y = 100 + 0.7x.$$

Нехай, крім того, нам відомі всі характеристики та дані, які ми не наводимо.

Для перевірки параметрів на значимість використаємо t -тест Ст'юдента, а для цього знайдемо спочатку оцінку дисперсії кожного параметра:

$$\hat{\sigma}_{b_0}^2 = \hat{\sigma}_\varepsilon^2 \frac{\sum_{i=1}^n x_i^2}{n \cdot \sum_{i=1}^n (x_i - \bar{x})^2};$$

$$\hat{\sigma}_{b_1}^2 = \hat{\sigma}_\varepsilon^2 \frac{1}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

Далі розраховуємо значення t -статистики, наприклад, для b_1 :

$$t^* = \frac{b_1}{\hat{\sigma}_{b_1}^2} = \frac{0.70}{0.21} \approx 3.3$$

і перевіряємо нуль-гіпотезу

$$H_0: \beta_1 = 0$$

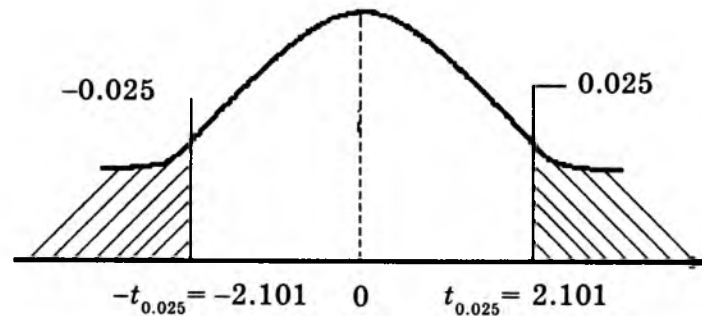
проти альтернативної гіпотези

$$H_1: \beta_1 \neq 0.$$

Критичні значення t для $(n - k) = 18$ дорівнюють відповідно:

$$\begin{aligned} t_1 &= -t_{0.025} = -2.1; \\ t_2 &= t_{0.025} = 2.1. \end{aligned}$$

Оскільки $t^* > t_{0.025}$, ми відкидаємо нуль-гіпотезу і робимо висновок, що параметр b_1 значно відрізняється від нуля, тобто знайдена його оцінка b_1 є статистично значимою (мал. 2.13.)



Малюнок 2.13. Встановлення критичної зони за t -статистикою Ст'юдента для параметра b_1

За спрощеним t -тестом маємо:

$t^*(b_1) = 3.3 > 2$, параметр b_1 є статистично значимий.

Повернемося тепер до нашого прикладу про залежність обсягів реалізації продукції від витрат на рекламу та перевіримо на значимість кожний параметр. Нагадаємо, що знайдена модель має вигляд: $y = 10 + 3x$.

Задамо рівень значимості $\alpha = 0.025$ (2.5%). Нагадаємо, що кількість спостережень у нашому прикладі дорівнює п'яти, отже, кількість ступенів вільності дорівнює відповідно трьом ($5 - 2$). За t -таблицею розподілу Ст'юдента знайдемо $t_{\alpha/2}$ - критичне з 3 ступенями вільності. Воно дорівнюватиме: $t_{0.0125} = \pm 3.182$.

Розрахуємо t -відношення для кожного параметра, встановивши спочатку оцінку дисперсії та середньоквадратичного відхилення.

Для параметра b_1 маємо:

$$\hat{\sigma}_{b_1}^2 = \frac{\hat{\sigma}_\varepsilon^2}{\sum_{i=1}^5 (x_i - \bar{x})^2} = \frac{10}{330} = 0.03 \Rightarrow \hat{\sigma}_{b_1} = 0.174,$$

для параметра b_0 :

$$\hat{\sigma}_{b_0}^2 = \frac{\hat{\sigma}_\varepsilon^2 \sum_{i=1}^5 x_i^2}{n \sum_{i=1}^5 (x_i - \bar{x})^2} = \frac{10 * 610}{3 * 10} = 3.697 \Rightarrow \hat{\sigma}_{b_0} = 1.92.$$

Тепер знайдемо t -відношення для кожного параметра.

$t_1 = \frac{b_1}{\hat{\sigma}_{b_1}} = \frac{3}{0.174} = 17.24 > t_{кр} = 3.182$, отже, параметр b_1 — статистично значимий.

Відповідно для параметра b_0 маємо:

$t_0 = \frac{b_0}{\hat{\sigma}_{b_0}} = \frac{10}{1.92} = 5.2 > t_{кр} = 3.182$, отже, параметр b_0 — також статистично значимий.

2.12.3. Знаходження інтервалів довіри для параметрів β_0 і β_1 за t -тестом Ст'юдента

Для того щоб визначити, як же параметри b_1 і b_0 пов'язані з параметрами β_0 і β_1 , потрібно побудувати інтервали довіри для параметрів.

Процедура побудови інтервалів довіри є аналогічною процедурі тестування значимості знайдених параметрів простої вибіркової лінійної регресії.

Спочатку розрахуємо t -статистику для кожного з параметрів:

$$t = \frac{b_i - \beta_i}{\hat{\sigma}_{b_i}}.$$

Потім обираємо рівень значимості (α або $\alpha \cdot 100\%$). Відповідно рівень довіри (*confidence level*) дорівнюватиме: $((1-\alpha)$ або $(1-\alpha) \cdot 100\%$). За t -таблицею Ст'юдента знаходимо значення $\pm t_{\alpha/2}$ з $(n-2)$ ступенями вільності. Тоді можна записати:

$$P(-t_{\alpha/2} < t < +t_{\alpha/2}) = (1 - \alpha). \quad (2.103)$$

У економетриці найчастіше використовують 95%-ний рівень довіри, що дозволяє переписати (2.103) у вигляді:

$$P(-t_{0.025} < t < +t_{0.025}) = 0.95 .$$

Проводячи заміну $t = (b_i - \beta_i) / \hat{\sigma}_{b_i}$, отримуємо:

$$P\left\{-t_{\alpha/2} < \frac{b_i - \beta_i}{\hat{\sigma}_{b_i}} < +t_{\alpha/2}\right\} = (1 - \alpha) \quad (2.104)$$

або:

$$P\left\{b_i - t_{\alpha/2} \cdot \hat{\sigma}_{b_i} < \beta_i < b_i + t_{\alpha/2} \cdot \hat{\sigma}_{b_i}\right\} \text{ з } n - 2 \text{ ступенями вільності}$$

або:

$$\beta_i = b_i \pm t_{\alpha/2} \cdot \hat{\sigma}_{b_i} \text{ з } (n - 2) \text{ ступенями вільності.}$$

Для 95%-ного рівня довіри це можна записати таким чином:

$$\beta_i = b_i \pm t_{0.025} \cdot \hat{\sigma}_{b_i} .$$

Для ілюстрації повернемося до нашого прикладу про залежність обсягів реалізації продукції від витрат на рекламу та побудуємо інтервали довіри для кожного параметра. Нагадаємо, що знайдена модель має вигляд: $y = 10 + 3x$.

Задамо рівень значимості $\alpha = 0.025$ (2.5%). Оскільки кількість спостережень у нашому прикладі дорівнює п'яти, отже, кількість ступенів вільності відповідно дорівнюватиме трьом (5 - 2). За t -таблицею розподілу Ст'юдента знайдемо $t_{\alpha/2}$ -критичне з 3-а ступенями вільності. Воно дорівнюватиме: $t_{0.0125} = \pm 3.128$.

Розрахуємо дисперсії кожного параметра:

$$\hat{\sigma}_{b_1}^2 = \frac{\hat{\sigma}_\varepsilon^2}{\sum_{i=1}^5 (x_i - \bar{x})^2} = \frac{10}{330} = 0.03 \Rightarrow \hat{\sigma}_{b_1} = 0.174 ;$$

$$\hat{\sigma}_{b_0}^2 = \frac{\hat{\sigma}_\varepsilon^2 \sum_{i=1}^5 x_i^2}{n \sum_{i=1}^5 (x_i - \bar{x})^2} = \frac{10 \times 610}{3 \times 10} = 3.697 \Rightarrow \hat{\sigma}_{b_0} = 1.92 .$$

Тепер знайдемо інтервали довіри для параметра β_1 :

$$\beta_1 = b_1 \pm t_{\alpha/2} \hat{\sigma}_{b_1} = 3 \pm 3.182 \times 0.174,$$

що означає:

$$p(2.45 < \beta_1 < 3.55) = 0.975 .$$

Відповідно, для параметра β_0 маємо:

$$\beta_0 = b_0 \pm t_{\alpha/2} \hat{\sigma}_{b_0} = 10 \pm 3.182 \times 1.92,$$

або

$$p(3.88 < \beta_0 < 16.12) = 0.975 .$$

Зробимо висновок, що інтервали довіри для параметра β_1 сталіші, ніж для параметра β_0 .

2.12.4. Тест Фішера для перевірки нуль-гіпотези $\beta_1 = 0$

Тестування гіпотези, чи дорівнює $\beta_1 = 0$, можна розглядати ще й по-іншому, а саме: вона дає відповідь, чи справді незалежна змінна x впливає на значення незалежної y . Тобто у разі простої лінійної регресії перевірка нуль-гіпотези $\beta_1 = 0$ аналогічна перевірці адекватності моделі за F -критерієм Фішера. Покажемо це.

Ми вже знаємо, що відношення

$$\frac{b_1 - \beta_1}{\sigma_\varepsilon / \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

розподіляється за нормальним законом розподілу з математичним сподіванням 0 і дисперсією 1.

За визначенням закону розподілу χ^2 маємо:

$\frac{(b_1 - \beta_1)^2}{\sigma_\varepsilon^2 / \sum_{i=1}^n (x_i - \bar{x})^2} \sim \chi^2$, що відповідає χ^2 -розподілу з 1 ступенем вільності, а

$$\sum_{i=1}^n (e_i^2 / (n-2)) / \sigma_\varepsilon^2 \sim \chi^2,$$

що відповідає χ^2 -розподілу з $(n-2)$ ступенями вільності незалежно від b_i .

З математичної статистики відомо, що відношення двох випадкових величин, розподілених за законом χ^2 і поділених на число їхніх ступенів вільності, розподіляється за законом розподілу Фішера з відповідними ступенями вільності.

Таким чином, величина

$$F = \frac{(b_1 - \beta_1)^2 \sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n e_i^2 / (n-2)} \sim F(1, n-2) \quad (2.105).$$

розподіляється за F -розподілом Фішера з $(1, n-2)$ ступенями вільності.

Якщо $\beta_1=0$, тоді

$$F = \frac{b_1^2 \sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n e_i^2 / (n-2)} \sim F(1, n-2). \quad (2.106)$$

Як показано раніше, суму квадратів, що пояснює регресію, можна подати у вигляді:

$$SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = b_1^2 \sum_{i=1}^n (x_i - \bar{x})^2, \quad (2.107)$$

а сума квадратів залишків відповідно дорівнюватиме:

$$SSE = \sum_{i=1}^n e_i^2. \quad (2.108)$$

З урахуванням цього вираз (2.106) можна переписати у звичайному для F -критерію Фішера вигляді перевірки на адекватність моделі:

$$F = \frac{SSR / 1}{SSE / (n-2)} \sim F(1, n-2).$$

Тестування за критерієм Фішера значимості змінної x , або адекватності моделі, складається з певних етапів.

Спочатку формулюється нуль-гіпотеза: $H_0: \beta_1 = 0$.

Задаємо $\alpha \cdot 100\%$ — рівень значимості (наприклад, 5%).

Обчислюємо F -відношення.

За таблицями F -розподілу Фішера знаходимо F -критичне значення при заданому рівні значимості (або помилки) та з $(1, n-2)$ ступенями вільності (для простої лінійної регресії).

Цю гіпотезу відкидаємо з $\alpha \cdot 100\%$ -ним ризиком помилитися (наприклад 5%-ний ризик), якщо:

$$F = \frac{SSR / 1}{SSE / (n-2)} > F_{0.95}(1, n-2),$$

де $F_{0.95}$ — значення F при 5%-ному ризику помилки (знаходимо за таблицями F -критерію Фішера з відповідними ступенями вільності і заданим рівнем значимості).

2.12.5. Порівняння F -критерію Фішера з t -критерієм Ст'юдента

F -критерій Фішера, як і t -критерій Ст'юдента, використовують для оцінки значимості x . Це одні й ті самі тести, але записані в різній формі. Для того, щоб визначити еквівалентність цих двох тестів, нагадаємо **тест Ст'юдента**:

відкинути $H_0: \beta_1 = 0$ з $\alpha \cdot 100\%$ -ним ризиком помилитись, якщо:

$$\left| \frac{b_1}{\hat{\sigma}_\varepsilon / \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} \right| > t_{\alpha/2}, \quad (2.110)$$

де $t_{\alpha/2}$ — критичне значення з таблиці t -розподілу Ст'юдента з $(n-2)$ ступенями вільності при заданому рівні значимості $\alpha \cdot 100\%$.

Тест Фішера має вигляд:

відкинути $H_0: \beta_1 = 0$ з $\alpha \cdot 100\%$ -ним ризиком, якщо

$$\frac{b_1^2 \sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n e_i^2 / (n-2)} > F_{(1-\alpha)}(1, n-2). \quad (2.111)$$

Якщо порівняти ліві частини нерівностей (2.110) та (2.111), то легко помітити, що друга є квадратом першої. Таким чином, виконується тождество:

$$F\text{-відношення} = (t\text{-відношенню})^2. \quad (2.112)$$

З математичної статистики також відомо, що

$$F\text{-критичне значення} = (t\text{-критичному значенню } t)^2. \quad (2.113)$$

З (2.112) і (2.113) випливає, що два тести еквівалентні.

Існує ще одна можливість розгляду F -критерію Фішера. Як ми показували раніше:

$$SSR = R^2 SST = R^2 \sum_{i=1}^n (y_i - \bar{y})^2;$$

$$SSE = (1 - R^2) SST = (1 - R^2) \sum_{i=1}^n (y_i - \bar{y})^2,$$

де R^2 — коефіцієнт детермінації, що є квадратом коефіцієнта кореляції r .

Згідно з (2.109) F -статистику можна записати у вигляді:

$$F = \frac{R^2 / 1}{(1 - R^2) / (n - 2)}. \quad (2.114)$$

Користуючись зв'язком між F і t (див. 2.113), можна отримати інший вираз для t статистики, яка розподіляється за законом розподілу Ст'юдента з $(n-2)$ ступенями вільності:

$$t = \frac{r\sqrt{(n-2)}}{(1-R^2)}. \quad (2.115)$$

Таким чином, ми отримали три різні версії одного і того самого тесту.

Тест (див. 2.115), який використовує r , можна використати для перевірки значимості коефіцієнта кореляції. Тест щодо b_1 можна використовувати, щоб визначити значимість нахилу регресії.

Проте всі вони (кожен по-своєму) відповідають на питання адекватності побудованої вибіркової простої регресійної моделі.

2.12.6. t -тест для оцінки значимості коефіцієнта кореляції

Ми знаємо, що коефіцієнт кореляції r вимірює щільність зв'язку між двома змінними. У випадку простої лінійної регресії він вимірює щільність зв'язку між незалежною змінною x і залежною змінною y . Задамося цілком встановити:

як пов'язані коефіцієнт кореляції r між вибірковими значеннями x і y з коефіцієнтом кореляції ρ всієї генеральної сукупності;

як перевірити значимість коефіцієнта кореляції і як для цього використати t -статистику (2.115).

Спробуємо спочатку відповісти на друге запитання, що автоматично дає відповідь і на перше запитання. Використаємо t -статистику у вигляді (2.115) для визначення, чи статистично значимо коефіцієнт кореляції всієї сукупності ρ відрізняється від нуля.

Спочатку розглянемо найпростіший випадок, а саме t -тест для оцінки значимості коефіцієнта кореляції (з припущенням, що $\rho = 0$).

Сформулюємо нуль-гіпотезу:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \sim t(n-2), \quad (2.116)$$

де r — вибірковий коефіцієнт кореляції між x і y ;

n — кількість спостережень.

Величина t^* розподілена за t -розподілом Ст'юдента з $(n-k)$ ступенями вільності (нагадаємо, що для простої лінійної регресії $k = 2$).

Розраховане значення t^* порівнюємо з критичним значенням $t_{\alpha/2}$ при $\alpha=100\%$ -ному рівні значимості й $(n-2)$ ступенях вільності. Якщо $t^* > t_{\alpha/2}$, відкидаємо нуль-гіпотезу $H_0: \rho = 0$ і приймаємо гіпотезу $H_1: \rho \neq 0$.

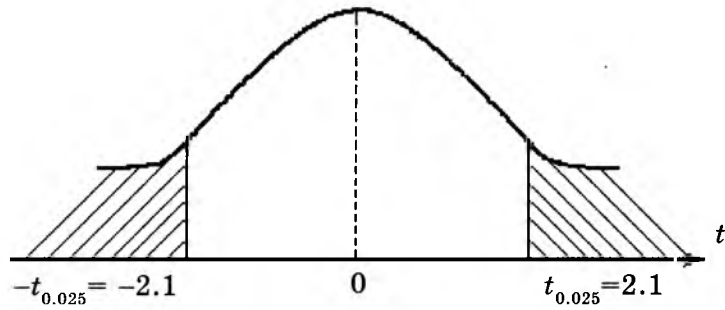
Приклад. Маємо вибірку значень x і y , яка складається з 20 спостережень; коефіцієнт кореляції $r = 0.82$. Ми хочемо перевірити при 5%-ному рівні значимості, чи значимо коефіцієнт кореляції відрізняється від нуля. Тобто перевіряємо нуль-гіпотезу:

$$H_0: \rho = 0$$

проти альтернативної:

$$H_1: \rho \neq 0.$$

$$t^* = \frac{0.82\sqrt{20-2}}{\sqrt{1-(0.82)^2}} = 6.04.$$



Малюнок 2.14. Перевірка на значимість коефіцієнта кореляції за t -тестом Ст'юдента

За t -таблицями Ст'юдента знаходимо теоретичне значення t з 18 ступенями вільності і 5%-ним рівнем значимості, яке дорівнює $t_{0.025} = 2.1$ (мал. 2.14). Оскільки $t^* > t$, ми відкидаємо нуль-гіпотезу і робимо висновок, що $\rho \neq 0$.

Трансформований у z -тест t -тест для оцінки значимості коефіцієнта кореляції (з припущенням, що $\rho \neq 0$)

Якщо $\rho \neq 0$, то t -розподіл з коефіцієнтом ρ є несиметричним. Тобто t -статистику не можна використовувати для перевірки таких гіпотез, як $\rho = 0.2$ або $\rho_1 - \rho_2 = 0$.

Цю проблему було вирішено в 1921 році Фішером, який трансформував ρ у величину h :

$$h = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right),$$

розподілену нормально з математичним сподіванням

$$E(h) = \mu_h = \frac{1}{2} \ln \left(\frac{1+\rho}{1-\rho} \right)$$

і дисперсією

$$\sigma_h^2 = \frac{1}{n-3},$$

що формалізовано можна записати

$$h \sim N \left(\left(\mu_h = \frac{1}{2} \ln \left(\frac{1+\rho}{1-\rho} \right) \right); \sigma_h = \sqrt{\frac{1}{n-3}} \right).$$

Як будь-яку нормально розподілену величину, її можна стандартизувати, тобто звести до нормально розподіленої величини з математичним сподіванням 0 і дисперсією 1 шляхом перетворення:

$$z_i = \frac{h_i - \mu_h}{\sigma_h} \sim N_{(0,1)}.$$

Приклад. 3 вибірки, що складається з 28 спостережень, знайдено, що $r = 0.70$.

Чи варто відкинути гіпотезу, що $\rho = 0.6$ (в генеральній сукупності) з ризиком помилитися 5%?

Тобто перевіряємо нуль-гіпотезу:

$$H_0: \rho = 0.60$$

проти

$$H_1: \rho \neq 0.60.$$

Для цього обчислимо величину:

$$h = 1.1513 \log_{10} \left(\frac{1+0.7}{1-0.7} \right) = 0.9730;$$

$$\mu_h = 1.1513 \log_{10} \left(\frac{1+0.6}{1-0.6} \right) = 0.6932;$$

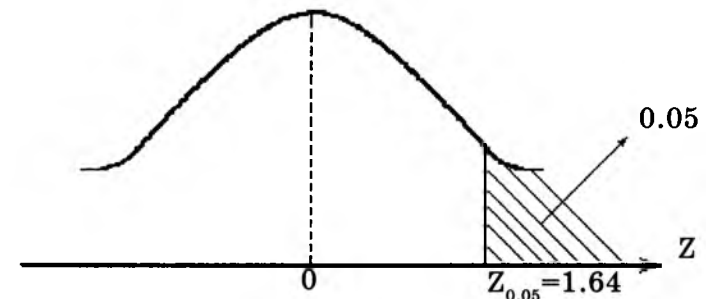
$$\sigma_h = \sqrt{\frac{1}{n-3}} = \sqrt{\frac{1}{25}} = 0.2.$$

Далі розрахуємо значення z -статистики:

$$z^* = \frac{h_i - \mu_h}{\sigma_h} = \frac{0.9730 - 0.6932}{0.2} = 1.28.$$

З огляду на те, що заданий рівень помилки $\alpha = 0.05$, за таблицями нормального розподілу знаходимо критичне значення $z_{0.05}$.

Табличне значення $z_{0.05} = 1.64$ (мал. 2.15).



Малюнок 2.15. Функція густини нормального закону розподілу

Оскільки $z^* < z_{0.05}$, приймаємо нуль-гіпотезу про те, що коефіцієнт кореляції усієї сукупності $\rho = 0.6$. Тобто розраховане значення $r = 0.7$ є статистично незначимим.

2.13. ПРОГНОЗУВАННЯ ЗА МОДЕЛЯМИ ПРОСТОЇ ЛІНІЙНОЇ РЕГРЕСІЇ

Повернемося до нашого прикладу. Ми маємо ряд значень незалежної змінної x (витрати на рекламу) та залежної змінної y (обсяги продажу продукції). На підставі цих даних ми побудували лінійну модель та оцінили параметри за методом найменших квадратів. Якщо наша модель адекватна, можемо прогнозувати зміну обсягів продаж продукції залежно від зміни витрат на рекламу. При цьому ми можемо отримати два типи прогнозів: точкові та інтервальні. Точковий прогноз дає значення залежної змінної, наприклад, для відповідного значення x_{n+1} з побудованої вибіркової моделі:

$$\hat{y}_{n+1} = b_0 + b_1 x_{n+1}. \quad (2.115)$$

При цьому, виходячи з узагальненої моделі, дійсне значення y для прогнозного періоду дорівнюватиме:

$$y_{n+1} = \beta_0 + \beta_1 x_{n+1} + \varepsilon_{n+1}, \quad (2.116)$$

де ε_{n+1} — значення випадкової величини, не спостережуваної в $(n+1)$ періоді.

Отже, прогнозне значення \hat{y}_{n+1} є оцінкою дійсного значення змінної y_{n+1} . Таким чином, за нашою вибірковою моделлю легко можна знаходити будь-яке прогнозне значення. Зазначимо, що таке прогнозне значення буде точковим. Як, використовуючи отриманий точковий прогноз, знайти інтервали довіри для дійсного значення залежної змінної, тобто побудувати інтервал, у який з певною заданою імовірністю потрапляє дійсне значення залежної змінної? Щоб відповісти на це запитання, розглянемо спочатку помилку прогнозного значення, знайдемо її математичне сподівання та дисперсію, а далі перейдемо до побудови інтервалів довіри.

Помилка прогнозу обчислюється за виразом:

$$e_{n+1} = y_{n+1} - \hat{y}_{n+1} = \varepsilon_{n+1} - (b_0 - \beta_0) - (b_1 - \beta_1)x_{n+1}. \quad (2.117)$$

Математичне сподівання помилки прогнозу:

$E(e_{n+1}) = 0$, бо за припущенням $E(\varepsilon_{n+1}) = 0$, β_0 та β_1 є константами, а математичне сподівання кожного знайденого параметра відповідно дорівнює:

$$E(b_0) = \beta_0, \quad E(b_1) = \beta_1.$$

Піднесемо до квадрата обидві частини (2.117), встановивши математичне сподівання цього виразу. Отримаємо значення дисперсії помилки:

$$\text{var}(e_{n+1}) = E(e_{n+1}^2) = \text{var}(\varepsilon_{n+1}) + \text{var}(b_0) + x_{n+1}^2 \text{var}(b_1) + 2x_{n+1} \text{cov}(b_0, b_1). \quad (2.118)$$

Дві інші коваріації зникають, бо за гіпотезою, що ε_{n+1} є незалежною від $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$, її коваріації з $(b_0 - \beta_0)$ та $(b_1 - \beta_1)$ дорівнюватимуть нулю.

Раніше ми вже показували, що дисперсія параметрів відповідно дорівнює:

параметра b_1 :

$$\text{var}(b_1) = \sigma_{b_1}^2 = \frac{\sigma_\varepsilon^2}{\sum (x_i - \bar{x})^2},$$

параметра b_0 :

$$\text{var}(b_0) = \sigma_{b_0}^2 = \sigma_\varepsilon^2 \left(\frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2} \right) = \sigma_\varepsilon^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right]$$

Покажемо, що $\text{cov}(b_0, b_1) = -\frac{\bar{x} \sigma_\varepsilon^2}{\sum (x_i - \bar{x})^2}$;

$$\begin{aligned} \text{cov}(b_0, b_1) &= E\{(b_0 - \beta_0)(b_1 - \beta_1)\} = E\{[\bar{\varepsilon} - (b_1 - \beta_1)\bar{x}](b_1 - \beta_1)\} = \\ &= -\bar{x} E\{(b_1 - \beta_1)^2\} \end{aligned}$$

тому, що $E\{(b_1 - \beta_1)\bar{\varepsilon}\} = 0$.

$$\text{Звідси маємо } \text{cov}(b_0, b_1) = -\bar{x} E\{(b_1 - \beta_1)^2\} = -\frac{\bar{x} \sigma_\varepsilon^2}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

Підставляючи у (2.118) явний вигляд дисперсій та коваріацій, отримаємо:

$$\begin{aligned} \text{var}(e_{n+1}) &= \sigma_\varepsilon^2 \left[1 + \frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2} + \frac{x_{n+1}^2}{\sum_{i=1}^n (x_i - \bar{x})^2} - \frac{2x_{n+1}\bar{x}}{\sum_{i=1}^n (x_i - \bar{x})^2} \right] = \\ &= \sigma_\varepsilon^2 \left[1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right] \end{aligned} \quad (2.119)$$

Як видно з (2.119), дисперсія помилки прогнозу буде мінімальною, якщо $x_{n+1} = \bar{x}$, та зростає нелінійно в міру того, як x_{n+1} віддаляється від свого середнього значення (\bar{x}).

З (2.117) видно, що e_{n+1} є функцією багатьох параметрів, розподілених за нормальним законом, звідки випливає, що помилка також розподілена нормально з математичним сподіванням 0 та дисперсією $\text{var}(e_{n+1})$.

Шляхом уже відомого перетворення зведемо випадкову величину до нормально розподіленої з математичним сподіванням 0 та дисперсією 1:

$$z = \frac{e_{n+1}}{\sigma_\varepsilon \sqrt{1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}}. \quad (2.120)$$

Замінюючи в (2.120) σ_ε на свою оцінку $\hat{\sigma}_\varepsilon = \sqrt{\sum_{i=1}^n e_i^2 / (n-2)}$, отримаємо:

$$\frac{y_{n+1} - \hat{y}_{n+1}}{\hat{\sigma}_\varepsilon \sqrt{1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}} \sim t(n-2). \quad (2.121)$$

Величина (2.121) буде вже розподілена за t -розподілом Ст'юдента з $(n-2)$ ступенями вільності, де єдина невідома — величина y_{n+1} . Для y_{n+1} знайдемо інтервал довіри вже за відомою схемою, задавши рівень значимості $\alpha 100\%$, а саме:

$$(b_0 + b_1 \cdot x_{n+1}) \pm t_{\alpha/2} \cdot \hat{\sigma}_\varepsilon \sqrt{1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}. \quad (2.122)$$

Таким чином, формула (2.122) дає нам інтервал довіри для дійсного значення залежної змінної.

На практиці більше застосовується побудова інтервалів довіри для математичного сподівання y_{n+1} , тобто побудова інтервалів довіри для:

$$E(y_{n+1}) = \beta_0 + \beta_1 x_{n+1}, \quad (2.123)$$

тому що, власне кажучи, немає великого сенсу прогнозувати точне значення y_{n+1} , беручи до уваги випадковий характер ε_{n+1} .

У такому разі, помилка прогнозу відповідно буде дорівнювати:

$$e_{n+1} = E(y_{n+1}) - \hat{y}_{n+1} = -(b_0 - \beta_0) - (b_1 - \beta_1)x_{n+1},$$

що дає:

$$\text{var}(e_{n+1}) = \sigma_\varepsilon^2 \left[\frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right] \quad (2.124)$$

і дозволяє побудувати інтервал довіри для $E(y_{n+1})$ при $\alpha 100\%$ -ному рівні значимості у вигляді:

$$b_0 + b_1 x_{n+1} \pm t_{\alpha/2} \cdot \hat{\sigma}_\varepsilon \sqrt{\frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}. \quad (2.125)$$

Проілюструємо наші викладки на прикладі залежності обсягів реалізації продукції від витрат на рекламу.

Нагадаємо, що побудована модель має вигляд: $y = 10 + 3x$.

Знайдемо прогнозне значення обсягів реалізації продукції при витратах на рекламу $x_j = 15$. Відповідно $y_j = 10 + 3 \times 15 = 55$. Побудуємо інтервали довіри для залежної змінної. Для цього знайдемо оцінку дисперсії залежної змінної, яка дорівнює, як ми показували раніше, оцінці дисперсії помилки:

$$\hat{\sigma}_{e_j}^2 = \hat{\sigma}_{y_j}^2 = 3.33 \left[1 + \frac{1}{5} + \frac{(15-10)^2}{110} \right] = 4.757 \Rightarrow \hat{\sigma}_{y_j} = 2.18.$$

Тепер задамо рівень значимості, наприклад $\alpha = 100\% = 2,5\%$. За таблицею t -розподілу Ст'юдента знаходимо $t_{\alpha/2}$, критичне при $(5 - 2 = 3)$ ступенях вільності, яке дорівнюватиме $t_{\alpha/2} = \pm 3.182$.

Відповідно інтервал довіри для залежної змінної дорівнюватиме

$$y_j = 55 \pm 3.183 \times 2.18.$$

2.14. ВЛАСТИВОСТІ МЕТОДУ НАЙМЕНШИХ КВАДРАТІВ

2.14.1. Найпоширеніші критерії порівняння методів оцінювання

Розглянемо властивості методу найменших квадратів та проаналізуємо, які оцінки невідомих параметрів b_0 та b_1 ми отримуємо, переходячи від окремої вибірки до множини вибірок.

Повернемося до нашого прикладу. Ми розглядали, як залежать обсяги реалізованої продукції від витрат на рекламу. Цю залежність ми встановили, маючи у розпорядженні лише кілька вибірових даних. Чи можна поширити наші висновки, наприклад, на ринок усєї країни? Тобто чи будуть наші висновки сталими? Як на них впливає збільшення або зменшення спостережень? Чи отримаємо ми ті самі результати, якщо зробимо таку саму вибірку, але в інші проміжки часу? На всі ці запитання можна дати позитивну відповідь у разі, якщо метод оцінювання невідомих параметрів буде найкращим з можливих при певних припущеннях. Отже, нам варто детально розглянути властивості методу найменших квадратів. Зразу ж зазначимо, метод найменших квадратів не є єдино можливим для оцінювання невідомих параметрів простої лінійної регресії, але він має ряд дуже важливих властивостей, що роблять його класичним методом оцінювання. В чому ж вони полягають? І взагалі, якими мають бути критерії, які свідчать про те, що один з методів оцінювання кращий за інші?

Найпоширенішими критеріями для аналізу методів оцінювання в економетриці є такі:

1) відсутність відхилення; 2) найменша дисперсія; 3) ефективність; 4) найкраща лінійна оцінка без відхилення (*BLUE*); 5) найменша середня квадратична помилка (*MSE*); 6) достатність.

Пояснимо ці критерії детальніше.

Метод оцінювання без відхилень

Розумітимемо під відхиленням для методу оцінювання різницю між очікуваним та дійсним значенням параметра.

$$\text{Відхилення} = E(b) - \beta.$$

Метод оцінювання вважається таким, що не має відхилення, коли відхилення дорівнює нулеві, тобто коли $E(b) = \beta$.

Оцінювання без відхилення є важливою властивістю, але не за своїм змістом, а в результаті комбінації з малою дисперсією.

Оцінювання з найменшою дисперсією

Метод оцінювання дає найкращі результати в тому разі, коли він забезпечує найменшу дисперсію порівняно з іншими методами.

Символічно b є найкращою, коли

$$E [b - E(b)]^2 < E[\tilde{b} - E(\tilde{b})]^2 \text{ або } \text{var}(b) < \text{var}(\tilde{b}),$$

де \tilde{b} є іншою, не обов'язково без відхилення, оцінкою параметра β .

Ефективна оцінка

Оцінка є ефективною, коли вона має обидві властивості, тобто не має відхилення та має найменшу дисперсію порівняно з іншими оцінками.

Символічно b є ефективною оцінкою, коли:

а) $E(b) = \beta$;

б) $E(b - E(b))^2 < E[b^* - E(b^*)]^2$,

де b^* — інша оцінка без відхилення, правильного значення β .

Іншими словами, ефективна оцінка є найкращою серед оцінок без відхилень.

Найкраща лінійна оцінка без відхилень (*BLUE*) (*best linear unbiased estimator*)

Оцінка є *BLUE*, коли вона без відхилення має найменшу дисперсію та є лінійною функцією від значень, які спостерігаються.

Наприклад, при значеннях y_1, y_2, \dots, y_n , лінійна оцінка має вигляд: $k_1 y_1 + k_2 y_2 + \dots + k_n y_n$, де $k_i, i = 1, n$ — константи.

Оцінка з мінімальною середньою квадратичною помилкою, *MSE*-оцінка

MSE — критерій є комбінацією властивостей оцінок без відхилень та мінімальною дисперсією.

Оцінка є мінімальною *MSE*-оцінкою, якщо вона має найменше значення: $MSE(b) = E[b - \beta]^2$.

Можна показати, що *MSE* дорівнює дисперсії оцінки плюс квадрат відхилення, тобто: $MSE(b) = \text{var}(b) + \text{відхилення}^2(b)$.

Покажемо це:

$$\begin{aligned} MSE &= E(b - \beta)^2 = E(b - E(b) + E(b) - \beta)^2 = \\ &= E[b - E(b)]^2 + [E(b) - \beta]^2 + 2E\{[b - E(b)][E(b) - \beta]\}, \end{aligned}$$

але

$$E(b - E(b))^2 = \text{var}(b),$$

$$[E(b) - \beta]^2 = \text{відхилення}^2(b).$$

$$E\{[b - E(b)][E(b) - \beta]\} = 0, \text{ тому що}$$

$$E\{[b - E(b)][E(b) - \beta]\} = E\{bE(b) - [E(b)]^2 - b\beta + \beta E(b)\} =$$

$$= [E(b)]^2 - [E(b)]^2 - \beta E(b) + \beta E(b) = 0.$$

$$\text{Тоді } MSE = \text{var}(b) + \text{відхилення}^2(b).$$

Достатність оцінки

Оцінка є достатньою, коли вона використовує всю вибірку інформацію. Достатність сама по собі не є важливою ознакою, але вона є необхідною умовою ефективною оцінки.

2.14.2. Властивості оцінок, отриманих за методом найменших квадратів

Теорема Гауса-Маркова показує, що оцінки, отримані за методом найменших квадратів, є BLUE-оцінками, коли виконуються основні припущення щодо випадкової величини ε . Тобто оцінки, розраховані за методом найменших квадратів, мають такі властивості: вони лінійні, без відхилень, мають найменшу дисперсію з усіх можливих методів оцінювання.

Таким чином, метод найменших квадратів є найкращим методом для оцінювання невідомих параметрів простої лінійної регресії.

ПІДСУМУЄМО ВИВЧЕНЕ

Після вивчення цього розділу ви зможете

1. Використовувати регресійний аналіз для знаходження взаємозв'язку між двома змінними.
2. Використовувати модель простої лінійної регресії і знати припущення для моделі простої лінійної регресії.
3. Оцінювати параметри вибіркової моделі.
4. Перевіряти регресійну модель на адекватність.
5. Застосовувати просту лінійну регресію до широкого кола проблем у бізнесі й економіці.
6. Тестувати значимість параметрів регресії b_1 і b_0 .
7. Розраховувати інтервали довіри для параметрів.
8. Прогнозувати за моделлю простої регресії.
9. Встановлювати інтервали довіри для прогнозного та середнього значення залежної змінної.

Короткий огляд розділу

1. Просту лінійну регресію використовують для встановлення лінійного зв'язку між двома змінними.
2. При простій лінійній регресії розглядають усього дві змінні. Незалежну змінну, наприклад x , яка пояснює залежну змінну, наприклад y .
3. При регресійному аналізі до уваги беруться всього два параметри: b_0 (перетин) і b_1 (нахил). Перетин показує значення y , коли x дорівнює нулю. Нахил вимірює вплив x на y .
4. Діаграму розподілу використовують для ілюстрації зв'язку між змінними x та y . Діаграма розподілу є просто нанесенням на графік значень змінних — x і y .
5. Одним із методів, які використовують для обчислення параметрів прямої, є метод найменших квадратів.
При цьому пряму проводять таким чином, щоб мінімізувати загальну суму квадратів відхилень фактичних значень від значень, які лежать на прямій (так званих теоретичних значень).
6. Основні припущення моделі лінійної регресії:
 - а) значеннями незалежної змінної x є або фіксовані числа, або, якщо вони випадкові змінні, то вони є статистично незалежними від випадкової величини ε ;
 - б) випадкова величина ε_i має нульове математичне сподівання і постійну дисперсію;
 - в) випадкові величини статистично незалежні одна від одної;
 - г) випадкові величини розподіляються за нормальним законом розподілу ($\varepsilon \sim N(0, \sigma^2)$).

7. Оцінюючи лінію регресії, прагнуть дізнатись, наскільки добре ця лінія представляє зв'язок між x та y . Щоб виміряти придатність лінії регресії, використовують стандартну помилку залишків або коефіцієнт детермінації. Стандартна помилка залишків показує відхилення змінної від лінії регресії. Коефіцієнт детермінації показує, наскільки варіація x пояснює варіацію y .

8. Під час оцінки параметрів регресії, нас цікавить їх значимість. Загалом, ми тестуємо гіпотези, щоб визначити, чи статистично значимо вони відрізняються від нуля. У деяких випадках оцінку проводять, щоб встановити, чи відрізняється нахил від певного заданого значення.

9. Тестування значимості нахилу (b_1) можна провести, використовуючи t -тест Ст'юдента.

10. Використовуючи регресійний аналіз для прогнозування величини y , іноді виникає потреба встановити інтервали довіри.

Зверніть увагу, що інтервали довіри стають ширшими в міру віддалення x від свого середнього значення, оскільки, чим далі від центра, тим менше впевненості у значенні спрогнозованої y .

Основні формули

Узагальнена регресійна модель:

$$y = \beta_0 + \beta_1 x + \varepsilon.$$

Вибіркова лінійна регресійна модель:

$$y = b_0 + b_1 x + e.$$

Оцінка параметрів, обчислених за методом найменших квадратів:

$$b_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})}, b_0 = \bar{y} - b_1 \bar{x}.$$

Сума квадратів, що пояснює регресію:

$$SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2.$$

Сума квадратів помилок:

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2.$$

Загальна сума квадратів:

$$SST = \sum_{i=1}^n (y_i - \bar{y})^2.$$

$$SST = SSR + SSE.$$

Коефіцієнт детермінації:

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}.$$

Коефіцієнт кореляції:

$$r = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y}.$$

t-тест Ст'юдента для перевірки, чи $b_i = \beta_i$:

$$t_{n-2} = \frac{b_i - \beta_i^*}{\hat{\sigma}_{b_i}}, i = 0, 1.$$

$$\hat{\sigma}_{b_0} = \sqrt{\hat{\sigma}_\varepsilon^2 \frac{\sum x^2}{n \sum (x - \bar{x})^2}}, \quad \hat{\sigma}_{b_1} = \frac{\hat{\sigma}_\varepsilon}{\sqrt{\sum (x - \bar{x})^2}},$$

$$\hat{\sigma}_\varepsilon = \sqrt{\frac{\sum e^2}{n-2}}.$$

t-тест Ст'юдента для перевірки, чи значимо b_i відрізняється від нуля:

$$H_0: \beta_i = 0, \quad H_1: \beta_i \neq 0.$$

$$t_{n-2} = \frac{b_i - 0}{\hat{\sigma}_{b_i}}.$$

F-тест для перевірки регресійної моделі на адекватність:

$$F_{1, n-2} = \frac{SSR / 1}{SSE / n - 2} = \frac{MSR}{MSE}.$$

t-тест для перевірки значимості коефіцієнта кореляції:

$$t_{n-2} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

Інтервал довіри для математичного сподівання значення y :

$$\hat{y}_{n+1} \pm t_{\alpha/2, n-2} \cdot \hat{\sigma}_\varepsilon \sqrt{\frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

Інтервал довіри для окремого значення y :

$$\hat{y}_{n+1} \pm t_{\alpha/2, n-2} \cdot \hat{\sigma}_\varepsilon \sqrt{1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

ВПРАВИ І КОМЕНТАРІ

Вправа 2.1. Пряма лінійна регресія

Припустимо, що ви збираєте дані про споживання домашніх господарств в Україні та їх дохід і оцінюєте таке рівняння:

$$C = 120 + 0.75y,$$

де C — споживання; y — дохід.

а) яка змінна є залежною, і яка змінна є незалежною?

б) поясніть взаємозв'язок між споживанням і доходом. На скільки зросте споживання, якщо дохід зросте на 1 одиницю?

Розв'язок

1. Оскільки споживання залежить від доходу, то споживання є залежною змінною, а дохід — незалежною змінною.

2. Дане рівняння свідчить про те, що споживання складається з двох частин. Перша частина представлена перетином 120, що означає, розмір споживання 120 при відсутності доходу. Друга частина складається з $0.75y$, тобто зростання доходу на одну одиницю зумовить зростання споживання на 0.75.

Вправа 2.2. Оцінка невідомих параметрів регресії.

1. Припустимо, що ви збираєте дані про річний продаж фірмою “Україна” продукції (y) і суми, які використано на наукові дослідження (x). Ви маєте таку статистику:

$$\text{cov}(x, y) = 300;$$

$$\text{var}(y) = 125;$$

$$\text{var}(x) = 880.$$

Середній річний продаж (\bar{y}) = 1200.

Середня сума на наукові дослідження (\bar{x}) = 895.

2. Підрахуйте коефіцієнт кореляції між продажем і сумою, використаною на наукові дослідження.

3. Визначте коефіцієнт детермінації для регресії.

4. Знайдіть параметри регресії b_0 і b_1 .

Розв'язок

$$\text{а) } r_{x,y} = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} = \frac{300}{\sqrt{125} \sqrt{880}} = 0.286;$$

$$\text{б) } R^2 = r^2 = 0.286^2 = 0.081;$$

$$\text{в) } b_1 = \frac{\text{cov}(x, y)}{\text{var}(x)} = \frac{300}{880} = 0.34;$$

$$b_0 = \bar{y} - b_1 \bar{x} = 1200 - 0.34 \cdot 895 = 895.7.$$

Вправа 2.3. Коефіцієнт детермінації

Проведіть оцінку регресії і підрахуйте SSE, SSR.

$$SSE = 53.27;$$

$$SSR = 202.91.$$

Розрахуйте SST, R^2 і r , використовуючи дану інформацію.

Розв'язок:

$$SST = SSE + SSR = 53.27 + 202.91 = 256.18;$$

$$R^2 = SSR / SST = 202.91 / 256.18 = 0.792;$$

$$r = \sqrt{0.792} = 0.889.$$

Вправа 2.4. Оцінка невідомих параметрів регресії

Є такі дані (табл. 2.9):

Таблиця 2.9

x	y	$(x_i - \bar{x})$	$(y_i - \bar{y})$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$
20	75				
14	85				
12	92				
20	88				
33	72				
38	99				

Заповніть пропуски і знайдіть параметри регресії b_0 і b_1 .

Розв'язок (табл. 2.10)

Таблиця 2.10

x	y	$(x_i - \bar{x})$	$(y_i - \bar{y})$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$
20	75	-2.83	-10.17	28.81	8.03
14	85	-8.83	-0.17	1.47	78.03
12	92	-10.83	6.83	-74.03	117.36
20	88	-2.83	2.83	-8.03	8.03
33	72	-10.17	-13.17	133.86	103.36
38	99	-15.17	13.83	209.81	230.03

Σ 137 511 24.17 544.84

$$\bar{x} = 22.83; \bar{y} = 85.16$$

$$b_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{224.17}{544.83} = 0.04436;$$

$$b_0 = \bar{y} - b_1 \bar{x} = 85.16 - 0.04436(22.83) = 84.15.$$

Вправа 2.5. *t*-тест для оцінки значимості параметрів регресії

Оцінюючи взаємозв'язок між ціною на товар та обсягами продажу (крива попиту) економісти іноді трансформують дані за цінами та обсягами продажу, беручи натуральні логарифми обох показників. У такому разі коефіцієнт нахилу b_1 можна інтерпретувати як еластичність попиту за ціною (чутливість попиту до змін у ціні).

Інформацію за цінами і обсяги продажу наведено у табл. 2.11.

Таблиця 2.11

Ціна, P	Кількість, Q
2.20	1020
1.85	1521
1.50	1755
1.21	2130
0.99	3200
0.79	4105

1. Оцініть еластичність попиту за цими даними.
2. Використайте *t*-тест для перевірки значимості b_1 .
3. Побудуйте 90%-ий інтервал довіри еластичності попиту за ціною.

Розв'язок (табл. 2.12)

Таблиця 2.12

$x=\ln(P)$	$y=\ln(Q)$	$(x - \bar{x})^2$	$(x - \bar{x})$	$(x - \bar{x})^2(y - \bar{y})^2$	\hat{y}	e	e^2
0.788	6.928	0.246	0.492	-0.348	6.986	-0.058	0.003
0.615	7.327	0.104	0.092	-0.098	7.211	0.116	0.014
0.406	7.470	0.113	0.026	-0.018	7.483	-0.013	0.000
0.191	7.664	0.010	0.001	-0.003	7.762	-0.098	0.010
-0.010	8.071	0.019	0.194	-0.133	8.022	0.048	0.002
-0.236	8.320	0.2779	0.476	-0.364	8.315	0.005	0.000

Середнє 0.292 7.640 0.124
 Σ 0.744 45.780 0.029

$$b_1 = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2} = \frac{-0.965}{0.744} = -1.298;$$

$$b_0 = \bar{y} - b_1 \bar{x} = 7.630 - (-1.298)(0.292) = 8.009;$$

$$\hat{\sigma}_e^2 = \frac{SSE}{(n-2)} = \frac{\sum e^2}{(n-2)} = \frac{0.0291}{(6-2)} = 0.0073;$$

$$\hat{\sigma}_e = \sqrt{0.0073} = 0.853;$$

$$\hat{\sigma}_{b_1}^2 = \frac{\hat{\sigma}_e^2}{\sum (x - \bar{x})^2} = \frac{0.0073}{0.744} = 0.0098;$$

$$\hat{\sigma}_{b_1} = \sqrt{0.0098} = 0.0989;$$

$$b_1 \pm t_{\alpha/2, n-2} \hat{\sigma}_{b_1};$$

$$-1.298 \pm 2.776(0.0989);$$

$$-1.572 \leq b_1 \leq -1.023.$$

Вправа 2.6. Інтервал довіри для прогнозу

Вивчаючи зміну попиту на йогурт залежно від його ціни, отримано такі результати:

$$y_i = 5149.268 - 2009.90 x_i;$$

$$\hat{\sigma}_e = 468.943, R^2 = 0.867;$$

$$\sum [x_i - \bar{x}]^2 = 1.426 ;$$

$$\bar{x} = 1.423 , n = 4 .$$

Фірма встановлює на йогурт ціну: 1.75 грн. Спрогнозуйте попит і побудуйте 90% -ний інтервал довіри для математичного сподівання прогнозу.

Розв'язок

$$\hat{y}_{n+1} = 5149.268 - 2009.90(1.75) = 1631.9 ;$$

$$\hat{y}_{n+1} \pm t_{\alpha/2} \hat{\sigma}_\varepsilon \sqrt{\frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} ;$$

$$1631.9 \pm 2.132(468.943) \sqrt{\frac{1}{4} + \frac{(1.75 - 1.423)^2}{1.426}} .$$

Отже, $1061.05 < E(y_{n+1}) < 2201.85$.

Вправа 2.7. Інтервал довіри для параметрів

1. Ви оцінюєте регресію y за x і маєте такі результати:

$$b_1 = 2.8 ,$$

$\sigma_{b_1} = 0.6$ — стандартна помилка параметра регресії (припускаємо, що вона відома).

2. Побудуйте 95% -ний інтервал довіри для b_1 .

3. Припустимо, що σ_{b_1} невідома. Чи зміниться спосіб побудови інтервалу довіри для b_1 ? Припустимо, що для оцінки регресії було використано 25 спостережень.

Розв'язок

1.

$$b \pm z_{\alpha/2} \sigma_{b_1} ;$$

$$2.8 \pm 1.96(0.6) .$$

Отже, інтервал становить: від 1.624 до 3.976.

2. Якщо σ_{b_1} невідоме, то використовуються оцінки для $\hat{\sigma}_{b_1}$. У цьому разі слід використовувати t -розподіл Ст'юдента, щоб побудувати інтервал довіри. Нехай $\hat{\sigma}_{b_1} = 0.6$.

$$b \pm t_{\alpha/2, n-2} \hat{\sigma}_{b_1} ;$$

$$2.8 \pm 2.069(0.6) .$$

Отже, інтервал становить від 1.559 до 4.041.

Вправа 2.8. Тест перевірки нахилу на значимість

1. Ви оцінюєте таку регресію:

$$y = 2300 + 10.12x ;$$

$$SSE = 28225 ;$$

$$n = 28 .$$

$$\sum (x - \bar{x})^2 = 3300 .$$

2. Перевірте значимість нахилу при 95% -ному рівні довіри.

3. Побудуйте 90% -ний інтервал довіри для нахилу.

Розв'язок

1.

$$\hat{\sigma}_\varepsilon^2 = SSE / (n - 2) = 28225 / (28 - 2) = 1085.58 ;$$

$$\hat{\sigma}_\varepsilon = \sqrt{1085.58} = 32.95 ;$$

$$\hat{\sigma}_{b_1} = \frac{\hat{\sigma}_\varepsilon}{\sqrt{\sum (x - \bar{x})^2}} = \frac{32.95}{\sqrt{3300}} = 0.5736 ;$$

$$t^* = \frac{b_1}{\hat{\sigma}_{b_1}} = \frac{10.12}{0.5736} = 17.64 .$$

Оскільки критичне значення величини $t = 2.056$, ми відкидаємо нульову гіпотезу про те, що $b_1 = 0$, і робимо висновок, що b_1 значимо відрізняється від нуля.

2.

$$b_1 \pm t_{\alpha/2, n-2} \hat{\sigma}_{b_1} ,$$

$$10.12 \pm 1.706(0.5736) .$$

Отже, інтервал становить від $9.14 \leq b_1 \leq 11.10$.

ВИКОНАЙТЕ САМОСТІЙНО

Вправи. Проста лінійна регресія

Вправа 2.9

Маємо статистику витрат родини залежно від кількості її членів (табл. 2.13):

Таблиця 2.13

Кількість членів родини, x	Витрати на відпустку, y
1	17
2	11
2	23
4	19
6	30

1. Побудувати лінійну регресію і розрахувати параметри.
 2. Розрахувати коефіцієнт кореляції і детермінації.
- Які висновки можна зробити щодо якості регресійної моделі?
3. Побудувати ANOVA-таблицю.

4. Розрахувати F -критерій Фішера, знайти за таблицею F -розподілу критичне значення. Який висновок щодо адекватності регресійної моделі можна зробити?

Вправа 2.10

Є 4 динамічні ряди. Якщо ці дані апроксимувати простою лінійною регресією, то вони дадуть однакові результати. Побудувати графіки і проаналізувати їх зображення. Побудувати лінійну модель, оцінивши параметри за методом найменших квадратів. Пояснити результати.

Таблиця 2.14

x_1	y_1	x_2	y_2	x_3	y_3	x_4	y_4
10	8.04	10	9.14	10	7.46	8	6.58
14	9.96	14	8.1	14	8.86	8	5.76
5	5.68	5	4.74	5	5.73	8	7.71
8	6.95	8	8.14	8	6.77	8	8.84
9	8.81	9	8.77	9	7.11	8	8.47
12	10.84	12	9.13	12	8.15	8	7.04
4	4.26	4	3.1	4	5.39	8	5.25
7	4.82	7	4.26	7	6.42	19	12.5
11	8.33	11	9.26	11	7.81	8	5.56
13	7.58	13	8.74	13	12.74	8	7.91
6	7.24	6	6.13	6	6.08	8	6.89

Вправа 2.11

У табл. 2.15 подано дані про рівень звільнень на 100 робітників та рівень безробіття у виробничій сфері східного регіону України (дані умовні) протягом 1980 — 1992 рр.

Примітка. Дані про звільнення стосуються людей, що покинули роботу за власним бажанням.

Таблиця 2.15

Рівень звільнень та безробіття у виробничій сфері східного регіону України, 1980 — 1992 рр.

Рік	Рівень звільнень на 100 робітників, y	Рівень безробіття (%), x
1980	1.3	6.2
1981	1.2	7.8
1982	1.4	5.8
1983	1.4	5.7
1984	1.5	5.0
1985	1.9	4.0
1986	2.6	3.2
1987	2.3	3.6
1988	2.5	3.3
1989	2.7	3.3
1990	2.1	5.6
1991	1.8	6.8
1992	2.2	5.6

1. Нанесіть дані на координатну площину;
2. Припустимо, що рівень звільнень y лінійно пов'язаний з рівнем безробіття x і цей зв'язок виражається моделлю $y_i = b_0 + b_1 x_i + e_i$. Обчисліть параметри b_0, b_1 та їх середні квадратичні відхилення;
3. Обчисліть коефіцієнт детермінації та кореляції;
4. Поясніть ваші результати;
5. Нанесіть на площину залишки e_i . Які висновки ви можете зробити?
6. За щорічними даними за період з 1986 по 1996 рік і такою самою моделлю, як у пункті 2, отримано такі результати:

$$\hat{y}_i = 3.1237 - 0.1714x_i;$$

$$\hat{\sigma}_{b_1} = 0.0210 \text{ та } R^2 = 0.8575.$$

Якщо ці результати відрізняються від отриманих у (2), то як ви поясните цю різницю?

Вправа 2.12

У таблиці 2.16 подано дані про кількість телефонів на 1000 осіб (y) та валовий внутрішній продукт (ВВП) на душу населення (x) у західному регіоні України (дані умовні) за 1976—1997 рр. Чи є зв'язок між цими двома змінними?

Таблиця 2.16

Кількість телефонів та ВВП на душу населення в західному регіоні України, 1976—1997 рр.

Рік	y	x	Рік	y	x
1976	36	1299	1988	90	2723
1978	37	1365	1989	102	3033
1979	38	1409	1990	114	3317
1980	41	1549	1991	126	3487
1981	42	1416	1992	141	3575
1982	45	1473	1993	163	3784
1983	48	1589	1994	196	4025
1984	54	1757	1995	223	4286
1985	59	1974	1996	269	4628
1986	67	2204	1997	291	5038
1987	78	2462			

Вправа 2.13

У таблиці 2.17 подано дані про ціни на золото, індекс споживчих цін (CPI) та індекс Нью-Йоркської фондової біржі (NYSE) у Сполучених Штатах за 1977 — 1991 рр. Індекс NYSE включає більшість акцій, зареєстрованих на NYSE, приблизно понад 1500.

Таблиця 2.17

Рік	Ціна на золото в Нью-Йорку, дол. за тройську унцію	Індекс споживчих цін (CPI), 1982-84 = 100	Індекс Нью-Йоркської фондової біржі (NYSE), 31.12.1965 = 100
1977	147.98	60.6	53.69
1978	193.44	65.2	53.70
1979	307.62	72.6	58.32
1980	612.51	82.4	68.10
1981	459.61	90.9	74.02
1982	376.01	96.5	68.93
1983	423.83	99.6	92.63
1984	360.29	103.9	92.46
1985	317.30	107.6	108.90
1986	367.87	109.6	136.00
1987	446.50	113.6	161.70
1988	436.93	118.3	149.91
1989	381.28	124.0	180.02
1990	384.08	130.7	183.46
1991	362.04	136.2	206.33

Джерело: дані про CPI і NYSE взято з "Economic Report of the President" (1993. — January. — Т. В-59, В-91). Дані про ціни на золото взято з "U.S. Department of Commerce, Bureau of Economic Analysis, Business Statistics, 1963—1991". — Р. 68.

1. Нанесіть на одну і ту саму діаграму ціни на золото, індекс споживчих цін та NYSE Index.

2. Інвестування вважається страховкою від інфляції, якщо ставка процента вища за рівень інфляції або перебуває на тому самому рівні. Щоб перевірити цю гіпотезу, використайте, якщо вони підходять, такі моделі:

$$\text{Gold price}_t = b_0 + b_1 \text{CPI}_t + e_t;$$

$$\text{NYSE Index}_t = b_0 + b_1 \text{CPI}_t + e_t.$$

Якщо гіпотеза правильна, яке значення параметрів ви б очікували?

3. Який вид страхування від інфляції кращий: золото чи фондова біржа?

ТЕСТИ

Тест 1

Вибрати правильну відповідь на запитання

1. Діаграма розподілу — це:

- лінія, що відображає зв'язок між незалежною і залежною змінними;
- інша назва простої регресії;
- інша назва множинної регресії;
- графік значень незалежної і залежної змінних;
- лінія з нахилом a і перетином b .

2. Лінійна регресія:

- лінія, що відображає зв'язок між незалежною і залежною змінними;
- інша назва простої регресії;
- лінія, яка завжди має нахил, що дорівнює 1;
- графік значень незалежної і залежної змінних;
- лінія, яка завжди має нахил, що дорівнює 0.

3. Нахил:

- точка, де лінія регресії перетинає вісь y ;
- вимірює придатність лінії регресії;
- вимірює зв'язок між залежною і незалежною змінними;
- завжди дорівнює 1;
- інша назва коефіцієнта детермінації.

4. Перетин:

- точка, де лінія регресії перетинає вісь y ;
- вимірює придатність лінії регресії;
- вимірює зв'язок між залежною і незалежною змінними;

17. З урахуванням співвідношення між заробітною платою (в гривнях) — y і освітою (в роках) — x , $y = 12.201 + 525x$, особа, що навчалася додатково нуль років, може очікувати на таку додаткову оплату:

- а) 12.201; з) 1.050;
б) 525; д) 12.201+525.
в) 24.402;

18. Якщо регресія має $R^2 = 0.80$, то регресійна лінія:

- а) пояснює 80% варіації змінної x ;
б) пояснює 80% варіації змінної y ;
в) матиме нахил 0.80;
г) матиме перетин 0.80;
д) не пояснює зв'язку між x і y .

Тест 2

1. При перевірці значимості коефіцієнта регресії (b_1) нульова гіпотеза представлена:

- а) $H_0: \beta_1 = 1$; б) $H_0: \beta_1 = -1$; в) $H_0: \beta_1 = 0$; г) $H_0: \beta_1 = 0.5$; д) $H_0: \beta_1 = -0.5$.

2. Для двостороннього тесту значимості b_1 альтернативною гіпотезою є:

- а) $H_1: \beta_1 \neq 1$; б) $H_1: \beta_1 > 1$; в) $H_1: \beta_1 \neq 0$; г) $H_1: \beta_1 > 0$; д) $H_1: \beta_1 < 0$.

3. Щоб перевірити, чи є нахил позитивним і таким, що значимо відрізняється від нуля, нульова гіпотеза має бути такою:

- а) $H_0: \beta_1 \leq 0$; б) $H_0: \beta_1 \geq 0$; в) $H_0: \beta_1 = 0$; г) $H_0: \beta_1 = 1$; д) $H_0: \beta_1 \geq 1$.

4. Щоб встановити, чи є нахил негативним і таким, що значимо відрізняється від нуля, нульова гіпотеза має бути такою:

- а) $H_0: \beta_1 \leq 0$; б) $H_0: \beta_1 \geq 0$; в) $H_0: \beta_1 = 0$; г) $H_0: \beta_1 = 1$; д) $H_0: \beta_1 = -1$.

5. При перевірці значимості параметра регресії використовуємо:

а) F -тест; б) χ^2 -тест; в) t -тест; г) біноміальний розподіл; д) будь-що із згаданого тут.

6. При перевірці значимості регресії ми:

- а) перевіряємо значимість перетину b_0 ;
б) перевіряємо значимість нахилу b_1 ;
в) однаково перевіряємо b_0 і b_1 ;
г) більше цікавимося дисперсією y ;
д) більше цікавимося дисперсією x .

7. Якщо нахил значимо відрізняється від нуля, то це означає, що:

- а) незалежна змінна відіграє важливу роль у поясненні залежної змінної;
б) незалежна змінна не відіграє важливої ролі у поясненні залежної змінної;
в) залежна змінна відіграє важливу роль у поясненні незалежної змінної;

г) залежна змінна не відіграє важливої ролі у поясненні незалежної змінної;
д) усі відповіді неправильні.

8. Якщо нахил регресії становить 2.4 і дисперсія нахилу 0.8, то величина t , що її використовують для перевірки $H_0: \beta_1 = 0$, становитиме:

- а) $0.8 / 2.4$; б) $2.4 / \sqrt{0.8}$; в) $(2.4 - 1) / \sqrt{0.8}$; г) $2.4 / 0.8$; д) $(2.4 - 1) / \sum 0.8$.

9. Якщо нахил для регресії становить 2.4 і дисперсія нахилу 0.8, то величина t , яку використовують для перевірки $H_0: \beta_1 = 1$, становитиме:

- а) $0.8 / 2.4$; б) $2.4 / \sqrt{0.8}$; в) $(2.4 - 1) / \sqrt{0.8}$; г) $2.4 / 0.8$; д) $(2.4 - 1) / \sum 0.8$.

10. За інших однакових умов, чим більша оцінка середнього квадратичного відхилення нахилу, тим:

- а) більша t -величина нахилу;
б) менша t -величина нахилу;
в) більша величина перетину;
г) менша t -величина перетину;
д) більший коефіцієнт нахилу.

11. Яке з поданих тверджень є правильним:

- а) $SSE + SSE > SST$; б) $R^2 = -0.5$; в) $R^2 = 1.83$; г) $\hat{\sigma}_e = -0.35$; д) $t = -2.3$?

12. У регресії завжди має бути:

- а) $r > 0$; б) $r < 0$; в) $t > 0$; г) $t < 0$; д) $\hat{\sigma}_e \geq 0$.

13. У регресії завжди має бути:

- а) $b_1 \geq 0$; б) $b_1 \leq 0$; в) $\hat{\sigma}_{b_1} \geq 0$; г) $\hat{\sigma}_{b_1} \leq 0$; д) $r \leq 0$.

14. У простій регресії $\hat{\sigma}_{b_0}^2$ дорівнює:

- а) $\hat{\sigma}_e = -0.35$ в) $\hat{\sigma}_e^2 / \left(\sum (y_i - \bar{y})^2 \right)$

$\hat{\sigma}_e^2 / \left(\sum (x_i - \bar{x})^2 \right)$; г) $\hat{\sigma}_e^2 (b_0 / b_1)$;

- б) $\hat{\sigma}_e^2$; д) $\hat{\sigma}_e^2 \left[\left(\sum x_i^2 \right) / \left(n \sum (x_i - \bar{x})^2 \right) \right]$.

15. У простій регресії $\hat{\sigma}_{b_1}^2$ дорівнює:

- а) $\hat{\sigma}_e^2 / \left(\sum (x_i - \bar{x})^2 \right)$; г) $\hat{\sigma}_e^2 (b_0 / b_1)$;

- б) $\hat{\sigma}_e^2$; д) $\hat{\sigma}_e^2 \left[\left(\sum x_i^2 \right) / \left(n \sum (x_i - \bar{x})^2 \right) \right]$.

- в) $\hat{\sigma}_e^2 / \left(\sum (y_i - \bar{y})^2 \right)$;

16. Для регресії з n спостережень інтервал довіри $1 - \alpha\%$ для перетину буде:

$$а) b_0 \pm t_{(\alpha, n-2)} \cdot \hat{\sigma}_{b_0}^2; \quad з) b_0 \pm t_{(\alpha/2, n-2)} \cdot \hat{\sigma}_{b_1}^2;$$

$$б) b_0 \pm t_{(\alpha, n-2)} \cdot \hat{\sigma}_{b_1}^2; \quad д) b_0 \pm t_{(\alpha/2, n-2)} \cdot \hat{\sigma}_\varepsilon.$$

$$в) b_0 \pm t_{(\alpha/2, n-2)} \cdot \hat{\sigma}_{b_0};$$

17. Для регресії з n спостережень інтервал довіри $1 - \alpha\%$ для нахилу буде:

$$а) b_1 \pm t_{(\alpha, n-2)} \cdot \hat{\sigma}_{b_1}^2; \quad з) b_1 \pm t_{(\alpha/2, n-2)} \cdot \hat{\sigma}_{b_1}^2;$$

$$б) b_1 \pm t_{(\alpha, n-2)} \cdot \hat{\sigma}_{b_1}; \quad д) b_1 \pm t_{(\alpha/2, n-2)} \cdot \hat{\sigma}_\varepsilon.$$

$$в) b_0 \pm t_{(\alpha/2, n-2)} \cdot \hat{\sigma}_{b_0}$$

ВІДПОВІСТІ НА ЗАПИТАННЯ: ТАК/НІ, ЯКЩО НІ, ПОЯСНІТЬ ЧОМУ

Частина 1

1. Діаграма розподілу показує взаємозв'язок між залежною і незалежною змінними.

2. У регресійному аналізі залежна змінна завжди знаходиться на осі x , а незалежна змінна — на осі y .

3. Придатність регресійної моделі може вимірюватись як коефіцієнтом детермінації, так і стандартною помилкою оцінювання (корінь з середнього квадрату залишків).

4. Множинну регресію використовуємо, коли на незалежну змінну впливає дві або більше залежних змінних.

5. Нахил регресії вимірює зв'язок між залежною і незалежною змінними.

6. Перетин у регресії — це значення залежної змінної, коли незалежна змінна дорівнює 0.

7. Якщо ми хочемо, за допомогою регресійного аналізу виміряти зв'язок між оцінками у вузі і оцінками у школі, то залежною змінною повинні бути оцінки у вузі, а незалежною змінною — оцінки у школі.

8. У кореляційному аналізі важливо знати, яка саме змінна є незалежною, а яка залежною.

9. Метод найменших квадратів визначає лінію регресії через мінімізацію суми квадратів відхилень.

10. Коефіцієнт детермінації вимірює частку дисперсії незалежної змінної, що пояснює регресію.

11. Якщо дві змінні мають коефіцієнт кореляції -1 , то зв'язок між ними зворотний.

12. Якщо економіст поміняє місцями залежну і незалежну змінні (припускаємо, що це можливо в економічному контексті), то R^2 зменшиться.

13. R^2 у простій регресії дорівнює квадрату коефіцієнта кореляції між x і y .

14. Якщо $b_1 < 0$, то це означає, що x має дуже малий вплив на y .

15. Якщо коефіцієнт кореляції дорівнює 1, то точки даних знаходяться на прямій лінії.

18. Коваріація завжди лежить між -1 та $+1$.

Частина 2

1. При перевірці значимості нахилу в регресії нульова гіпотеза завжди: $\beta_1 = 1$.

2. Від'ємний і значимий нахил означає, що є непрямої зв'язок між незалежною і залежною змінними.

3. Якщо рівень значимості $\alpha = 0.05$, двосторонній тест значимості b_1 матиме менше абсолютне критичне значення, ніж односторонній тест.

4. Прогнози є одним із найважливіших застосувань регресійного аналізу.

5. Прогнозне значення y завжди дорівнюватиме дійсному значенню y .

6. Значення t -статистики не може бути від'ємним.

7. У простій лінійній регресії великий R^2 означає значимий нахил.

8. Проблема помилок у змінних виникає через неточності незалежної змінної.

9. Від'ємна t -статистика для нахилу означає, що незалежна змінна не є значимою.

10. t -статистику можна використати, щоб перевірити значимість коефіцієнта кореляції між двома випадковими змінними.

ЗАПИТАННЯ І ЗАДАЧІ

1. Дано таку інформацію про просту лінійну регресію:

$$SSE = 28.6; \quad SSR = 30.1; \quad n = 50.$$

а) Обчисліть коефіцієнт детермінації.

б) Обчисліть оцінку середньоквадратичного відхилення випадкової величини (стандартну помилку оцінювання).

2. Дано таку інформацію про дві змінні — x та y :

$$\bar{x} = 32, \bar{y} = 6; \quad \text{cov}(x, y) = 81; \quad \text{var}(x) = 21.$$

Розрахуйте нахил і перетин регресії y за x .

3. Встановіть взаємозв'язок між доходом (y) і кількістю років навчання оцінюючи таку регресію:

$$y = 12 + 125x.$$

(доход вимірюємо в гривнях, навчання — в роках). Поясніть результати регресії.

4. Підрахуйте коефіцієнт кореляції між x та y , на основі такої інформації (табл. 2.18):

Таблиця 2.18

	x	y
1	10	22
2	9	31
3	11	19
4	6	25

5. За даними пункту 4 побудуйте діаграму.

6. Використайте інформацію пункту 4 для оцінки регресії y за x .

7. Зробіть оцінку функції споживання:

$$y = 12.25 + 0.82x.$$

На скільки зросте споживання, якщо дохід (x) збільшиться на 1 гривню? Яким буде споживання, якщо дохід буде нульовим?

8. Припустимо, що R^2 з регресії y завданні 7 становить 0.75. Поясніть, що це означає.

9. Є така інформація:

$$\sum_{i=1}^n x_i^2 = 92, \quad \sum (x - \bar{x})^2 = 3.7, \quad \hat{\sigma}_e^2 = 1.25, \quad n = 40.$$

а) знайдіть σ_{b_0} ;

б) знайдіть $\hat{\sigma}_{b_1}$.

10. Припустимо, що ви оцінюєте залежність доходу відповідно до кількості років навчання, використовуючи 30 спостережень. Середньоквадратичні відхилення параметрів подано в дужках.

$$y = 11.929 + 421x \\ (4.825) \quad (127).$$

а) перевірте значимість нахилу при 5%-ному рівні значимості.

б) побудуйте 95%-ний інтервал довіри для нахилу.

11. Припустимо, що в регресії з питання 10 $SSE = 75$ і $SSR = 81$. Використайте F -тест для перевірки значимості регресії. Використайте 5%-ний рівень значимості.

12. Припустимо, що ви підраховали кореляцію між двома випадковими змінними, яка дорівнює 0.62. Якщо для оцінки коефіцієнта кореляції було використано 25 спостережень, використайте 5%-ний рівень значимості, щоб перевірити значимість коефіцієнта кореляції.

ВІДПОВІДІ

Тести

Тест 1

- 1) г; 5) г; 9) а; 13) д; 17) а;
 2) а; 6) б; 10) б; 14) г; 18) б.
 3) в; 7) б; 11) в; 15) в;
 4) а; 8) г; 12) в; 16) б;

Тест 2

- 1) в; 5) в; 9) в; 13) в; 17) б.
 2) в; 6) б; 10) б; 14) д;
 3) а; 7) а; 11) д; 15) а;
 4) б; 8) б; 12) д; 16) в;

Так/Ні

Частина 1

- Ні. Діаграма розподілу — це нанесені на графік незалежна і залежна змінні.
- Ні. x — незалежна змінна та y — залежна змінна.
- Так.
- Ні. Множинну регресію використовуємо, коли є дві чи більше незалежних змінних.
- Так.
- Так.
- Так.
- Ні. В кореляційному аналізі немає значення, яка змінна є залежною, а яка незалежною.
- Так.
- Ні. Він вимірює долю дисперсії залежної змінної, яка пояснює регресію.
- Так.
- Ні. R^2 залишається такою самою.
- Так.
- Ні. Воно може мати значний від'ємний вплив.
- Так.
- Ні. Кореляція знаходитиметься між -1 і $+1$; коваріація ж може набути будь-якого значення.

Частина 2

- Ні. $\beta_1 = 0$.
- Так.
- Ні. Абсолютне критичне значення буде більшим.
- Так.
- Ні. Прогнозне значення — це тільки можливе значення, і тому воно не

обов'язково дорівнює фактичній величині.

6. Ні. Коли параметр від'ємний, t -статистика теж буде від'ємною.

7. Так.

8. Ні. Може бути значущим і негативним.

Запитання і задачі

$$1. \text{ а) } R^2 = \frac{SSR}{SST} = \frac{30.1}{28.6 + 30.1};$$

$$6) \hat{\sigma}_\varepsilon = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{28.6}{50-2}} = 0.772.$$

$$2. b_1 = \frac{\text{cov}(x,y)}{\text{var}(x)} = \frac{81}{21} = 3.86, \quad b_0 = \bar{y} - b_1\bar{x} = 6 - 3.86(32) = -117.52.$$

3. Значення перетину показує, що людина без освіти отримуватиме 12 гривень. З коефіцієнта кута нахилу видно, що за кожен додатковий рік навчання доход підвищуватиметься на 125 гривень.

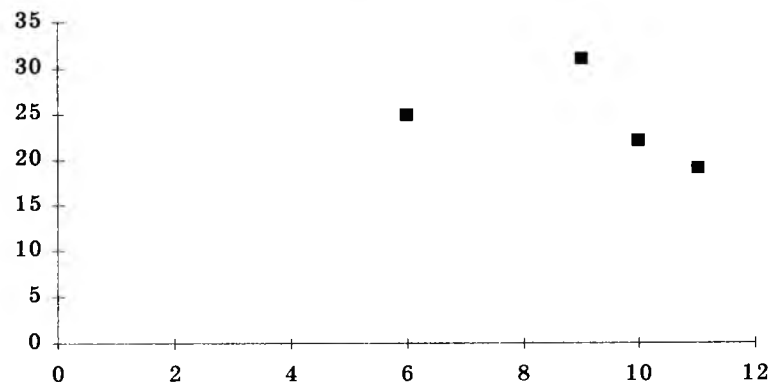
4.

Таблиця 2.19

	x	y	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
1	10	22	-2.25	1	5.06
2	9	31	0	0	45.56
3	11	19	-10.5	4	27.56
4	6	25	-2.25	9	0.56
Середнє Сума	9	24.2	-15	14	78.75

$$r_{x,y} = \frac{-15}{\sqrt{14} \sqrt{78.75}} = -0.452.$$

5.



Малюнок 2.18. Графічне зображення даних

$$6. b_1 = \frac{-15}{14} = -1.07;$$

$$b_0 = 24.2 - (-1.07)(9) = 33.83.$$

7. Зростання доходу на 1 гривню приведе до збільшення споживання на 0.82 гривні. Якщо доход становить 0, то споживання буде 12.25 гривні.

8. Те, що R^2 становить 0.75, означає, що 75% споживання залежить від доходу.

$$9. \text{ а) } \hat{\sigma}_{b_0} = \sqrt{\frac{\sigma_\varepsilon^2}{n \sum (x_i - \bar{x})^2}} = \sqrt{\frac{1.25}{(40)(3.7)}} = \sqrt{0.777} = 0.881;$$

$$6) \hat{\sigma}_{b_1}^2 = \frac{\hat{\sigma}_\varepsilon^2}{\sum (x_i - \bar{x})^2} = \frac{1.25}{3.7} = 0.338.$$

$$10. \text{ а) } t_{b_1} = 421 / 127 = 3.31;$$

$$t_{0.05/2, 30-2} = 2.048,$$

тобто ми відкидаємо: $H_0: \beta_1 = 0$.

$$6) 160.90 \leq \beta_1 \leq 681.10.$$

$$11. F_{1, n-2} = \frac{SSR/1}{SSE/(n-2)} = \frac{81/1}{75/(30-2)} = 30.24;$$

$$F_{0.05, 1, 28} = 4.20,$$

тобто регресія значима і ми відкидаємо: $H_0: \beta_1 = 0$.

$$12. t_{n-1} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.62\sqrt{25-2}}{\sqrt{1-0.62^2}} = 4.83;$$

$$t_{0.025; 24} = 2.064, \quad t_{n-1} > t_{0.025; 24},$$

тобто ми відкидаємо $H_0: \rho = 0$. Коефіцієнт кореляції значимий.

ЗАСТОСУВАННЯ ЕКОНОМІЧНИХ ЗНАТЬ

Модель оцінки капітальних активів (МОКА-модель)

Модель оцінки капітальних активів (у англійському варіанті CAPM — Capital Asset Pricing Model) дає змогу аналізувати надійність вкладів і оцінювати співвідношення ризику та доходів від різних інвестицій. Вона встановлює зв'язок між доходами окремих власників цінних паперів і загальними доходами ринку цінних паперів.

Мінливість ринку цінних паперів є загальним визначником оцінки ризику індивідуальних активів і цінних паперів. Якщо доходи від певних акцій зростають або спадають у процентному відношенні більше, ніж у цілому на ринку, то їх називають ризикованішими за сам ринок.

МОКА-модель у формі “премія за ризик” можна вважати прикладом регресії, що проходить через початок координат. Вона має такий вигляд:

$$(R_i - r_f) = \beta_i (R_m - r_f), \quad (2.126)$$

де R_i — очікувана норма прибутку з безпечної ставки i ;

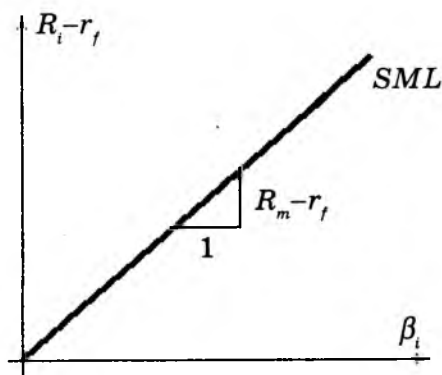
R_m — очікувана норма прибутку ринкового “портфеля”, представленого, наприклад, складним біржовим індексом S&P500;

r_f — позаризикова (вільна від ризику) норма прибутку, наприклад, прибуток з 90-денних державних облігацій;

β_i — коефіцієнт, що вимірює систематичний ризик, тобто ризик, який не можна усунути за допомогою диверсифікації. Цей коефіцієнт також вимірює рівень, до якого i -та безпечна норма прибутку рухається разом з ринком. Якщо $\beta_i > 1$, то вважається, що безпека нестійка і загрозна, тоді як при $\beta_i < 1$ безпека більш стійка або, як кажуть, захисна. (Зуваження. Не треба змішувати параметр β_i з коефіцієнтом нахилу b_1).

Якщо ринок капіталу працює ефективно, тоді МОКА-модель стверджує, що безпека i -ої премії за ризик $(R_i - r_f)$ дорівнює β_i -коефіцієнту цієї безпеки, помноженому на очікувану ринкову премію за ризик $(R_m - r_f)$, і ми маємо ситуацію, зображену на мал. 2.19. Пряма, яку зображено на графіку, відома під назвою *лінії безпеки ринку (security market line, SLM)*.

Для емпіричного застосування модель (2.126) часто записують як:



Малюнок 2.19. Модель систематичного ризику

$$R_i - r_f = \beta_i (R_m - r_f) + \varepsilon_i \quad (2.127)$$

або

$$R_i - r_f = \alpha_i + \beta_i (R_m - r_f) + \varepsilon_i. \quad (2.128)$$

Остання з двох моделей відома як модель ринку. МОКА-модель, передбачає, що α_i повинне дорівнювати нулю (мал.2.20).

З викладеного раніше зауважимо, що у (2.128) залежна змінна y є не що інше, як $(R_i - r_f)$, а незалежна змінна x — це β_i , тобто нестійкий коефіцієнт, а не $(R_m - r_f)$. Тому для того, щоб використати регресію (2.128), спочатку треба знати β_i .

Цей приклад показує, що є випадки, які можуть бути описані моделлю з нульовим перетиним, наприклад, гіпотеза неперервного (постійного) прибутку Мілтона Фрідмана, за якою постійне споживання пропорційне постійному прибутку; теорія аналізу витрат, яка стверджує, що змінні витрати виробництва пропорційні випуску продукції; у деяких випадках монетаристська теорія стверджує, що рівень зміни цін (тобто рівень інфляції) пропорційний рівневі зміни пропозиції грошей.

Як оцінювати моделі, у яких перетин дорівнює нулеві; які явища вони відображають? Для того, щоб відповісти на ці запитання, запишемо їх у формі вибіркової простої лінійної регресії, тобто:

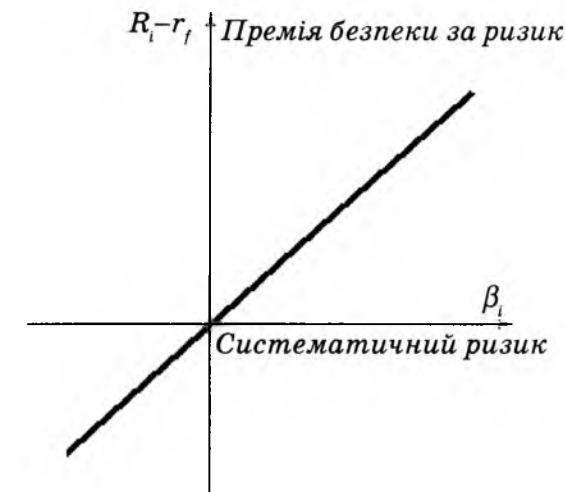
$$y_i = b_1 x_i + e_i. \quad (2.129)$$

Застосувавши метод найменших квадратів до (2.129), отримаємо такі формули для b_1 і його модифікацій:

$$b_1 = \frac{\sum x_i y_i}{\sum x_i^2}; \quad (2.130)$$

$$\text{var}(b_1) = \frac{\sum \hat{\sigma}^2}{\sum x_i^2}, \quad (2.131)$$

де $\hat{\sigma}^2$ оцінюється як:



Малюнок 2.20. Модель ринку “теорії портфеля”

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{N-1}. \quad (2.132)$$

Цікаво порівняти ці формули з тими, що ми отримали, коли величина перетину була наявна у моделі:

$$b_1 = \frac{\sum \tilde{x}_i \tilde{y}_i}{\sum \tilde{x}_i^2}; \quad (2.133)$$

$$\text{var}(b_1) = \frac{\hat{\sigma}^2}{\sum \tilde{x}_i^2}, \quad (2.134)$$

де $\tilde{x}_i = (x_i - \bar{x})$; $\tilde{y}_i = (y_i - \bar{y})$.

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{N-2}. \quad (2.135)$$

Відмінність між цими двома наборами формул є очевидною: в моделі, де величина перетину відсутня, ми використовуємо просту суму квадратів і суму добутків, а в моделі з наявною величиною перетину використовуємо зважену (за середнім) суму квадратів і суму добутків. По-друге, ступені вільності для обчислення $\hat{\sigma}_e^2$ у першому випадку $(N-1)$, а в другому випадку $(N-2)$.

Модель з відсутнім або нульовим перетином також можна застосувати у деяких випадках, проте тут потрібно пам'ятати декілька специфічних рис. По-перше, $\sum e_i$, яка завжди дорівнює нулю в моделі з наявною величиною перетину (загальноприйнята модель), не завжди повинна дорівнювати нулю для регресії, що проходить через початок координат. По-друге, R^2 — коефіцієнт детермінації, що є завжди додатним у загальноприйнятій моделі, у деяких випадках може перетворюватись на від'ємний для регресії, що проходить через початок координат. Цей аномальний результат ми отримуємо тому, що R^2 недвозначно передбачає, що перетин включається в модель. Тому R^2 , який обчислено за правилами, може не відповідати регресійній моделі, що проходить через початок координат.

Саме через ці специфічні риси моделі регресії, що проходить через початок координат, кожен, хто її використовує, має бути дуже обережним. Якщо немає достатньої попередньої упевненості в її використанні, краще використати загальноприйнятую модель із наявним перетином. Це має подвійну перевагу. По-перше, якщо ми включаємо в модель величину перетину, але вона статистично незначима (тобто статистично дорівнює нулю), для усіх практичних цілей ми маємо регресію, що проходить через початок координат. По-друге, і це найважливіше, якщо фактично перетин існує, але ми наполягаємо на тому, щоб застосувати регресію, що проходить через початок координат, то ми припускаємо помилку специфікації, тим самим порушуючи припущення 5 класичної моделі лінійної регресії.

Модель залежності зміни частки заробітної плати населення України у загальному доході

Розглянемо “грошові доходи населення України з вересня 1993 по серпень 1994 року”¹ (табл. 2.21).

Таблиця 2.21

Грошові доходи населення України у 1993—1994 рр. (млрд. крб.)

Місяць	Зарплата у державному секторі	Пенсії, допомоги, стипендії та інші трансферні доходи	Всього грошових доходів	Частка зарплати у загальному доході, %
Вересень — 93	3040	3551	7768	39.1
Жовтень — 93	4349	3253	9101	47.8
Листопад — 93	5382	5932	12997	41.4
Грудень — 93	11493	6601	20369	56.4
Січень — 94	10994	6978	19330	56.9
Лютий — 94	13896	7224	22746	61.1
Березень — 94	17513	7595	27100	64.6
Квітень — 94	20497	8528	31644	64.8
Травень — 94	17646	8061	28578	61.7
Червень — 94	22599	8913	34657	65.2
Липень — 94	24064	10377	38969	61.8
Серпень — 94	23703	11590	41696	56.8

Побудуємо лінійну регресійну модель $\hat{y} = a + bx$, де y — частка зарплати у загальному доході населення, x — місяць.

За даними таблиці, розрахована лінійна модель має вигляд:

$$y = 44.17 + 1.89x.$$

Як бачимо, з кожним місяцем частка зарплати населення в загальному доході збільшувалася на 1.89 млн. крб. (грошових одиниць 1994 року — карбованцях).

Розрахований коефіцієнт кореляції (r) = 0.756882, а коефіцієнт детермінації (R^2) = 57.29%. Оскільки значення коефіцієнта детермінації невелике, для впевненості перевіримо адекватність лінійної моделі за F -критерієм Фішера. Для цього за статистичними таблицями F -розподілу Фішера для 5%-ного рівня значимості (задаємо довільно) та при ступенях вільності відповідно 1 і 10 знайдемо критичне значення $F_{кр} = 4.96$. Розраховане значення F -критерію Фішера відповідно дорівнює: $F = 13.4121$. Як можна помітити, розраховане значення F більше, ніж критичне: $F = 13.4121 > F_{кр} = 4.96$. Отже, можна зробити висновок про адекватність моделі за F -критерієм Фішера.

¹ Тенденції української економіки // Місячний бюлетень. — 1994. — Вересень.

Побудова інтервалів довіри для параметрів

Інтервали довіри для параметрів розраховуються за формулами:

$$\text{для перетину: } b_0 \pm t_{\alpha/2, n-2} \hat{\sigma}_{b_0};$$

$$\text{для нахилу: } b_1 \pm t_{\alpha/2, n-2} \hat{\sigma}_{b_1}.$$

Розраховуємо середнє квадратичне відхилення для перетину і параметра, які відповідно дорівнюватимуть: $\hat{\sigma}_{b_0} = 3.8$ та $\hat{\sigma}_{b_1} = 0.52$.

Значення t -статистики знаходимо за статистичними таблицями t -розподілу Ст'юдента при рівні значимості $\alpha = 0.05$ та ступенях вільності $(n - 2)$, тобто $(12 - 2 = 10)$:

$$t_{(0,025;10)} = 2.22.$$

Таким чином, можна визначити інтервали довіри:

$$\text{для перетину } 44.17 \pm 2.22 \times 3.8;$$

$$\text{для нахилу } 1.8 \pm 2.22 \times 0.52.$$

Як ми бачимо, інтервал довіри для нахилу нестійкий, тому такі моделі потрібно дуже обережно використовувати для прогнозу.

Для ілюстрації все ж таки побудуємо прогноз на вересень 1994 р., тобто на 13-й місяць ($x = 13$). Звідси маємо:

$$y = 44.17 + 1.89 \times 13 = 68.74.$$

Побудуємо інтервали довіри для окремого значення y :

$$\hat{y}_{n+1} \pm t_{\alpha/2, n-2} \hat{\sigma}_e \sqrt{1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}.$$

$$\hat{y}_{n+1} = 68.74,$$

$$t_{(\alpha/2, n-2)} = 2.22,$$

$$\hat{\sigma}_e = 6.17662.$$

Розрахуємо величину під коренем:

$$\sqrt{1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} = 1.13.$$

Отже, остаточно інтервали довіри для окремого значення y в $(n + 1)$ період мають вигляд:

$$68 \pm 15.5.$$

Розглянутий приклад дає основну ідею проведення розрахунків за моделями простої лінійної регресії. Нагадаємо, що це лише інструмент для прогнозів, бо побудова моделі, інтерпретація даних у першу чергу залежить від економічної кваліфікації дослідника.

РОЗДІЛ 3. КРИВІ ЗРОСТАННЯ

3.1. ПОНЯТТЯ ПРО КРИВІ ЗРОСТАННЯ

Криві зростання описують різні тенденції економічних процесів, наприклад, життєвий цикл товару, процес нагромадження капіталу, маркетингові зусилля фірм тощо. Економічна практика вже накопичила певний досвід і певні типи кривих, які найчастіше використовуються в макро- та мікроекономічних дослідженнях. До таких кривих відносяться:

експоненційна: $y = \alpha\beta^x$;

степенева (мультиплікативна): $y = \alpha x^\beta$;

зворотна: $y = \beta_0 + \beta_1 \frac{1}{x}$;

квадратична: $y = b_0 + b_1x + b_2x^2$;

модифікована експонента: $y = \alpha\beta^x + \gamma$;

крива Гомперця: $y = e^{\alpha\beta^x + \gamma}$;

логістична крива: $y = \frac{1}{\alpha\beta^x + \gamma}$.

У загальному випадку однофакторну економетричну модель можна подати у вигляді:

$$y = f(x) + e,$$

де $f(x)$ — одна з функцій зростання; e — випадкова величина.

Як і у випадку з простою лінійною регресією, основне завдання полягає в розрахунку невідомих параметрів кривих зростання і подальшому аналізі обраної моделі. Оцінку невідомих параметрів проводять по-різному: експоненційні функції шляхом логарифмічних перетворень зводять до лінійної регресії, квадратичні функції зводять до багатофакторної регресії, для інших використовують ітеративні методи, метод трьох точок, метод Тейла тощо. В цьому розділі ми особливу увагу звернемо на функції зростання, які шляхом перетворень зводяться до лінійної регресії. Для таких функцій зберігається вся методологія досліджень, яку ми детально розглядали у випадку простої лінійної регресії. Крім того, ми розглянемо найпоширеніші в сучасному економічному аналізі криві, які не зводяться шляхом елементарних перетворень до лінійної регресії, але параметри яких можна легко роз-

рахувати спеціальними спрощеними методами. Слід зазначити, що всі криві зростання, які розглядаються, представлені в пакеті STATGRAPH, тому їх практичне застосування не складне.

3.2. НАЙПРОСТІШІ ПЕРЕТВОРЕННЯ НЕЛІНІЙНИХ МОДЕЛЕЙ У ЛІНІЙНІ. ЕКСПОНЕНЦІЙНА ФУНКЦІЯ. ПРИКЛАДИ ЗАСТОСУВАННЯ ЕКСПОНЕНЦІЙНИХ ФУНКЦІЙ У БІЗНЕСІ ТА ФІНАНСАХ. ЗВЕДЕННЯ ДО ЛІНІЙНОЇ РЕГРЕСІЇ

3.2.1. Експоненційна функція. Приклади застосування експоненційних функцій у бізнесі та фінансах

Експоненційна функція може набирати різних чисельних еквівалентних форм:

$$y = \alpha\beta^x, \text{ основна форма } \beta > 0; \quad (3.1)$$

$$y = \alpha e^{b_1 x}, \beta \text{ замінюємо на } e^{b_1}, \text{ де } b_1 = \ln(\beta); \quad (3.2)$$

$$y = \alpha(1-r)^x, \beta \text{ замінюємо на } 1-r, \text{ де } r = \beta - 1; \quad (3.3)$$

$$y = e^{b_0 + b_1 x}, \alpha \text{ замінюємо на } e^{b_0} \text{ та } \beta \text{ на } e^{b_1}, \text{ де } b_0 = \ln(\alpha) \text{ і } b_1 = \ln(\beta); \quad (3.4)$$

$$y = 10^{a+bx}, \alpha \text{ замінюємо на } 10^a \text{ і } \beta \text{ на } 10^b, \text{ де } a = \log(\alpha), b = \log(\beta). \quad (3.5)$$

Усі ці форми використовуються на практиці для опису різних економічних процесів, наприклад, форму (3.3) найчастіше використовують у фінансах. У цьому разі r інтерпретується як норма річного відсотка. Розглянемо декілька прикладів застосування експоненційної функції у бізнесі та фінансах.

Приклад 1. Припустимо, що капітал C_0 знаходиться у банку протягом t років із річним відсотком r . Він змінюється відповідно до функції (3.3), тобто через t років капітал дорівнюватиме:

$$C_t = C_0(1+r)^t.$$

З експоненційної форми (3.2) легко отримати так зване **правило 70**, поширене в фінансових розрахунках. Правило 70 дає значення часу t , через яке змінна подвоїть своє значення відносно часу 0. Припустимо, що у початковий період часу ми маємо капітал C_0 , а в період часу t — капітал C_t . Тоді подвоєння капіталу можна записати, переходячи від форми (3.3) до форми (3.2):

$$C_t = 2C_0 = C_0 e^{b_1 t}, \text{ звідки } \Rightarrow t = \frac{\ln 2}{b_1} \approx \frac{0.6931}{b_1} \approx \frac{0.70}{b_1}.$$

Отже, капітал подвоїться через $\frac{0.70}{b_1}$ років. Правило має назву 70, тому що $\ln 2$ приблизно дорівнює 0.70.

Приклад 2. Приведення витрат до поточного часу

Порівнюючи різні інвестиційні проекти, ми повинні порівняти в першу чергу кошти, які на них витрачаються, та майбутні прибутки, які вони приносять. Найпростіше це зробити, оцінюючи майбутні витрати та прибутки у вартості поточного року. Для ілюстрації розглянемо приклад. Заощаджуючи в банку A гривень у поточному році при ставці процента r , через t років ми отримаємо:

$$B = Ae^{rt}. \quad (3.6)$$

Відповідно B гривень, отриманих у t році при ставці відсотка r , у поточному році коштуватимуть:

$$A = Be^{-rt}. \quad (3.7)$$

Величина називається приведеними витратами B гривень через t років до поточного часу (при нормі відсотка r).

Дискретною версією виразу (3.6) є відома формула

$$B = A(1+r)^t,$$

яку ми розглядали у першому прикладі, а зворотною до неї відповідно:

$$A = B / (1+r)^t = B(1+r)^{-t}.$$

З дискретною версією легше працювати, коли t — є цілим числом. Якщо t набуває неперервних значень, перевагу слід віддати виразу (3.6) або (3.7).

Приведені витрати можуть також визначатись для потоку платежів. При нормі відсотка r приведені витрати платежу B_1 , який ми мали б через t_1 років, платежу B_2 через t_2 років, ..., платежу B_n відповідно через t_n років мають вигляд:

$$A = B_1e^{-rt_1} + B_2e^{-rt_2} + \dots + B_n e^{-rt_n}. \quad (3.8)$$

Приклад 3. Довічна рента

Довічна рента — це послідовність однакових платежів у рівновіддалені інтервали для певного проміжку часу. Приведені витрати ренти A , яка виплачується наприкінці кожного з N років при постійній нормі відсотка r , що нараховується неперервно, дорівнює:

$$PV = Ae^{-r1} + Ae^{-r2} + \dots + Ae^{-rN} = A(e^{-r} + e^{-r2} + \dots + e^{-rN}), \quad (3.9)$$

де PV — приведені витрати.

Враховуючи, що вираз $(a + \dots + a^n)(1 - a) = a - a^{n+1}$, його легко переписати у вигляді:

$$(a + \dots + a^n) = \frac{a(1 - a^n)}{1 - a}. \quad (3.10)$$

Якщо у виразі (3.9) припустити, що $a = e^{-r}$, а $n = N$, тоді приведені витрати довічної ренти можна записати у вигляді:

$$PV = A \frac{e^{-r}(1 - e^{-rN})}{1 - e^{-r}} = \frac{A(1 - e^{-rN})}{e^r - 1}. \quad (3.11)$$

За припущення, що ми будемо отримувати ренту довічно, тобто нескінченну кількість років ($N \rightarrow \infty$), вираз (3.11) набуде вигляду:

$$PV = \frac{A}{e^r - 1} \quad (3.12)$$

за припущення, що $e^{-rN} \rightarrow 0$ та $N \rightarrow \infty$.

Якщо ж відсоток нараховується щорічно, приведену довічну ренту до поточного часу краще подавати у відомому вигляді:

$$PV = \frac{A}{1+r} + \dots + \frac{A}{(1+r)^N}. \quad (3.13)$$

Якщо ми скористаємось формулою (3.10) та покладемо $a = 1/(1+r)$, $n = N$, то отримаємо:

$$PV = A \frac{1/(1+r)}{r/(1+r)} \left(1 - \left(\frac{1}{1+r} \right)^N \right) = \frac{A}{r} \left(1 - \left(\frac{1}{1+r} \right)^N \right). \quad (3.14)$$

Якщо $N \rightarrow \infty$, то вираз (3.14) остаточно набере вигляду: $PV = \frac{A}{r}$.

Взагалі експоненційні криві використовуються для опису швидко зростаючих або спадаючих економічних процесів. При цьому, якщо $\beta > 0$ ($b_1 > 0$) — функція зростає до нескінченності,

якщо $\beta < 0$ ($b_1 < 0$) — функція спадає до 0.

Параметр b можна інтерпретувати як коефіцієнт зростання у часі.

3.2.2. Зведення експоненційної кривої до простої лінійної функції

Звичайно, перед нами постає запитання: як можна розрахувати невідомі параметри експоненційної кривої? Покажемо, що шляхом логарифмічного перетворення можна легко звести експоненційну криву у будь-якій формі до лінійної функції, що дає змогу розраховувати параметри методом найменших квадратів та використовувати подальший аналіз моделі, як і в разі простої лінійної регресії. Отже, маємо:

$$\begin{aligned} y &= \alpha\beta^x, & \ln(y) &= \ln(\alpha) + x \ln(\beta); \\ y &= \alpha e^{b_1 x}, & \ln(y) &= \ln(\alpha) + b_1 x; \\ y &= \alpha(1-r)^x, & \ln(y) &= \ln(\alpha) + x \ln(1-r); \\ y &= e^{b_0 + b_1 x}, & \ln(y) &= b_0 + b_1 x; \\ y &= 10^{b_0 + b_1 x}, & \ln(y) &= b_0 + b_1 x. \end{aligned}$$

Проводячи необхідну заміну змінних, переходимо до моделі лінійної регресії, розрахунок параметрів та всі необхідні дослідження якої ми розглядали детально у попередньому розділі.

3.3. СТЕПЕНЕВА (МУЛЬТИПЛІКАТИВНА) ФУНКЦІЯ. ЗВЕДЕННЯ ДО ЛІНІЙНОЇ РЕГРЕСІЇ. ПРИКЛАДИ ЗАСТОСУВАННЯ СТЕПЕНЕВИХ ФУНКЦІЙ У БІЗНЕСІ ТА ФІНАНСАХ

Степенева функція є однією з найпоширеніших у практиці кривих зростання і описує дуже широкий спектр економічних процесів. Вона має такий вигляд:

$$y = \alpha x^\beta. \quad (3.15)$$

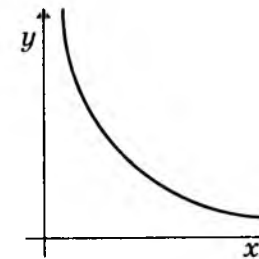
Ми розглядатимемо випадок, коли параметр $\alpha \geq 0$, що є типовим для економічних процесів. Якщо значення параметра β — не ціле число, то розглядають лише випадок, коли $x \geq 0$. При цьому залежно від знака параметра β степенева функція описуватиме різні економічні процеси: прискорене зростання, уповільнене зростання та спад. Слід зазначити, що якщо $\beta=1$, степенева функція перетворюється на лінійну. Ці різні ситуації зображені на мал. 3.1 (а, б, в, г).

Якщо параметр β степеневої функції — ціле число, то залежно від того, парне чи непарне його значення, графік функції має різний вигляд. Якщо β — парне, тобто його можна записати у вигляді: $\beta = 2k$, $k \in \mathbb{Z}_+$ (розглянемо тільки випадок, коли $k > 0$), тоді $y \in [0, +\infty)$, а графік функції симетричний відносно осі ординат (мал. 3.2, а).

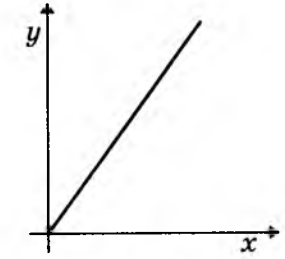


а) $\beta > 1$ (Прискорене зростання)

б) $0 < \beta < 1$ (Уповільнене зростання)



в) $\beta < 0$

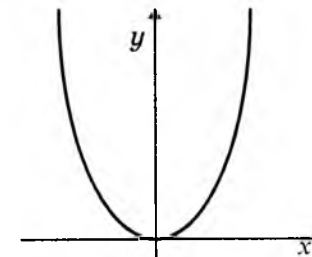


г) $\beta = 1$

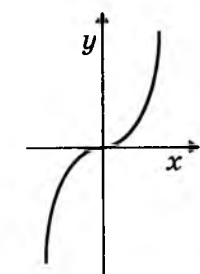
Малюнок 3.1. Вигляд степеневої функції, коли β — не ціле число, $x \geq 0$

Якщо параметр β — непарне число, тобто його можна подати у вигляді: $\beta = 2k - 1$, $k \in \mathbb{Z}_+$ (розглянемо тільки випадок, коли $k > 0$), тоді $y \in \mathbb{R}$, а графік функції симетричний відносно початку координат (0,0) (мал. 3.2, б).

Звичайно, основне питання полягає в тому, щоб розрахувати невідомі параметри мультиплікативної (степеневої) кривої. Покажемо, що шляхом логарифмічного перетворення ми можемо легко звести степеневу криву,



а) β — парне



б) β — непарне

Малюнок 3.2. Графік степеневої функції, коли β — ціле число

так само, як і експоненційну, до лінійної функції, що дає нам змогу розраховувати параметри методом найменших квадратів. Справді, логарифмуючи праву та ліву частини (3.15), отримаємо:

$$\ln y = \ln \alpha + \beta \ln x,$$

$$\ln y = z, \ln \alpha = \beta_0, \ln x = x_1 \Rightarrow z = \beta_0 + \beta x_1.$$

На практиці степеневі функції використовуються для опису різних економічних процесів. Найвідомішою з них є виробнича функція Кобба — Дугласа. Крім того, вони застосовуються для опису кривих байдужості, а також попиту на товари різних категорій, так звана крива Торнквіста та ін.

3.4. ЗВОРОТНІ ПЕРЕТВОРЕННЯ. ПРИКЛАДИ ЇХ ЗАСТОСУВАННЯ НА ПРАКТИЦІ

Узагальнена зворотна модель має вигляд:

$$y_i = \beta_0 + \beta_1 \left(\frac{1}{x_i} \right) + \varepsilon_i. \quad (3.16)$$

Вона нелінійна за змінною x , але лінійна за параметрами β_0 і β_1 , і тому є лінійною регресійною моделлю.

Справді, позначивши $\frac{1}{x_i} = z_i$, отримаємо:

$$y_i = \beta_0 + \beta_1 z_i + \varepsilon_i. \quad (3.17)$$

Вибіркова зворотна модель, враховуючи позначення попереднього розділу, може бути записана у вигляді:

$$y = b_0 + b_1 \frac{1}{x} + e, \quad (3.18)$$

де b_0 та b_1 — невідомі параметри, які необхідно знайти; e — помилка.

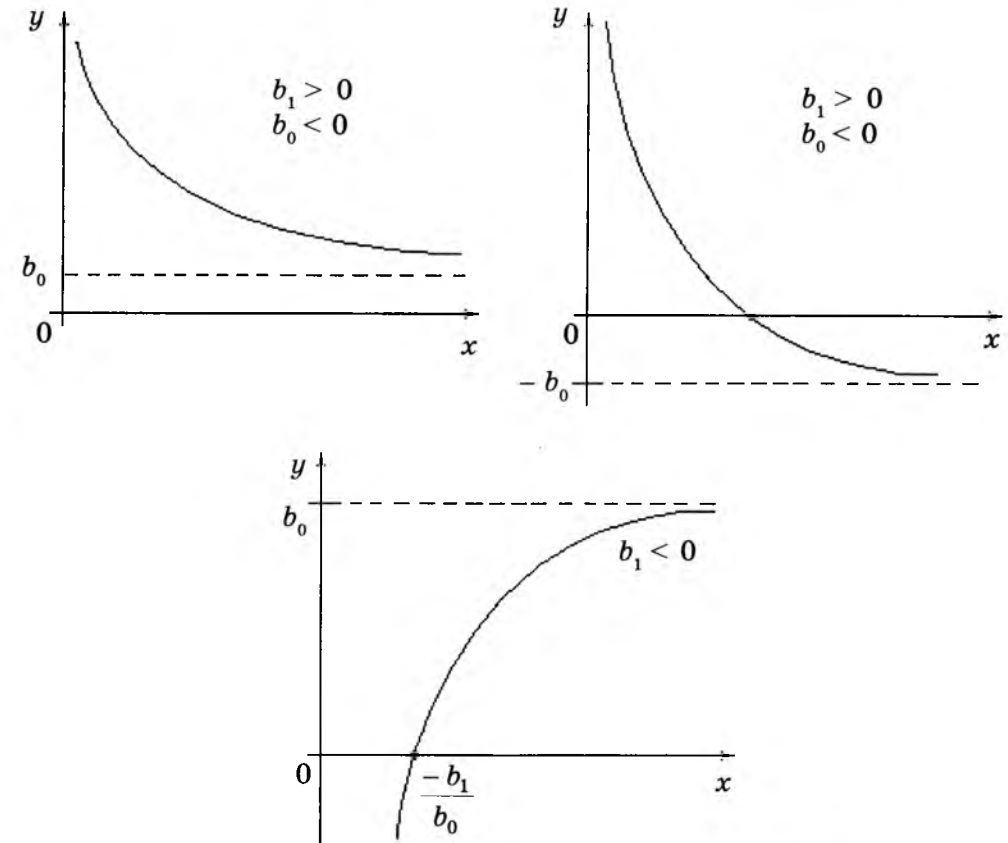
Модель (3.18) має свої особливості, на відміну від простої лінійної регресії: коли x прямує до нескінченності, величина $b_1 \left(\frac{1}{x} \right)$ прямує до нуля, а y прямує до граничного значення.

Вигляд моделі (3.18) значною мірою залежить від знака параметрів b_0 і b_1 (мал. 3.3).

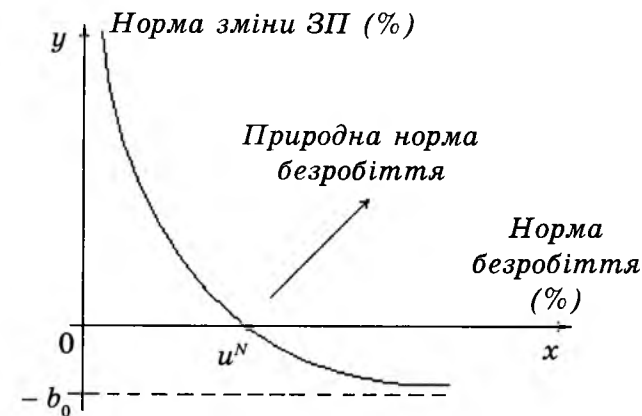
Нахил моделі (3.18) визначаємо за такою формулою: $\frac{dy}{dx} = -b_1 \left(\frac{1}{x^2} \right)$. Він є додатним, коли $b_1 < 0$, і від'ємним, коли $b_1 > 0$.

Яскравим прикладом використання зворотних моделей в економіці (зокрема в макроекономіці) є відома крива Філіпса (мал. 3.2).

Базуючись на даних норми процента зміни заробітної плати і процента безробіття для Англії за період з 1861 по 1975 рр., Філіпс побудував криву,



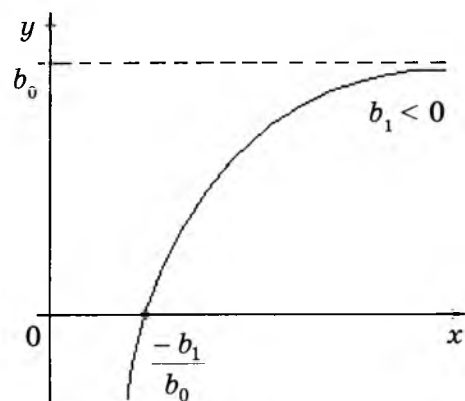
Малюнок 3.3. Зворотна функція: $y = b_0 + b_1 \left(\frac{1}{x} \right)$



Малюнок 3.4. Крива Філіпса

яку в сучасній інтерпретації подано на мал. 3.4.

Як показано на мал. 3.4, асимптота (границя зміни заробітної плати) пов'язана з параметром b_0 . Точка U^N є значенням природної норми безробіття; коли $x < U^N$, то норма зміни заробітної плати додатна, а коли значення x стає більшим, ніж природна норма безробіття, y (норма зміни заробітної



Малюнок 3.5. Крива Енгеля

платні) буде від'ємною.

Крива Філіпса дає змогу розрахувати мінімальну заробітну плату, компенсацію за безробіття тощо.

Інший, не менш важливий випадок використання зворотної кривої (мал. 3.5) — крива витрат Енгеля, яка пов'язує споживчі витрати на товари із загальними витратами або доходом.

Якщо ми позначимо через y витрати на споживання, а через x — доход, то крива Енгеля для певного товару виявить такі особливості:

а) критичний рівень доходу, нижче від якого товар не буде куплено (це значення на мал. 3.5 дорівнює $-(b_1/b_0)$;

б) "стелю" (межу) насичення, яку не можна збільшити, як би не зростав доход (на мал. 3.5 це значення дорівнює b_0).

3.5. КВАДРАТИЧНІ ФУНКЦІЇ

Квадратичні функції використовуються для опису дуже широкого спектру економічних процесів, завдяки їхнім універсальним властивостям. Справді, у загальному випадку квадратична функція має вигляд:

$$y = b_0 + b_1x + b_2x^2. \tag{3.19}$$

Якщо фактор x інтерпретувати як змінну часу t , то (3.19) можна переписати у вигляді:

$$y = b_0 + b_1t + b_2t^2. \tag{3.20}$$

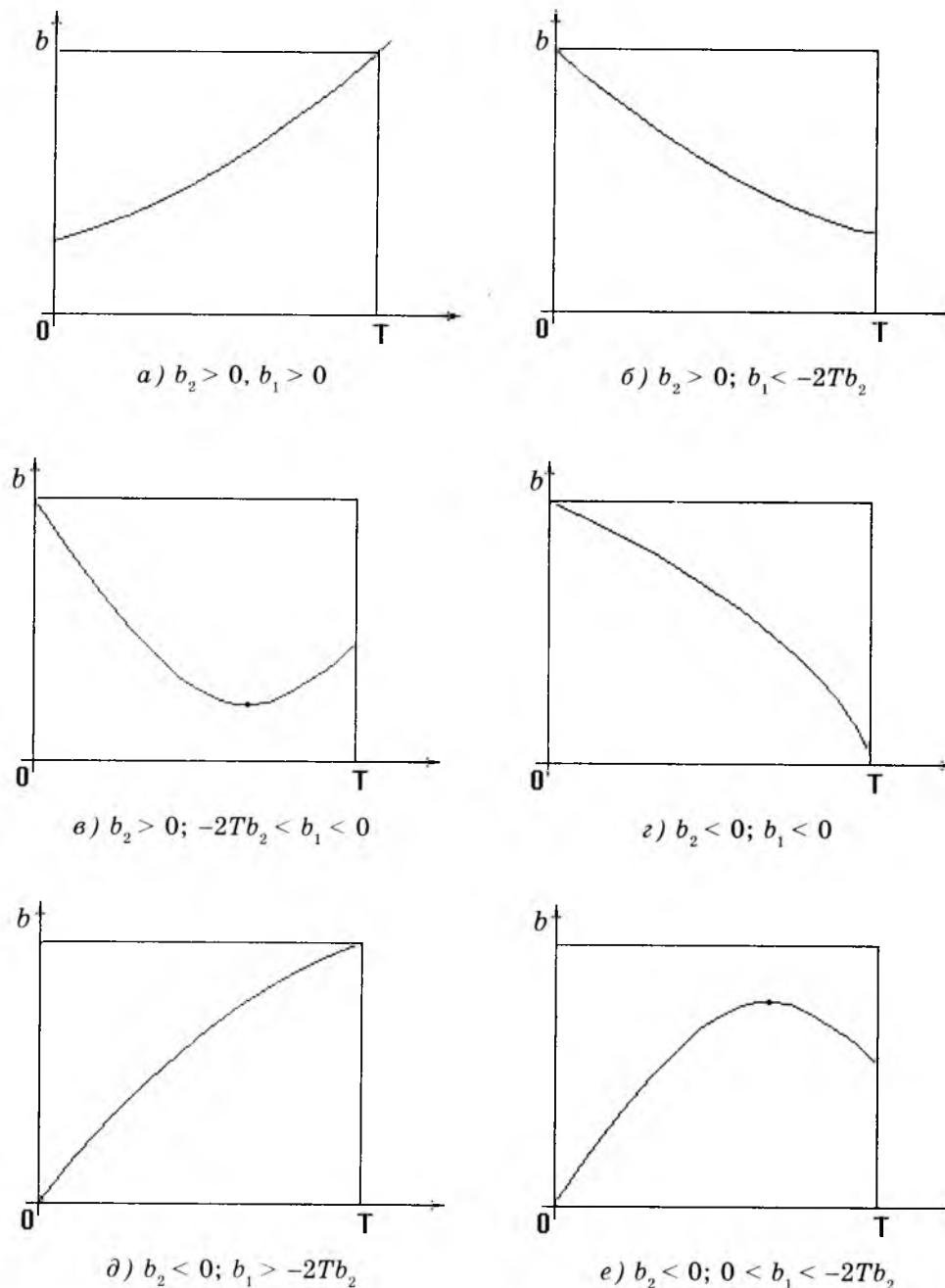
Виходячи із значень b_0, b_1, b_2 , крива може відобразити еволюцію — дуже різну на інтервалах часу від 1 до T .

Якщо $b_2 > 0$, маємо параболу на міні, при $b_2 < 0$ — параболу на максимумі. Екстремум функції досягається у точці $t = -b_1/2b_2$.

Проілюструємо різні можливості квадратичної функції. Взагалі можливі шість різних випадків.

1. Коли параметри $b_2 > 0, b_1 > 0$, квадратична функція описує поліпшене зростання. У цьому разі вершина розташована перед часом 0 (мал. 3.6, а).

2. Коли параметри $b_2 > 0; b_1 < -2Tb_2$, квадратична функція описує уповільнений спад (вершина після часу T) (мал. 3.6, б).



Малюнок 3.6. Квадратична функція

3. Коли параметри $b_2 > 0$; $-2Tb_2 < b_1 < 0$, маємо класичний випадок — параболу з \min (мал. 3.6, в).

4. Коли параметри $b_2 < 0$; $b_1 < 0$, квадратична функція описує прискорений спад (вершина перед 0, мал. 3.6, з).

5. Коли параметри $b_2 < 0$; $b_1 > -2Tb_2$, квадратична функція описує уповільнене зростання (вершина після часу T , мал. 3.6, д).

6. Коли параметри $b_2 < 0$; $0 < b_1 < -2Tb_2$, ми також маємо класичний випадок — параболу на \max (мал. 3.6, е).

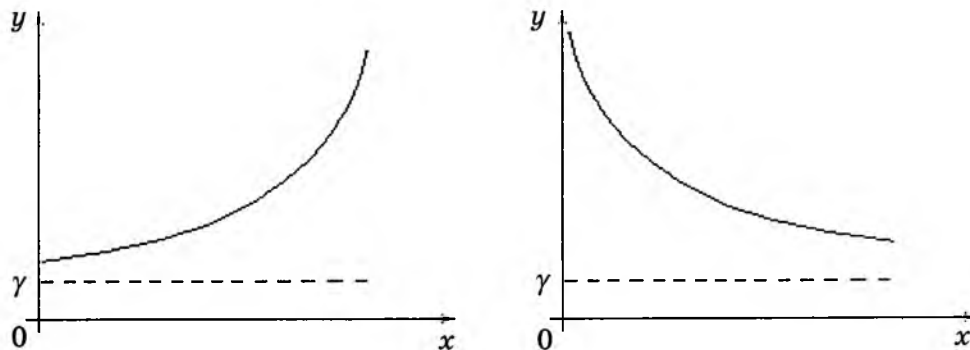
Шляхом елементарної заміни змінних, а саме: $x_1 = t$; $x_2 = t^2$ квадратичну функцію зводять до багатофакторної регресії, тому невідомі параметри розраховуються як для випадку багатофакторної регресії (див. розділ 4).

3.6. ЕКСПОНЕНЦІЙНА МОДИФІКОВАНА КРИВА

Експоненційна модифікована крива, або, як її частіше називають, модифікована експонента, використовується для опису економічних процесів, які обмежені знизу (мал. 3.7). Її широко використовують у дослідженні ринку. Модифікована експонента має вигляд:

$$y = \alpha\beta^x + \gamma. \quad (3.21)$$

При $\beta > 1$ модифікована експонента спочатку повільно, а потім швидко зростає і обмежена знизу прямою $x = \gamma$. При $\beta < 1$ модифікована експонента спадає, та її значення також обмежене знизу значенням γ .



Малюнок 3.7. Модифікована експонента

Наявність параметра γ змінює проблему оцінки параметрів і заважає використовувати лінійну регресію шляхом логарифмування правої та лівої частин (3.21). Справді, прологарифмувавши (3.21), отримаємо:

$$\ln(y) = \ln(\alpha\beta^x + \gamma). \quad (3.22)$$

Логарифм у правій частині (3.22) ми далі розбити не можемо, отже, не можемо зробити ніякої заміни змінних, як у випадку простої експоненти.

Якщо γ відоме, то рівняння (3.21) набуває вигляду:

$y - \gamma = \alpha\beta^x$, тобто є випадком експоненційної кривої, яку ми легко отримаємо, замінивши $y - \gamma$ на z . У цьому разі ми знаходимо невідомі параметри шляхом логарифмування і зведення до простої лінійної регресії.

Якщо γ не відомо, то не можна обчислити $\ln(y - \gamma)$, або розбити вираз $\ln(y) = \ln(\alpha\beta^x + \gamma)$. Для обчислення невідомих параметрів необхідно застосувати методи нелінійного оцінювання, які в даному підручнику ми не розглядаємо.

3.7. КРИВА ГОМПЕРЦЯ

Крива Гомперця використовується в основному для опису процесів з насиченням і дуже поширена в демографії, маркетингу, дослідженні ринку, збуту продукції. Вона має такий вигляд:

$$y = e^{\alpha\beta^x + \gamma}, \text{ де } 0 < \beta < 1. \quad (3.23)$$

Зазначимо спочатку, що шляхом логарифмування і подальшої заміни змінних криву Гомперця легко звести до модифікованої експоненти: $\ln(y) = \alpha\beta^x + \gamma$.

Дослідимо властивості кривої Гомперця.

1. Граничні точки, коли $x \Rightarrow \pm\infty$, будуть такими:

$$\lim_{x \rightarrow \infty} y = e^\gamma, \quad \lim_{x \rightarrow -\infty} y = \begin{cases} +\infty, & \alpha > 0 \\ 0, & \alpha < 0 \end{cases}$$

2. Перша похідна функції має вигляд:

$$y' = (\alpha\beta^x + \gamma)' e^{\alpha\beta^x + \gamma} = \alpha \ln(\beta) \beta^x y.$$

Її знак збігається із знаком параметра α , коли $0 < \beta < 1$ ($\ln(\beta) < 0$). Звідси ми можемо зробити висновок, що функція спадає, коли $\alpha > 0$, і зростає, коли $\alpha < 0$.

3. Друга похідна дорівнює:

$$y'' = \alpha \ln(\beta) [\ln(\beta) \beta^x y + \beta^x y'] = \alpha \times \ln(\beta) \beta^x [\ln(\beta) y + \alpha \ln(\beta) \beta^x y] = \alpha \ln^2(\beta) \beta^x y (1 + \alpha \beta^x).$$

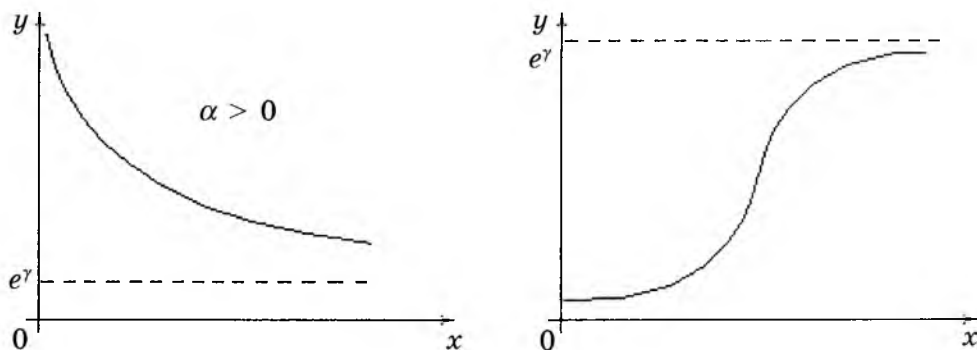
Знак цієї похідної збігається із знаком виразу — $\alpha(1 + \alpha\beta^x)$.

Якщо $\alpha > 0$, то функція опукла, в протилежному випадку, коли $\alpha < 0$, крива Гомперця є s-кривою та має точку перегину, коли $1 + \alpha\beta^x = 0$, тобто $x = \ln(-1/\alpha)/\ln(\beta)$.

У цій точці $y = e^{-1+\gamma}$. Точка перегину входить до інтервалу спостережень $[1, x]$, якщо $\alpha\beta < -1$.

Якщо $\alpha > 0$, крива спадає асимптотично до e^γ (мал. 3.8).

Якщо $\alpha < 0$, вона має форму s-кривої, тобто спочатку зростає швидко, а потім повільно. Такою функцією можна описати типову еволюцію продажу товару.



Малюнок 3.8. Крива Гомперця

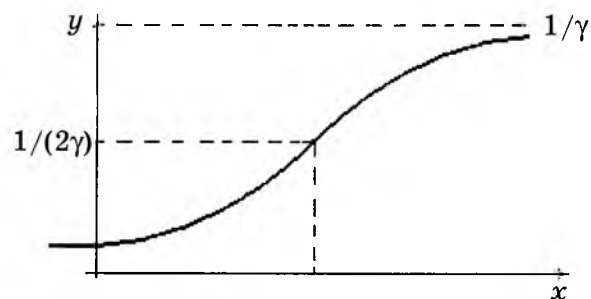
3.8. ЛОГІСТИЧНА КРИВА

Логістична крива є зворотною функцією до модифікованої експоненти, вона має такий вигляд:

$$y = \frac{1}{\alpha\beta^x + \gamma}, \text{ де } 0 < \beta < 1, \alpha > 0, \gamma > 0. \quad (3.24)$$

Логістична крива є типовою s-подібною кривою, у якій точка перетину припадає на середину стелі (мал. 3.9).

Дослідимо властивості цієї кривої.



Малюнок 3.9. Логістична крива

1. Асимптоти функції дорівнюють:

$$\lim_{x \rightarrow \infty} y = \frac{1}{\gamma}; \quad \lim_{x \rightarrow -\infty} y = 0.$$

2. Перша похідна функції:

$$y' = \frac{-(\alpha\beta^x + \gamma)'}{(\alpha\beta^x + \gamma)^2} = -\alpha \ln(\beta) \beta^x y^2,$$

коли $\alpha > 0, 0 < \beta < 1 \Rightarrow \ln(\beta) < 0$ функція зростає.

3. Друга похідна має вигляд:

$$\begin{aligned} y'' &= -\alpha \ln(\beta) [\ln(\beta) \beta^x y^2 + 2\beta^x y y'] = \\ &= -\alpha \ln^2(\beta) \beta^x y^3 \left[\frac{1}{y} - 2\alpha\beta^x \right] = -\alpha \ln^2(\beta) \beta^x y^3 (\gamma - \alpha\beta^x). \end{aligned}$$

Знак другої похідної збігається із знаком виразу: $-\alpha(\gamma - \alpha\beta^x)$. Логістична крива має точку перегину, коли $\gamma - \alpha\beta^x = 0$, відповідно $x = \ln(\gamma/\alpha) \times \ln(\beta)$. У цій точці $y = 1/(2\gamma)$, тобто ордината точки перегину є серединою “стелі”.

Таким чином, логістична крива — це s-подібна крива з “підлогою” 0 на початку і “стелею”, що дорівнює $1/\gamma$.

3.9. ДЕКІЛЬКА ПРОСТИХ МЕТОДІВ ОБЧИСЛЕННЯ НЕВІДОМИХ ПАРАМЕТРІВ НЕЛІНІЙНИХ МОДЕЛЕЙ. МЕТОД ТРЬОХ ТОЧОК

Оцінку параметрів кривих зростання з трьома параметрами (модифікована експонента, крива Гомперця і логістична крива) можна провести на основі нелінійної регресії, яка вимагає ітеративної процедури мінімізації функції (сума квадратів помилок) багатьох змінних (параметрів α, β, γ).

В особливих випадках невідомі параметри цих функцій можна розрахувати досить простими методами, не використовуючи нелінійних методів. До числа спрощених методів належить, наприклад, метод трьох точок, який ми і розглянемо. Метод трьох точок є зручнішою процедурою і часто взагалі достатньою.

Припустимо, що дані є доступними у період $t = 1, T$ і що функція має вигляд модифікованої експоненти:

$$y = \alpha\beta^t + \gamma.$$

У випадку кривої Гомперця і логістичної кривої формули адаптуємо через заміну y_t на $\ln(y_t)$ і $(1/y_t)$.

Алгоритм методу трьох точок

Алгоритм складається з певних етапів.

Етап 1. На першому етапі необхідно розділити дані на 3 підмножини I, II, III, однакові за кількістю елементів. Точніше:

а) якщо вся кількість елементів без залишку ділиться на 3, тобто $T = 3\tau$, маємо три підмножини з кількістю елементів у кожній: $T/3 = \tau$;

б) якщо $T = 3\tau + 1 \Rightarrow$ кількість елементів $(\tau + 1)$ входить до II, а множини I і III складатимуться з τ елементів;

в) якщо $T = 3\tau + 2$, тоді τ входить до II множини, а $(\tau + 1)$ до I, і до III множин.

Етап 2. На другому етапі необхідно обчислити значення медіан у трьох підмножинах. Позначимо ці значення відповідно: y_I, y_{II}, y_{III} .

Етап 3. Передбачає розв'язання системи 3-х нерівностей (нелінійних) з трьома невідомими:

$$y_I = \alpha\beta^{t_I} + \gamma \quad (3.25)$$

$$y_{II} = \alpha\beta^{t_{II}} + \gamma \quad (3.26)$$

$$y_{III} = \alpha\beta^{t_{III}} + \gamma \quad (3.27)$$

Виходячи з того, що рівняння є нелінійними, систему можна розв'язати таким чином:

а) спочатку визначити різницю між рівняннями (3.27) і (3.26) і між рівняннями (3.26) і (3.25):

$$y_{III} - y_{II} = \alpha(\beta^{t_{III}} - \beta^{t_{II}}); \quad (3.28)$$

$$y_{II} - y_I = \alpha(\beta^{t_{II}} - \beta^{t_I}) \quad (3.29)$$

б) поділити почленно рівняння (3.28) на рівняння (3.29), записавши $\Delta = t_{III} - t_{II} = t_{II} - t_I$:

$$\frac{y_{III} - y_{II}}{y_{II} - y_I} = \frac{\beta^{t_{III}} - \beta^{t_{II}}}{\beta^{t_{II}} - \beta^{t_I}} = \frac{\beta^{t_{II}}(\beta^\Delta - 1)}{\beta^{t_I}(\beta^\Delta - 1)} = \beta^\Delta,$$

$$\text{звідки } \ln(\beta) = \frac{1}{\Delta} \ln\left(\frac{y_{III} - y_{II}}{y_{II} - y_I}\right).$$

в) визначивши β , повертаємось до (3.28):

$$\alpha = \frac{y_{III} - y_{II}}{\beta^{t_{III}} - \beta^{t_{II}}},$$

і, виходячи з (3.25)

$$\gamma = y_I - \alpha\beta^{t_I}.$$

Таким чином, ми оцінили всі невідомі параметри. Інколи для практичних розрахунків можна замінити медіани середніми величинами на етапі 2.

3.10. ЗВ'ЯЗОК МІЖ КОЕФІЦІЄНТАМИ ЕЛАСТИЧНОСТІ І ПАРАМЕТРАМИ КРИВИХ ЗРОСТАННЯ

В економетриці значну роль відіграють коефіцієнти еластичності, які можна отримати через оцінені параметри кривих зростання. Нагадаємо загальну формулу для розрахунку коефіцієнтів еластичності:

$$(dy/y)/(dx/x) = (dy/dx)(x/y).$$

У випадку лінійної моделі коефіцієнт еластичності дорівнюватиме $b_1 \left(\frac{x}{y}\right)$. Зазначимо, що він залежить не тільки від параметра b_1 , а й від значень x і y (на практиці, коли x і y не специфіковано, еластичність розраховуємо для середніх значень \bar{x} і \bar{y}).

Для кривої, яка шляхом логарифмування зводиться до вигляду:

$$\ln y = b_0 + b_1 \ln x, \quad (3.30)$$

коефіцієнт еластичності є сталою величиною і дорівнює b_1 .

Справді, ми маємо $d(\ln x)/dx = 1/x$, або $d(\ln x) = dx/x$.

Якщо брати до уваги дуже малі зміни, то можна записати приблизно: $(\ln x_t - \ln x_{t-1}) = (x_t - x_{t-1})/x_{t-1}$ — відносна зміна x . Звідси можна отримати:

1) абсолютну зміну:

$$(x_t - x_{t-1});$$

2) відносну зміну:

$$\frac{(x_t - x_{t-1})}{x_{t-1}} = \frac{\Delta x_t}{x_{t-1}};$$

3) відносну зміну у процентах, або процент зростання:

$$\left(\frac{x_t - x_{t-1}}{x_{t-1}}\right) \times 100\%.$$

Модель типу (3.30) називається *log-linear-моделлю*.

Для кривої зростання, яка шляхом перетворення зводиться до вигляду:

$$\ln y = b_0 + b_1 x, \quad (3.31)$$

або до вигляду:

$$y = b_0 + b_1 \ln x, \quad (3.32)$$

коефіцієнти еластичності мають відповідно вигляд: $b_1(x)$ і $b_1\left(\frac{1}{y}\right)$.

Модель типу (3.31) відома під назвою *log-lin-моделі*, а (3.32) — *lin-log-моделі*.

Для *зворотної моделі* коефіцієнт еластичності дорівнює

$$-b_1\left(\frac{1}{xy}\right).$$

Зв'язок між нахилом і коефіцієнтом еластичності для основних типів моделей наведено в табл. 3.1.

Таблиця 3.1

Модель	Загальний вигляд	Нахил (= dy/dx)	Коефіцієнт еластичності (= $dy/dx \cdot x/y$)
Лінійна	$y = b_0 + b_1x$	b_1	$b_1(y/x)$
log-linear	$\ln y = b_0 + b_1 \ln x$	$b_1(y/x)$	b_1
log-lin	$\ln y = b_0 + b_1x$	$b_1(y)$	$b_1(x)^*$
lin-log	$y = b_0 + b_1 \ln x$	$b_1(1/x)$	$b_1(1/y)^*$
Зворотня	$y = b_0 + b_1(1/x)$	$-b_1(1/x^2)^*$	$-b_1(1/xy)^*$

Зазначимо, що коефіцієнт еластичності є змінною, тобто залежить від y , або x , або від обох цих значень одразу. На практиці, коли ці значення не специфіковано, їх дуже часто замінюють на середні значення (\bar{x} і \bar{y}).

ПІДСУМУЄМО ВИВЧЕНЕ

Після вивчення цього розділу ви зможете

1. Використовувати криві зростання для опису нелінійного зв'язку між двома змінними.
2. Зводити шляхом перетворень криві зростання до простої лінійної регресії.
3. Розраховувати невідомі параметри кривих зростання.
4. Прогнозувати за кривими зростання.
5. Мати уявлення про криві зростання, які шляхом перетворень неважко звести до простої регресії.
6. Вміти обчислювати невідомі параметри для кривих методом трьох точок (особливий випадок).
7. Визначити коефіцієнти еластичності через оцінені параметри кривих зростання.

Короткий огляд розділу

1. Криві зростання використовують для опису нелінійного зв'язку між двома змінними.
2. Певний клас кривих зростання шляхом перетворень можна звести до простої лінійної регресії. Для таких кривих залишаються правильними всі дії, які були правильними для простої лінійної регресії, а саме: обчислювання невідомих параметрів проводиться методом найменших квадратів; параметри тестуються на значимість за t -критерієм Ст'юдента; будуються інтервали довіри для параметрів і для прогнозного значення; модель тестується на адекватність за F -критерієм Фішера.
3. Певний клас кривих зростання (модифікована експонента, крива Гомперца, логістична крива) є нелінійними за своїми параметрами. Для розрахунку невідомих параметрів у таких випадках використовуються методи нелінійної регресії. У особливих випадках можливе також застосування спрощених методів оцінювання, наприклад методу трьох точок.
4. В економетриці значну роль відіграють коефіцієнти еластичності, які можна отримати через оцінені параметри кривих зростання. Загальна формула для розрахунку коефіцієнтів еластичності має вигляд:

$$(dy/y)/(dx/y) = (dy/dx)(x/y).$$

Основні формули**Узагальнена регресійна модель за кривою зростання:**

$$y = f(x) + \varepsilon.$$

Вибіркова регресійна модель за кривою зростання:

$$y = f(x) + e.$$

Основні типи кривих зростання:експоненційна: $y = \alpha\beta^x$;степенева: $y = \alpha x^\beta$;зворотна: $y = b_0 + b_1 \frac{1}{x}$;квадратична: $y = b_0 + b_1 x + b_2 x^2$;модифікована експонента: $y = \alpha\beta^x + \gamma$;крива Гомперця: $y = e^{\alpha\beta^x + \gamma}$;логістична крива: $y = \frac{1}{\alpha\beta^x + \gamma}$.**Форми експоненційної кривої:** $y = \alpha\beta^x$, основна форма $\beta > 0$. $y = \alpha e^{b_1 x}$, замінюємо на e^{b_1} , де $b_1 = \ln(\beta)$. $y = \alpha(1-r)^x$, β замінюємо на $(1-r)$, де $r = \beta - 1$. $y = e^{b_0 + b_1 x}$, α замінюємо на e^{b_0} , β на e^{b_1} , де $b_0 = \ln(\alpha)$, $b_1 = \ln(\beta)$. $y = 10^{b_0 + b_1 x}$, α замінюємо на 10^{b_0} і β на 10^{b_1} , де $b_0 = \log(\alpha)$, $b_1 = \log(\beta)$.**Перетворення для експоненційної кривої:**

$$y = \alpha\beta^x, \quad \ln(y) = \ln(\alpha) + x\ln(\beta).$$

$$y = \alpha e^{b_1 x}, \quad \ln(y) = \ln(\alpha) + x b_1.$$

$$y = \alpha(1-r)^x, \quad \ln(y) = \ln(\alpha) + x\ln(1-r).$$

$$y = e^{b_0 + b_1 x}, \quad \ln(y) = b_0 + b_1 x.$$

$$y = 10^{b_0 + b_1 x}, \quad \log(y) = b_0 + b_1 x.$$

Перетворення для степеневі функції:

$$y = \alpha x^\beta, \quad \ln y = \ln \alpha + \beta \ln x.$$

Перетворення для зворотної функції:

$$y = b_0 + b_1 \frac{1}{x}, \quad y = b_0 + b_1 z.$$

Перетворення для квадратичної функції:

$$y = b_0 + b_1 x + b_2 x^2, \quad y = b_0 + b_1 z_1 + b_2 z_2.$$

Коефіцієнт еластичності:

$$(dy/y)/(dx/x) = (dy/dx)(x/y).$$

Зведена таблиця коефіцієнтів еластичності (табл. 3.1).**ВПРАВИ І КОМЕНТАРІ****Вправа 3.1. Крива Філіпса для Англії 1950—1966 рр.**Побудувати залежність між річною зміною заробітної плати у відсотках (y) і нормою безробіття (x) для Англії 1950—1966 рр. (табл. 3.2).

Таблиця 3.2

Рік	Норма зміни заробітної плати, % y	Безробіття, % x
1950	1.8	1.4
1951	8.5	1.1
1952	8.4	1.5
1953	4.5	1.5
1954	4.3	1.2
1955	6.9	1.0
1956	8.0	1.1
1957	5.0	1.3
1958	3.6	1.8
1959	2.6	1.9
1960	2.6	1.5
1961	4.2	1.4
1962	3.6	1.8
1963	3.7	2.1
1964	4.8	1.5
1965	4.3	1.3
1966	4.6	1.4

Розв'язок

Залежність між зміною норми заробітної плати та нормою безробіття в макроекономіці описується кривою Філіпса. Проведені за даними (таблиці) розрахунки дали змогу побудувати таку модель:

$$\hat{y}_t = -1.4284 + 8.7243 \frac{1}{x_t}$$

$$(2.0675) (2.8478) , R^2 = 0.38 , F_{1,15} = 9.39 ,$$

де в дужках наведено стандартне квадратичне відхилення для параметрів.

Оцінена крива дає нам граничне значення \hat{y} (-1.43%), тобто, коли x нескінченно зростає, відсоток заробітної плати знижується, але не може переходити межу 1.43% на рік.

ВИКОНАЙТЕ САМОСТІЙНО**Вправи. Криві зростання**

Вправа 3.2. У табл. 3.3 наведено умовні дані за дефлятором ВВП (валовий внутрішній продукт) для вітчизняних товарів і дефлятор для імпортованих товарів за 1978—1992 рр. Дефлятор часто використовується як показник інфляції споживчих цін (індекс споживчих цін).

Таблиця 3.3

Рік	ВВП дефлятор для вітчизняних продуктів y	ВВП дефлятор для імпорту x
1978	1000	100.0
1979	1023	104.2
1980	1040	109.2
1981	1087	110.5
1982	1146	111.0
1983	1285	125.7
1984	1485	174.9
1985	1521	177.0
1986	1543	188.9
1987	1567	197.4
1988	1592	201.5
1989	1714	226.0
1990	1841	262.1
1991	1959	277.7
1992	2033	273.5

Щоб дослідити відношення між внутрішніми та світовими цінами, можна використати такі моделі:

- $y_t = \alpha_0 + \alpha_1 x_t + \varepsilon_t$;
- $y_t = \beta_1 x_t + \varepsilon_t$,

де $y = ВВП$ — дефлятор для вітчизняної продукції і $x = ВВП$ — дефлятор для імпортованої продукції.

- Яку з моделей ви виберете як найкращу?
- Застосуйте кожну з моделей до даних та визначте, яка з них краще підходить для опису даних.
- Яка інша модель може бути прийнятною для наведених даних?

Вправа 3.3. Регресія із стандартизованими змінними

Нехай $x_i^* = (x_i - \bar{x}) / S_x$ $y_i^* = (y_i - \bar{y}) / S_y$,

де S_x і S_y — середньоквадратичні відхилення x та y у вибірці. Покажіть, що для моделі

$$y_i^* = \beta_0 + \beta_1 x_i^* + \varepsilon_i$$

оцінка параметра β_0 дорівнює нулю ($b_0=0$), а оцінка параметра β_1 дорівнює коефіцієнту кореляції між x та y ($b_1=r$). Поясніть причину використання регресійної моделі із стандартизованими змінними.

Примітка. y_i^* та x_i^* , визначені вище, є стандартизованими змінними. Змінна є стандартизованою, або заданою в одиницях стандартного відхилення, якщо вона виражена в формі відхилення від її середнього значення (зміна від початкового значення) і поділена на її середньоквадратичне відхилення (зміна масштабу). Таким чином, стандартизація включає як зміну від початкового значення, так і зміну масштабу.

Стандартизовані змінні мають нульове середнє значення і дисперсію 1. Як результат, зміна на одиницю, скажімо x_i^* , означає зміну одного середньоквадратичного відхилення. Тоді нахил в моделі, поданій вище, може інтерпретуватися як такий, що показує кількість середньоквадратичних відхилень, на яку в середньому змінюється залежна змінна, коли пояснювана змінна змінюється на одне стандартне відхилення. Коефіцієнт нахилу у вищевказаній моделі називається бета-коефіцієнтом (не змішувати з бета-коефіцієнтом у портфельній теорії).

Вправа 3.4

Дано такі моделі:

$$\text{модель 1: } y_i = \beta_0 + \beta_1 x_i + \varepsilon_i.$$

$$\text{модель 2: } y_i^* = \alpha_0 + \alpha_1 x_i^* + \varepsilon_i,$$

де y^* і x^* — стандартизовані змінні, як було визначено у вправі 3.2. Як і раніше, вважатимемо, що оцінка параметра $\alpha_1 = a_1$, а оцінка параметра $\beta_1 = b_1$. Покажіть, що $a_1 = b_1(S_x / S_y)$.

Доведіть, що, хоча параметри регресії не залежать від зміни початкового значення, вони є залежними від зміни масштабу.

Вправа 3.5

Розглянемо такі моделі:

$$\ln y_i^* = \alpha_0 + \alpha_1 \ln x_i^* + \varepsilon_i^*;$$

$$\ln y_i = \beta_0 + \beta_1 \ln x_i + \varepsilon_i,$$

де $y_i^* = w_1 y_i$ та $x_i^* = w_2 x_i$, w — константою.

1. Встановіть відношення між двома наборами параметрів регресії та їхніми середньоквадратичними відхиленнями.
2. Чи є R^2 різним в обох моделях?

Вправа 3.6

Припустимо, ви застосовуєте таку версію кривої Філіпса до даних із таблиці 3.1:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t,$$

де y — річна зміна рівня заробітної плати (%) та x — рівень безробіття (%).

1. Яким попередньо є очікуваний знак β_1 ?
2. Оцініть невідомі параметри регресії, використовуючи дані вправ 3.1.
3. Як можна отримані результати порівняти з такими ж у моделі з вправ 3.1? Чи є між ними якась суперечність?
4. Чи можете ви порівняти два значення R^2 ?
5. Якій з моделей ви надаєте перевагу? Чому?
6. Чи можна оцінити природний рівень безробіття з оціненої кривої Філіпса у вправі 3.1? Якщо можна, то як?

Вправа 3.7

Дано таку модель:

$$y_t = y_0(1 + g)^t e^{\varepsilon_t},$$

де y_t — значення змінної y в період часу t ; y_0 — початкове значення y ; g — річний темп зростання; ε — випадкова величина; t — час.

Якщо не брати до уваги випадкову величину, то наведена модель є простою формулою складного процента у фінансах.

Логарифмуючи модель, отримаємо:

$$\ln y_t = \ln y_0 + t \ln(1 + g) + \varepsilon_t = \beta_0 + \beta_1 t + \varepsilon_t, \quad (3.23)$$

де $\beta_0 = \ln y_0$ і $\beta_1 = \ln(1 + g)$.

1. Яка відмінність між рівнянням (3.23) та log-lin-моделлю

$$\ln y_t = \beta_0 + \beta_1 x_t + \varepsilon_t, \quad (3.24)$$

хоч стосовно регресії (після того як допускається $x = t$) обидва рівняння виглядають однаково?

2. Якщо ви використовуєте модель (3.23), як би ви оцінили параметр g , темп зростання?

3. Як цей темп зростання відрізняється від отриманого у (3.24)? Чи є між ними якась відмінність, якщо так, то чи вона суттєва?

Вправа 3.8

Як ви обчислите еластичність за доходом $(dy/dx)(x/y)$ для log-lin-моделі (3.24), якщо y є величиною спожитих товарів і x доходом споживачів?

Вправа 3.9

Логарифмічна зворотна модель має вигляд:

$$\ln y_i = \beta_0 - \beta_1(1/x_i) + \varepsilon_i.$$

1. Які властивості цієї моделі? Підказка: використайте нахил, асимптоти і т.д.
2. Нехай x — час. Який вигляд має крива, що описується цією моделлю?
3. У якому випадку використовується ця модель?

Вправа 3.10

У табл. 3.4 подаються такі дані:

Таблиця 3.4

y_i	x_i	y_i	x_i
86	3	62	35
79	7	52	45
76	12	51	55
69	17	51	70
65	25	48	120

Побудуйте за наведеними даними модель вигляду:

$$(100 / 100 - y_i) = \beta_0 + \beta_1(1/x_i).$$

Оцініть невідомі параметри, проінтерпретуйте результати.

Вправа 3.11

Щоб виміряти еластичність між капіталом та витратами праці Ероу, Ченері, Мінас та Солоу, автори відомої тепер виробничої функції ПЕЗ (функція з постійною еластичністю заміни — CES, constant elasticity of substitution), використовували таку модель:

$$\log(V/L) = \log \beta_0 + \beta_1 \log W + \varepsilon,$$

де (V/L) — вартість, що додається кожною додатковою одиницею праці; L — кількість праці; W — реальна заробітна плата. Коефіцієнт β_1 показує еластичність заміни між працею та капіталом. За даними табл. 3.5, яка подана нижче, пересвідчіться, що еластичність дорівнює 1.3338 і що її значення не відрізняється статистично значимо від 1.

Таблиця 3.5

Промисловість	$\log(V/L)$	$\log W$
Борошняна	3.6973	2.9617
Цукрова	3.4795	2.8532
Фарби	4.0004	3.1158
Цементна	3.6609	3.0371
Скляних виробів	3.2321	2.8727
Керамічна	3.3418	2.9745
Лісова	3.4308	2.8287
Бавовняна	3.3158	3.0888
Вовняна	3.5062	3.0086
Льняна	3.2352	2.9680
Хімічна	3.8823	3.0909
Алюмінієва	3.7309	3.0881
Металургійна	3.7716	3.2256
Велосипедів	3.6601	3.1025
Ткацьких верстатів	3.7554	3.1354

Вправа 3.12

У табл. 3.6 наведені дані про норми інфляції в умовних країнах (%).

Таблиця 3.6

Рік	Країна 1	Країна 2	Країна 3
1982	5.9	6.5	7.6
1983	4.3	9.5	6.3
1984	3.6	6.8	4.9
1985	6.2	8.4	12.0
1986	10.9	16.0	24.6
1987	9.2	24.2	11.7
1988	5.8	16.5	9.3
1989	6.4	15.9	8.1
1990	7.6	8.3	3.8
1991	11.4	13.4	3.6
1992	13.6	18.0	8.0

Для кожної країни оцініть темп зростання інфляції, за такою моделлю:

$$\ln y_t = \beta_0 + \beta_1 t + \varepsilon_{it},$$

де y — норма інфляції.

Вправа 3.13

Будівельна компанія спеціалізується на будівництві нерухомості. У табл. 3.7 наведено дані за 8 років щодо об'єктів, будівництво яких розпочато.

1. Для кожної змінної визначити регресію за часом.
2. Визначити загальну площу (кількість будинків та їх загальну площу).

Чи можна використати лінійну регресію?

Таблиця 3.7

Рік	1	2	3	4	5	6	7	8	9
Число	2	3	4	3	6	5	6	3	?
Середня площа	3	4	7	10	11	12	15	18	?

3. Визначити прогноз загальної площі на 9-ий рік.

Вправа 3.14

У табл. 3.8 наведено дані про населення США протягом 1850 — 1990 рр. Використати різні криві зростання для прогнозу, оцінити одержані результати, використовуючи критерії MSE і MAPE.

Таблиця 3.8

1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950
23.2	31.4	39.8	50.2	62.9	76.0	92.0	105.7	122.8	131.7	151.1

Порівняйте результати прогнозу і реальні дані за 1960—1990 роки (табл. 3.9).

Таблиця 3.9

1960	1970	1980	1990
179	203	226	246

Вправа 3.15

Проаналізуємо гіпотетичний випуск продукції двох країн — України та Литви за 1986—1993 р. (табл. 3.10). Економічне зростання України здається скромнішим порівняно з експоненціальним економічним зростанням Литви.

Таблиця 3.10

Випуск продукції України та Литви (млрд. єку)

Рік	Україна			Литва		
	y_t	$1/y_t$	$\ln(y_t)$	y_t	$1/y_t$	$\ln(y_t)$
1986	31.3	0.0319	3.4434	230.9	0.00433	5.442
1987	36.7	0.0272	3.6028	280.4	0.00357	5.6362
1988	44.4	0.0225	3.7932	320.1	0.00312	5.7686
1989	49.8	0.0200	3.9080	339.1	0.00295	5.8263
1990	59.6	0.01678	4.0877	398.8	0.00251	5.9885
1991	68.0	0.0147	4.2195	453.3	0.00221	6.1166
1992	74.4	0.0134	4.3095	504.4	0.00198	6.2234
1993	79.0	0.01266	4.3694	557.6	0.00179	6.3236

Якщо побудувати графіки за звичайною і логарифмічною шкалами, то можна помітити, що ефект зростання виробництва України, замаскований ефектом масштабу, зникне.

1. Побудувати графіки за звичайною і логарифмічною шкалами. Порівняти результати.

2. Знайти параметри за методом трьох точок для модифікованої експоненти, логістичної кривої і кривої Гомперця (використати медіани і середні).

3. Знайти параметри за методом МНК для експоненти.

4. Зробити прогноз на 1994 р. та дати оцінку якості моделі за критерієм MAPE. Порівняти результати.

Вправа 3.16

У табл. 3.11 та 3.12 наведено дані щодо випуску промислової продукції та поквартальної заробітної плати в Україні.

Підібрати криві, які найповніше описують тенденцію.

Таблиця 3.11

Випуск промислової продукції (помісячно) в Україні в 1994 р.

Дата	Випуск продукції, млрд.крб
Січень 1994	46516
Лютий 1994	52928
Березень 1994	57928
Квітень 1994	58827
Травень 1994	59978
Червень 1994	65169
Липень 1994	64513

Таблиця 3.12

Поквартальна заробітна плата у держсекторі України (1992 — 1994 рр.)

Квартал	Заробітна плата без премій, млрд.крб.
1992 (1)	115
1992 (2)	238
1992 (3)	362
1992 (4)	761
1993 (1)	1223
1993 (2)	2220
1993 (3)	6366
1993 (4)	21224
1994 (1)	42043
1994 (2)	54313

ТЕСТИ

Вибрати правильну відповідь на запитання

1. Регресійна модель вважається лінійною, коли вона:

- а) лінійна за змінними;
- б) лінійна за параметрами;
- в) лінійна за змінними та параметрами.

2. Припустимо, що для опису одного економічного процесу придатні дві моделі. Обидві адекватні за F-критерієм Фішера. Якій надати перевагу, тій у якій:

- а) більший коефіцієнт детермінації;
- б) менший коефіцієнт детермінації;
- в) більше значення F-критерія Фішера;
- г) менше значення F-критерія Фішера?

3. Припустимо, що для опису одного економічного процесу прийнятні дві моделі. Обидві адекватні за F-критерієм Фішера. Якій надати перевагу, тій у якій:

- а) менше значення MAPE;
- б) більше значення MAPE;
- в) більше значення F-критерія Фішера;
- г) менше значення F-критерія Фішера?

4. Значення ME для лінійної регресії повинно прямувати:

- а) до 1;
- б) до нескінченності;
- в) до 0.

5. Крива витрат Енгеля показує відношення витрат споживача до його загального доходу. Позначаючи через y витрати споживача на товари, через x дохід споживача, отримуємо такі моделі:

- а) $y = \beta_0 + \beta_1 x + \varepsilon$;
- б) $y = \beta_0 + \beta_1 \frac{1}{x} + \varepsilon$;
- в) $y = \beta_0 + \beta_1 \ln x + \varepsilon$;
- г) $\ln y = \ln \beta_0 + \beta_1 \ln x + \varepsilon$.

Яку з цих моделей ви б вибрали для кривої Енгеля?

6. Крива Філіпса краще за все описується моделлю:

- а) $\ln y = \ln \beta_0 + \beta_1 \frac{1}{x} + \varepsilon$;
- б) $\ln y = \ln \beta_0 + \beta_1 \ln x + \varepsilon$;

$$в) y = \beta_0 + \beta_1 x + \varepsilon;$$

$$г) y = \beta_0 + \beta_1 \frac{1}{x} + \varepsilon.$$

7. Припустимо, що залежність витрат від доходу описується функцією:

$$y = b_0 + b_1 x.$$

Середнє значення $y = 2$, середнє значення $x = 6$, а $b_1 = 3$. Тоді коефіцієнт еластичності витрат від доходу дорівнює:

- 8;
- 1;
- 9;
- 4.

8. Припустимо, що залежність витрат від доходу описується функцією:

$$\ln y = b_0 + b_1 \ln x.$$

Середнє значення $y = 15$, середнє значення $x = 7$, а $b_1 = 4$. Тоді коефіцієнт еластичності витрат від доходу дорівнює:

- 38;
- 1/7;
- 9;
- 4.

ВІДПОВІСТІ НА ЗАПИТАННЯ: ТАК/НІ; ЯКЩО НІ, ПОЯСНІТЬ ЧОМУ

- Крива Гомперца є нелінійною за параметрами функцією.
- Логістична крива шляхом перетворень зводиться до лінійної регресії.
- Регресійна модель є нелінійною, якщо вона нелінійна за своїми змінними.
- Крива Філіпса має вигляд:

$$y = \beta_0 + \beta_1 \frac{1}{x} + \varepsilon.$$

- Життєвий цикл товару частіше за все описують s-подібними кривими.
- Чи можна шляхом перетворень звести експоненту до вигляду:

$$\ln y = \ln \alpha + x \ln \beta?$$

- Чи можна параметри модифікованої експоненти розрахувати за методом найменших квадратів?
- Чи можна параметри експоненти розрахувати за методом найменших квадратів?
- У log-linear-моделі коефіцієнт еластичності дорівнює нахилу.

10. У lin-log-моделі коефіцієнт еластичності дорівнює нахилу.

ВІДПОВІДІ

Тест

- б;
- а;
- а;
- в;
- б;
- г;
- в;
- г.

Так/ні

- Так.
- Ні. Логістична крива є нелінійною за параметрами моделлю.
- Ні. Регресійна модель є нелінійною, якщо вона нелінійна за своїми параметрами.
- Так.
- Так.
- Так.
- Ні. Модифікована експонента є нелінійною за своїми параметрами, і шляхом перетворень її не можна звести до лінійної регресії.
- Так.
- Так.
- Ні. В lin-log-моделі, яка має вигляд: $y = b_0 + b_1 \ln x$, коефіцієнт еластичності дорівнює: $b_1(1/y)$.

ЗАСТОСУВАННЯ ЕКОНОМІЧНИХ ЗНАТЬ

Зростання малих приватизованих підприємств в Україні

Розглянемо модель динаміки приватизації підприємств в Україні, зупинившись лише на найпростішій моделі зростання кількості приватизованих підприємств за часом. У реальному випадку при моделюванні процесу малої приватизації в Україні перш за все необхідно дослідити вплив багатьох факторів на процес малої приватизації, тобто побудувати та дослідити модель багатфакторної регресії (до неї ми повернемося в наступному розділі).

У табл. 3.16 наведено дані про щомісячну кількість приватизованих підприємств в Україні з січня 1994 по червень 1995 року.

Дані взято із збірників "Тенденції української економіки" Європейського Центру макроекономічного аналізу України та економічного бюлетеня "Україна в цифрах".

Якщо відобразити ці дані на числовій осі, то отримаємо приблизний графік, характерний для квадратичної функції (мал. 3.15).

Отже, для нашого прикладу можна припустити, що залежність між кількістю приватизованих малих підприємств та часом може бути описана квадратичною функцією.

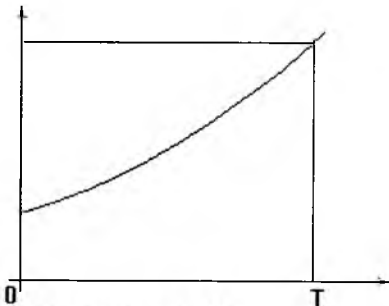
За наведеними даними, з використанням пакета STATGRAPH було отримано таку квадратичну модель:

$$y = 3262.05 + 913.819t - 17.5813t^2.$$

$$ME = 00.00, \\ MAPE = 2.58.$$

Зробимо прогноз зростання кількості приватизованих підприємств за наведеною моделлю, починаючи з жовтня 1995 року, та порівняємо його з уже відомими фактичними значеннями (табл. 3.17).

Як бачимо з таблиці, прогнозні дані виявилися значно заниженими порівняно з фактичними даними. Це можна легко пояснити тим, що у другій половині 1995 року уряд намагався адміністративними діями штурмувати процес приватизації, тому, звичайно, інерційність процесу була порушена.



Малюнок 3.10

Таблиця 3.16

Місяць, рік	Кількість приватизованих підприємств
15.01.94	4300
15.02.94	4967
15.03.94	5442
15.04.94	6438
15.05.94	7557
15.06.94	8402
15.07.94	8910
15.08.94	9450
15.09.94	10214
15.10.94	10910
15.11.94	11100
15.12.94	11552
15.01.95	12375
15.02.95	12450
15.03.95	12802
15.04.95	12975
15.05.95	13100
15.06.95	14957

Таблиця 3.17

Період	15.10.95	15.11.95	15.12.95
Результати, отримані за квадратичною моделлю	15138	15291	15406
Фактичні дані	18344	19026	19714

Прогнозування середнього темпу зростання обсягів продажу товару

На цьому прикладі ми покажемо, як на практиці можна ув'язати використання простої лінійної моделі та експоненційної. Припустимо, що відоме в Україні підприємство "Вимпел", що випускає автомобілі, хоче спрогнозувати обсяги реалізації продукції в майбутньому, а також знайти середній темп зростання обсягів продажу своєї продукції. Підприємство має статистику обсягу реалізації продукції з 1974 по 1996 рр. (умовні дані у табл. 3.18). Виходячи з цих даних, відділ маркетингу проводить відповідні дослідження.

Якщо відобразити ці дані на числовій осі, то отримаємо графік, характерний для лінійної функції. Отже, прогнозувати обсяги продажу продукції можна за моделлю простої лінійної регресії:

$$x_t = b_0 + b_1t + e_t. \quad (3.25)$$

Для того, щоб знайти середній темп зростання, згадаємо, що між обсягами продажу продукції у t -році та базовому є така залежність:

$$x_t = x_0 e^{gt}, \quad (3.26)$$

де x_0 — продаж продукції у базовому році; x_t — продаж продукції в t -році; g — темп зростання продажу продукції.

Використовуючи логарифмічне трансформування формули (3.26), отримаємо:

$$\log_e x_t = \log_e (x_0 e^{gt}) = \log_e x_0 + gt. \quad (3.27)$$

Рівняння (3.27) відоме як лог-лінійна (log-lin) регресійна модель. Справді, вводячи таку заміну змінних: $a_0 = \log_e x_0$, $a_1 = g = \text{норма зростання продажу}$, отримаємо:

$$\log_e x_t = a_0 + a_1t + e'_t. \quad (3.28)$$

Використовуючи моделі (3.25) та (3.28), відділ маркетингу розраховує прогноз майбутніх обсягів продажу, а також оцінює середній темп зростання обсягу продажу. Невідомі параметри моделі (3.25) обчислюються методом най-

Таблиця 3.18

Щорічні обсяги продажу автомобілів

Рік	t	Обсяги продажу, млн.грн. x_t
1973	1	14755.6
1974	2	14979.9
1975	3	16433.0
1976	4	20194.4
1978	5	23015.1
1979	6	23620.6
1980	7	24009.1
1981	8	28839.6
1982	9	37841.5
1983	10	42784.1
1984	11	43513.7
1985	12	37085.5
1986	13	38247.1
1987	14	37067.2
1988	15	44454.6
1989	16	52366.4
1990	17	52774.4
1991	18	62868.3
1992	19	72797.2
1993	20	82193.0
1994	21	82879.4
1995	22	81844.0

менших квадратів. Після розрахунків за допомогою пакета прикладних програм STATGRAPH отримаємо (див. табл. 3.19):

Таблиця 3.19

Оцінені параметри

Змінна	Оцінені параметри	Стандартна помилка	t для H_0	Prob> t
Перетин	3239.94	2948.62	1.099	0.2843
Нахил	3167.10	215.05	14.727	0.0001

Таблиця 3.20

Аналіз дисперсій

Джерело	DF	Сума квадратів	Середній квадрат	F-значення	Prob> t
Модель	1	101150900144	101150900144	216.892	0.001
Помилка	20	982834678	46801651		
Загальна	21	11133734822			

$$\begin{aligned} \text{Root MSE} &= 6841.17; \\ R^2 &= 0.9075. \end{aligned}$$

Оцінений коефіцієнт детермінації (\bar{R}^2) = 0.9075.

Отже, модель для прогнозування обсягів продажу продукції з оціненими параметрами має вигляд:

$$\hat{x}_t = b_0 + b_1 t = 3239.94 + 3167.101t \quad (3.29)$$

з $R^2 = 0.9075$.

Підставляючи замість t необхідний період, наприклад (23, 24, ...), отримаємо відповідно прогнози обсягів продажу на 1996, 1997 рік і т.д.

Для того, щоб оцінити середній темп зростання обсягів продажу підприємства за 1973 — 1995 рр., використаємо log-lin-модель (3.28) та дані таблиці. Після розрахунку невідомих параметрів за методом найменших квадратів отримаємо:

$$\log_e \hat{x}_t = a_0 + a_1 t = 9.4834 + 0.0827 t \quad (3.30)$$

$$(0.516) (0.0038)$$

$$R^2 = 0.958.$$

Зазначимо, що величини в дужках — це стандартні помилки параметрів, тобто середнє квадратичне відхилення параметрів.

Результат показує, що середній темп зростання обсягів продажу для досліджуваного підприємства дорівнює $a_1 = g = 8.27\%$. Інакше кажучи, середній темп зростання обсягу продажу автомобілів підприємства “Вимпел” за 1973 — 1995 рр. становив 8.27%.

РОЗДІЛ 4. БАГАТОФАКТОРНА РЕГРЕСІЯ

4.1. ВСТУП. ПРИКЛАДИ ВИКОРИСТАННЯ БАГАТОФАКТОРНОГО РЕГРЕСІЙНОГО АНАЛІЗУ НА ПРАКТИЦІ

На практиці економічний процес змінюється під впливом багатьох різноманітних факторів, які треба вміти виявити та оцінити. Якщо ми повернемося до нашого прикладу аналізу обсягу продаж на фірмі, то справді, було б великим спрощенням допускати, що зміни їх залежать тільки від витрат на рекламу. На обсяги продажу впливає частина ринку, яку утримує фірма; якість продукції; імідж марки продукції серед населення; середня заробітна плата населення у регіонах продажу та інші фактори.

До складу доходу консолідованого бюджету України, входять прямі податки (на доходи) підприємств та домашніх господарств і непрямі податки (ПДВ — податок на додану вартість, акцизні збори). Доход також може залучати стягнення із зарплати (Чорнобильський податок, пенсійний фонд) та нефіскальні стягнення (наприклад, доходи від приватизації) і т.ін. Отже, при аналізі та прогнозуванні доходу консолідованого бюджету України необхідно дослідити вплив на його величину податків на додану вартість, податків на доходи підприємств, податків з населення, акцизного збору, Чорнобильського, пенсійного фонду та інші, тобто здійснити багатофакторний аналіз.

При дослідженні процесу малої приватизації великим спрощенням було б припускати (див. розділ 3), що кількість приватизованих підприємств змінюється тільки з часом. Звичайно вона змінюється під впливом багатьох факторів, а саме: обсягів кредитів, індексу реальної продукції, рівня інфляції оптових цін, кількості громадян, які скористалися приватизаційними сертифікатами, та ін.

Проблема лібералізації цін залишається досить актуальною для України. На відміну від стабілізації валюти та приватизації, необхідність яких була загальноновизнаю, лібералізація цін довго залишалася неузгодженим питанням для прихильників різних поглядів на розвиток економіки України. Для дослідження процесу лібералізації цін на ринках України також важливо виявити та дослідити вплив різноманітних факторів, таких, наприклад, як відхилення у відсотках ринкового обмінного курсу валют від встановленого державою (можна вважати цифру відхилення аналогією до відповідного “податку” на операції з обміну валют), від обмеження експорту (цей показник розраховується як частка експорту, що здійснюється за ліцензіями та встановленими державою квотами), від адмі-

ністративного контролю за внутрішньою торгівлею (частка загального випуску продукції, що здійснюється за державними замовленнями і контрактами, а не за рішенням підприємств), від контролю за цінами роздрібною торгівлі (частка загального випуску продукції, що підлягає державному управлінню), від контролю за оптовими цінами, від частки валютних операцій, що здійснюються через неринкові, адміністративні механізми, та ін.

За допомогою багатофакторної регресії можна аналізувати численні соціальні проблеми. Наприклад, досліджувати, як впливають різні соціальні фактори на кількість зареєстрованих в Україні злочинів протягом останнього десятиліття. Для аналізу та побудови багатофакторної регресійної моделі можна обрати такі фактори, як загальна кількість населення в Україні, забезпеченість населення житлом, продаж алкогольних напоїв у розрахунку на душу населення, кількість людей, що мають вищу та середню освіту в розрахунку на 1000 осіб, реальну середню заробітну плату робітників та службовців, рівень безробіття, рівень культурного розвитку нації та інші.

Досить актуальною в Україні залишається проблема аналізу народжуваності, а саме впливу різноманітних факторів на кількість народжених. Серед таких факторів виділимо кількість жіночого населення віком від 16 до 47 років, частку пенсіонерів у загальній кількості населення України, реальні грошові доходи населення, кількість постійних дошкільних закладів, кількість зареєстрованих шлюбів, кількість зареєстрованих розлучень, кількість абортів, кількість хворих з діагнозом на алкоголізм та наркоманію тощо.

Підсумовуючи розглянуті вище приклади, зазначимо, що саме багатофакторний регресійний аналіз допомагає знайти явний вигляд залежності досліджуваного показника від численних факторів, що впливають на його зміну, а також кількісно оцінити їхній вплив.

Але треба підкреслити, що складність розрахунків та узагальнення інформації призводять до необхідності широкого використання обчислювальної техніки. Тому побудова та аналіз багатофакторних регресійних моделей базуються на сучасних пакетах прикладних програм. Економіст-статистик повинен уміти аналізувати отримані результати та робити за ними висновки, вміти оцінити найкращу модель для взаємозв'язку вихідних статистичних даних.

4.2. КЛАСИЧНА ЛІНІЙНА БАГАТОФАКТОРНА МОДЕЛЬ. ОСНОВНІ ПРИПУЩЕННЯ У БАГАТОФАКТОРНОМУ РЕГРЕСІЙНОМУ АНАЛІЗІ

Узагальнена багатофакторна лінійна регресійна модель може бути записана у такому вигляді:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon, \quad (4.1)$$

де y — залежна змінна;

x_1, x_2, \dots, x_p — незалежні змінні (або фактори);

$\beta_0, \beta_1, \dots, \beta_p$ — параметри моделі (константи), які потрібно оцінити;

ε — неспостережувана випадкова величина.

Нагадаємо, що узагальнена регресійна модель — це модель, яка дійсна для всієї генеральної сукупності. Невідомі параметри узагальненої моделі є константами, а випадкова величина — неспостережувана, і ми можемо зробити тільки припущення відповідно до закону її розподілу. На відміну від узагальненої регресійної моделі, вибіркова модель будується для певної вибірки; невідомі параметри вибіркової моделі є випадковими величинами, математичне сподівання яких дорівнює параметрам узагальненої моделі (випадок класичної лінійної регресії), випадкові величини (помилки) можна оцінити, виходячи з вибірових даних.

Відповідно до позначень, введених у розділі 2, **вибіркова лінійна багатофакторна модель** має такий вигляд:

$$y = b_0 + b_1 x_1 + \dots + b_p x_p + e,$$

де y — залежна змінна;

x_1, x_2, \dots, x_p — незалежні змінні (або фактори);

b_0, b_1, \dots, b_p — оцінки невідомих параметрів узагальненої моделі;

e — випадкова величина (помилка).

Лінійною регресійною моделлю називається модель, лінійна за своїми параметрами.

За введеними нами позначеннями, багатофакторна лінійна регресійна модель має p незалежних змінних, або факторів, які впливають на залежну змінну y , та $(p + 1)$ невідомих параметрів, які потрібно оцінити.

Основні припущення у багатофакторному регресійному аналізі

Як було вже сказано, у разі узагальненої регресійної моделі, тобто моделі, дійсної для всієї генеральної сукупності, випадкова величина ε є неспостережуваною величиною, і ми можемо зробити лише деякі припущення щодо її поведінки та закону розподілу. Для класичної багатофакторної регресійної моделі, яка є узагальненням простої лінійної регресійної

моделі, всі основні класичні припущення зберігаються, але дещо модифікуються. Розглянемо ці припущення.

Припущення 1

Математичне сподівання випадкової величини ε дорівнює 0.

$$E(\varepsilon_i | x_{1i}, x_{2i}, \dots, x_{pi}) = 0 \text{ для кожного } i.$$

Припущення 2

Випадкові величини незалежні між собою.

$$\text{cov}(\varepsilon_i, \varepsilon_j) = 0;$$

$$i \neq j.$$

Припущення 3

Модель гомоскедастична, тобто має однакову дисперсію для будь-якого спостереження:

$$\text{var}(\varepsilon_i) = \sigma^2.$$

Припущення 4

Коваріація між випадковою величиною ε_i та кожною незалежною змінною x дорівнює 0.

Зазначимо, що властивість 4 виконується автоматично, якщо $x_i (i = \overline{1, p})$ не стохастичні та припущення 1 має силу.

Припущення 5

Модель повинна бути правильно специфікованою.

Припущення 6

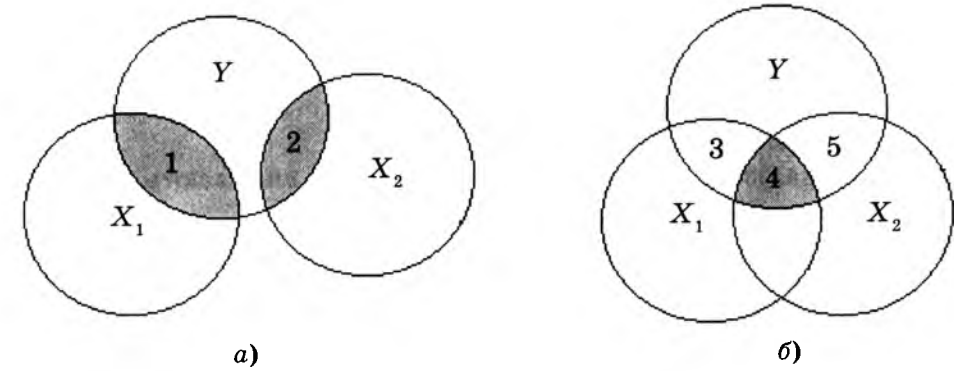
Випадкова величина ε відповідає нормальному закону розподілу з нульовим математичним сподіванням і постійною дисперсією.

Припущення 7

Відсутність мультиколінеарності між факторами x , тобто фактори повинні бути незалежними між собою. Іншими словами, не повинно бути точного лінійного зв'язку між двома або більше факторами.

Слід зазначити, що припущення 7 не прийнятне для простої лінійної регресії, але воно надзвичайно важливе для багатофакторної регресії. Зважаючи на це, розглянемо його детальніше.

Припустимо, що є лінійна залежність між факторами x_1 та x_2 . В такому випадку неможливо точно визначити окремий вплив кожного з цих факторів на залежну змінну y . Графічно це можна подати, виходячи з кругової діаграми (мал. 4.1).



Малюнок 4.1. Зв'язок між факторами: а — відсутність залежності між факторами x_1 та x_2 ; б — наявність такої залежності.

На малюнку 4.1 зображено два випадки. У випадку а відсутня залежність або колінеарність між x_1 та x_2 ; у випадку б — вона наявна. У випадку а підмножина 1 описує окремий вплив фактора x_1 на залежну змінну y , а підмножина 2 — окремий вплив фактора x_2 . При наявності колінеарності (випадок б) підмножина 3 описує окремий вплив фактора x_1 , а підмножина 5 — окремий вплив фактора x_2 . Підмножина 4 характеризує спільний вплив обох факторів на змінну y , який не можна відокремити. Саме підмножина 4 графічно описує ситуацію колінеарності. Коли ця підмножина дорівнює 0, то колінеарності немає, що ми і бачимо на мал. 4.1 у випадку а.

Математично відсутність колінеарності між двома факторами, наприклад факторами x_1 та x_2 , визначається таким чином, що не існує чисел γ_1 та γ_2 , які одночасно не дорівнюють 0, для яких би виконувалась тотожність

$$\gamma_1 x_{1i} + \gamma_2 x_{2i} = 0, i = \overline{1, n}. \quad (4.2)$$

Іншими словами, якщо тотожність (4.2) виконується тільки тоді, коли $\gamma_1 = \gamma_2 = 0$, то x_1 та x_2 лінійно незалежні, або неколінеарні. У протилежному разі має місце колінеарність.

Інтуїтивно встановлюємо, що коли між двома змінними, наприклад x_1 та x_2 , є лінійний зв'язок, то йдеться не про дві, а одну незалежну змінну, бо неможливо знайти окремий вплив кожної з цих змінних на y .

Покажемо це. Нехай залежність між x_1 та x_2 має вигляд: $x_2 = 3x_1$. Тоді вираз (4.1) можна переписати у вигляді:

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_{1i} + \beta_2 3x_{1i} + \beta_3 x_{3i} + \dots + \beta_p x_{pi} + \varepsilon_i = \\ &= \beta_0 + x_{1i} (\beta_1 + 3\beta_2) + \beta_3 x_{3i} + \beta_p x_{pi} + \varepsilon_i x_{3i} + \dots + \beta_p x_{pi} + \varepsilon_i. \end{aligned} \quad (4.3)$$

Як бачимо з (4.3), кількість змінних зменшилась на 1, а параметр $(\beta_1 + 3\beta_2)$ є оцінкою спільного впливу x_1 та x_2 на y , який не можна розділити.

4.3. ЕТАПИ ПОБУДОВИ БАГАТОФАКТОРНОЇ РЕГРЕСІЙНОЇ МОДЕЛІ

Процес побудови багатофакторної регресійної моделі більш складний, ніж процес побудови простої лінійної регресії. Він складається з багатьох етапів. Серед них можна виділити такі.

1. Вибір та аналіз усіх можливих факторів, які впливають на процес (або показник), що вивчається.
2. Вимір та аналіз знайдених факторів.
3. Математико-статистичний аналіз факторів.
4. Вибір методу та побудова регресійної багатофакторної моделі.
5. Оцінка невідомих параметрів регресійної моделі.
6. Перевірка моделі на адекватність.
7. Розрахунок основних характеристик та побудова інтервалів довіри.
8. Аналіз отриманих результатів, висновки.

Розглянемо детально кожний з етапів побудови та аналізу багатофакторної регресійної моделі.

Перший етап складається з вибору всіх можливих факторів, які впливають на процес або показник, що вивчається. На цьому етапі дослідник повинен глибоко зрозуміти сам економічний процес, розглянути його з макроекономічних та мікроекономічних позицій; виявити якомога більше факторів, які в конкретному випадку можуть справити суттєвий або несуттєвий вплив на його зміну. На цьому етапі можуть знадобитися поради практиків, які працюють у галузі або на фірмі, що вивчається, і т.ін. Після того, як множина всіх факторів окреслена, переходять до **другого етапу** — кількісного аналізу відібраних факторів.

На етапі кількісного аналізу дослідник повинен оцінити можливість кількісного вираження відібраних факторів, провести вимірювання або зібрати статистику для кількісних факторів; підібрати або розробити балоу шкалу оцінок для якісних даних. Якщо деякі фактори неможливо кількісно виразити, наприклад імідж продукції у населення, їх треба вилучити з подальшого розгляду. З подальшого розгляду вилучаються також фактори, за якими немає або недоступна статистика.

Після того, як усі фактори проаналізовано, подано у кількісному вигляді, тобто у вигляді динамічних або варіаційних рядів, переходять до **третього етапу** — етапу математико-статистичного аналізу.

Етап математико-статистичного аналізу є найважливішим підготовчим етапом для побудови регресійної багатофакторної моделі. Це заключний етап формування необхідної інформаційної бази.

При наявності у динамічних рядах недостатньої інформації саме на цьому етапі за допомогою спеціальних методів проводиться її відтворення. На цьому етапі проводиться перевірка основних припущень класичного регресійного аналізу, крім того, здійснюється найважливіша процедура багатофакторного аналізу — перевірка факторів на мультиколіне-

арність. Для цього спочатку будується матриця коефіцієнтів парної кореляції, яка є симетричною і має такий вигляд:

$$R = \begin{pmatrix} & y & x_1 & x_2 & \dots & x_k \\ y & r_{y^2} & r_{yx_1} & r_{yx_2} & \dots & r_{yx_k} \\ x_1 & r_{yx_1} & r_{x_1^2} & r_{x_1x_2} & \dots & r_{x_1x_k} \\ x_2 & r_{yx_2} & r_{x_1x_2} & r_{x_2^2} & \dots & r_{x_2x_k} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_k & r_{yx_k} & r_{x_1x_k} & \dots & \dots & r_{x_k^2} \end{pmatrix},$$

де R — матриця кореляції;

$r_{x_i x_j} = r_{x_j x_i}$; $i = j = 1, k$ — коефіцієнт парної кореляції між i -м та j -м факторами;

r_{yx_j} — коефіцієнт кореляції між залежною змінною y та j -м фактором.

Потім аналізуються коефіцієнти парної кореляції між факторами. Якщо значення деяких з них близьке до 1, це вказує на щільний зв'язок між ними, або на мультиколінеарність. Тоді один з факторів необхідно залишити, а інший вилучити із подальшого розгляду. Постає питання: який саме? Це залежить від конкретної ситуації. Найчастіше залишають той фактор, який з економічної точки зору більш вагомий для аналізу впливу на залежну змінну. Можна також залишити фактор, який має більший коефіцієнт кореляції із залежною змінною y . Такий аналіз проводиться для кожної пари залежних між собою факторів. Результатом етапу математико-статистичного аналізу є знаходження множини основних незалежних між собою факторів, які є базою для побудови регресійної моделі.

Метод побудови регресійної багатофакторної моделі неможливо відокремити від самої моделі, вони найтіснішим чином пов'язані між собою. Іншими словами, саме обраний метод впливає на остаточний вигляд регресійної моделі. Це ми розглядатимемо дещо пізніше.

Оцінка невідомих параметрів $\beta_0, \beta_1, \dots, \beta_p$ здійснюється у лінійних регресійних моделях за методом найменших квадратів, уже відомим із розділу простої лінійної регресії.

Після того, як параметри знайдено, проводиться перевірка моделей на адекватність за допомогою F -критерію Фішера, а також перевірка значимості знайдених параметрів за t -критерієм Ст'юдента. Якщо модель неадекватна, то необхідно повернутися до етапу побудови моделі і, можливо, від лінійної моделі перейти до нелінійної, або ввести додаткові фактори.

Якщо модель адекватна, то можемо працювати далі: робити прогнозування, вивчати вплив окремих факторів на залежний показник, будувати

інтервали довіри, аналізувати та інтерпретувати отримані результати. Для того, щоб розглянути, як можна проінтерпретувати параметри регресійної моделі, повернемося до загальної моделі багатофакторного регресійного аналізу (4.1) та знайдемо математичне очікування обох частин. Виходячи з основних припущень, отримаємо:

$$E(y_i | x_{1i}, x_{2i}, \dots, x_{pi}) = \beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi}. \quad (4.4)$$

Рівняння (4.4) дає умовне математичне сподівання y при фіксованих значеннях x .

Параметри $\beta_0, \beta_1, \dots, \beta_p$ ще називають частковими коефіцієнтами регресії. Кожний з них вимірює вплив відповідної змінної за умови, що всі інші залишаються без змін, тобто є константами.

4.4. РОЗРАХУНОК НЕВІДОМИХ ПАРАМЕТРІВ БАГАТОФАКТОРНОЇ РЕГРЕСІЇ ЗА МЕТОДОМ НАЙМЕНШИХ КВАДРАТІВ (МНК)

Нехай маємо ряд спостережень за залежною змінною $y = \{y_1, y_2, \dots, y_n\}$ та за незалежними змінними, або факторами:

$$x_1 = \{x_{11}, x_{12}, \dots, x_{1n}\}; x_2 = \{x_{21}, x_{22}, \dots, x_{2n}\}, \dots, \\ x_p = \{x_{p1}, x_{p2}, \dots, x_{pn}\}.$$

На підставі цих спостережень побудуємо лінійну вибірку багатофакторну модель, а саме:

$$y = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_p x_p + e, \quad (4.5)$$

де y — залежна змінна; x_1, \dots, x_p — незалежні змінні, або фактори;

b_0, b_1, \dots, b_p — невідомі параметри; e — випадкова величина, або помилка.

Як і у випадку простої лінійної регресії, знайдемо невідомі параметри за методом найменших квадратів, тобто мінімізуючи суму квадратів відхилень фактичних даних від теоретичних (даних, які ми отримуємо з регресійної моделі):

$$F(b_0, b_1, \dots, b_p) = \min \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_{1i} - \dots - b_p x_{pi})^2. \quad (4.6)$$

Для того, щоб знайти мінімум виразу (4.6), необхідно прив'язати до нуля часткові похідні функції F за аргументами b_0, b_1, \dots, b_p . Отримаємо систему нормальних рівнянь. Зважаючи на досить громіздкий вигляд системи нормальних рівнянь у загальному випадку, ми не будемо її наводити. Загальний вираз для розрахунку невідомих параметрів моделі роз-

глянемо пізніше, коли повернемося до матричного підходу в багатофакторному аналізі. Зазначимо тільки, що перетин (параметр b_0) розраховується аналогічно до простої регресії за допомогою середніх значень:

$$b_0 = \bar{y} - b_1 \bar{x}_1 - \dots - b_p \bar{x}_p. \quad (4.7)$$

4.5. ВЛАСТИВОСТІ МЕТОДУ НАЙМЕНШИХ КВАДРАТІВ

Властивості методу найменших квадратів у випадку багатофакторної регресії збігаються з його властивостями у випадку простої лінійної регресії:

Властивість 1. Багатофакторна регресійна модель правильна для середніх точок $\bar{y}; \bar{x}_1; \bar{x}_2; \dots; \bar{x}_p$.

Тобто, для моделі

$$y_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + \dots + b_p x_{pi} + e_i$$

$$\text{маємо: } b_0 = \bar{y} - b_1 \bar{x}_1 - \dots - b_p \bar{x}_p. \quad (4.8)$$

Властивість 2. Середнє значення оцінки дорівнює середньому значенню фактичних даних, тобто $\hat{y} = y$.

Це легко показати:

$$\begin{aligned} \hat{y}_i &= b_0 + b_1 x_{1i} + b_2 x_{2i} + \dots + b_p x_{pi} = \\ &= (\bar{y} - b_1 \bar{x}_1 - \dots - b_p \bar{x}_p) + b_1 x_{1i} + b_2 x_{2i} + \dots + b_p x_{pi} = \\ &= \bar{y} + b_1 (x_{1i} - \bar{x}_1) + \dots + b_p (x_{pi} - \bar{x}_p). \end{aligned} \quad (4.9)$$

Просумуємо обидві частини (4.9) за i ($i = \overline{1, n}$) і, виходячи з того, що

$$\sum_{i=1}^n (x_{ji} - \bar{x}_j) = 0; \text{ для } j = \overline{1, p}, \text{ отримаємо } \bar{\hat{y}} = \bar{y}.$$

Для пояснення решти властивостей введемо позначення.

Позначимо $(x_{ji} - \bar{x}_j) = \tilde{x}_{ji}$, тоді рівність (4.9) можна переписати:

$$\hat{y}_i = b_1 \tilde{x}_{1i} + b_2 \tilde{x}_{2i} + \dots + b_p \tilde{x}_{pi}, \quad (4.10)$$

де $\hat{y}_i = (\hat{y}_i - \bar{y})$, $\tilde{y}_i = (y_i - \bar{y})$.

На основі (4.10) багатофакторну вибірку модель (4.5) можна записати у формі:

$$\tilde{y}_i = \hat{y}_i + e_i. \quad (4.11)$$

Властивість 3. Сума помилок дорівнює нулю. $\sum_{i=1}^n e_i = \bar{e} = 0$. (Це випливає з (4.11)).

Властивість 4. Помилки e_i некорельовані з $x_{1i}, x_{2i}, \dots, x_{pi}$, тобто

$$\sum_{i=1}^n e_i x_{1i} = \sum_{i=1}^n e_i x_{2i} = \dots = \sum_{i=1}^n e_i x_{pi} = 0.$$

Властивість 5. Помилки e_i некорельовані з \hat{y}_i , тобто $\sum_{i=1}^n e_i \hat{y}_i = 0$.

Властивість 6. Якщо правильні припущення класичної лінійної регресійної моделі, то МНК-оцінки є не тільки лінійними, без відхилень оцінками, а й мають найменшу дисперсію, тобто є BLU-оцінками.

4.6. КОЕФІЦІЄНТ МНОЖИНОЇ КОРЕЛЯЦІЇ ТА ДЕТЕРМІНАЦІЇ

Корисною мірою ступеня відповідності даних $\{\hat{y}_i, i = \overline{1, n}\}$, отриманих з регресійної моделі, фактичним даним $\{y_i, i = \overline{1, n}\}$ є коефіцієнт множинної кореляції, який визначається як коефіцієнт кореляції між y та \hat{y} і має вигляд:

$$R = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}}. \quad (4.12)$$

Квадрат коефіцієнта множинної кореляції, як і у випадку простої регресії, називають коефіцієнтом детермінації і позначають через R^2 . Можна показати, що вигляд коефіцієнта детермінації у випадку багатофакторної регресії ідентичний коефіцієнту детермінації простої лінійної регресії:

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}. \quad (4.13)$$

Розглянемо вибірку багатофакторну модель

$$y_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + \dots + b_p x_{pi} + e_i = \hat{y}_i + e_i, \quad (4.14)$$

де \hat{y}_i — оцінка фактичного значення, або прогнозне чи теоретичне значення.

Замінюючи b_0 на його вираз через середні значення

$$b_0 = \bar{y}_i - b_1 \bar{x}_1 - b_2 \bar{x}_2 - \dots - b_p \bar{x}_p$$

і підставляючи даний вираз у (4.14), отримаємо:

$$(y_i - \bar{y}) = b_1(x_{1i} - \bar{x}) + b_2(x_{2i} - \bar{x}) + \dots + b_p(x_{pi} - \bar{x}) + e_i. \quad (4.15)$$

Замінімо $(y_i - \bar{y})$ на \tilde{y}_i , а $(x_{ji} - \bar{x})$ на \tilde{x}_{ji} , тоді (4.15) можна переписати у вигляді:

$$\tilde{y}_i = b_1 \tilde{x}_{1i} + b_2 \tilde{x}_{2i} + \dots + b_p \tilde{x}_{pi} + e_i = \tilde{\hat{y}}_i + e_i. \quad (4.16)$$

Піднесемо обидві частини (4.16) до квадрата і просумуємо за всіма значеннями. Отримаємо:

$$\sum_{i=1}^n \tilde{y}_i^2 = \sum_{i=1}^n \tilde{\hat{y}}_i^2 + \sum_{i=1}^n e_i^2 + 2 \sum_{i=1}^n \tilde{\hat{y}}_i e_i = \sum_{i=1}^n \tilde{\hat{y}}_i^2 + \sum_{i=1}^n e_i^2. \quad (4.17)$$

Виходячи з позначень, які ми раніше вводили для простої лінійної регресії, отримаємо:

$$SST = SSR + SSE, \quad (4.18)$$

де $SST = \sum_{i=1}^n (y_i - \bar{y})^2$; $SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$; $SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$.

Тепер легко можна побачити, що, як і у випадку простої лінійної регресії, коефіцієнт детермінації дорівнює:

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}. \quad (4.19)$$

4.7. КОЕФІЦІЄНТ ДЕТЕРМІНАЦІЇ R^2 ТА ОЦІНЕНИЙ КОЕФІЦІЄНТ ДЕТЕРМІНАЦІЇ \bar{R}^2

Важливою властивістю коефіцієнта детермінації R^2 є те, що він — неспадна функція від кількості факторів, які входять до моделі. Якщо кількість факторів зростає, то R^2 також зростає і ніколи не зменшується. Тобто, якщо ми додаємо новий фактор у регресійну модель, це тільки збільшує значення коефіцієнта детермінації R^2 , що легко побачити з його визначення:

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} = 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}. \quad (4.20)$$

У виразі (4.20) знаменник не залежить від кількості факторів x , тоді як чисельник, навпаки, залежить. Інтуїтивно можна зрозуміти, що якщо кількість факторів x зростає, величина $\sum_{i=1}^n e_i^2$ спадає (або хоча б не зростає).

Якщо ми порівнюватимемо дві регресійні моделі з однаковою залежною змінною, але різною кількістю факторів x , то, звичайно, віддамо перевагу тій, яка має більше значення R^2 .

Зразу ж постає запитання, що робити, якщо ми хочемо порівняти значення коефіцієнтів детермінації в різних моделях. У таких випадках потрібно коригувати коефіцієнт кореляції з урахуванням кількості факторів x , які входять у різні моделі, тобто зменшити вплив залежності значення коефіцієнта детермінації від кількості факторів. Для цього вводиться спеціальний оцінений коефіцієнт детермінації, який має вигляд:

$$\bar{R}^2 = 1 - \frac{\sum_{i=1}^n e_i^2 / (n-k)}{\sum_{i=1}^n (y_i - \bar{y})^2 / (n-1)}, \quad (4.21)$$

де k — кількість параметрів регресійної моделі, включаючи перетин.

На відміну від простого коефіцієнта детермінації, оцінений коефіцієнт детермінації коригується з урахуванням ступенів вільності суми квадратів залишків та загальної суми квадратів. Як бачимо у виразі (4.21), суми квадратів у чисельнику та знаменнику діляться на відповідні ступені вільності, в яких ураховується кількість факторів, що входять до моделі.

Вираз (4.21) можна записати ще таким чином:

$$\bar{R}^2 = 1 - \left(\frac{\hat{\sigma}_\varepsilon^2}{S_y^2} \right), \quad (4.22)$$

де $\hat{\sigma}_\varepsilon^2$ — оцінена дисперсія залишків; S_y^2 — вибіркова дисперсія незалежної змінної y .

Легко помітити, що оцінений коефіцієнт детермінації \bar{R}^2 та коефіцієнт детермінації R^2 пов'язані між собою такою залежністю:

$$\bar{R}^2 = 1 - (1 - R^2) \frac{n-1}{n-k}. \quad (4.23)$$

З виразу (4.23) видно, що якщо $k > 1$, то $\bar{R}^2 < R^2$. Крім того, якщо кількість факторів x зростає, оцінений коефіцієнт детермінації зростає повільніше, ніж просто коефіцієнт детермінації. Таким чином, зменшується вплив кількості факторів на величину коефіцієнта детермінації, тому на практиці більше використовують оцінений коефіцієнт детермінації, особливо при порівнянні різних регресійних моделей. Слід зазначити, що оцінений коефіцієнт детермінації може бути і негативним, на відміну від R^2 , який має завжди позитивне значення. Крім того, коли $R^2 = 1$, оцінений

коефіцієнт кореляції також дорівнює одиниці. Коли \bar{R}^2 прямує до негативної величини, R^2 прямує до нуля.

Для того, щоб розібратись, як на практиці порівнюють значення коефіцієнтів детермінації в різних моделях, розглянемо приклад. Ще раз нагадаємо, що порівняти значення двох або більше коефіцієнтів детермінації (оцінених або ні) можна лише за однакових залежних змінних, які можуть набирати різних функціональних форм.

Наприклад, нехай ми маємо такі дві моделі:

$$\ln y_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + \dots + b_p x_{pi} + e_i, \quad (4.24)$$

$$y_i = a_0 + a_1 x_{1i} + a_2 x_{2i} + \dots + a_p x_{pi} + e_i. \quad (4.25)$$

Розраховані коефіцієнти детермінації R^2 в цих моделях не можна порівняти між собою. Пояснимо чому. За означенням, коефіцієнт детермінації є частиною дисперсії, що пояснює регресію в загальній дисперсії (дисперсії залежної змінної). Таким чином, коефіцієнт детермінації моделі (4.24) вимірює частку дисперсії $\ln y$, яку можна пояснити факторами $x_{1i}, x_{2i}, \dots, x_{pi}$, тоді як у моделі (4.25) він вимірює частку дисперсії y . Але це не одне й те саме. Зміна в $\ln y$ забезпечує відносну зміну y , тоді як зміна в самому y є абсолютною зміною. Таким чином, величина $\text{var}(\hat{y}_i) / \text{var}(y_i)$ не буде дорівнювати величині $\text{var}(\ln \hat{y}_i) / \text{var}(\ln y_i)$. Справді, відповідні коефіцієнти детермінації дорівнюватимуть для моделі (4.25):

$$1 - R^2 = \frac{SSE}{SST} = \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (4.26)$$

а для моделі (4.24):

$$1 - R^2 = \frac{SSE}{SST} = \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (\ln y_i - \ln \bar{y})^2}. \quad (4.27)$$

Знаменники y (4.26) та (4.27) різні, отже, коефіцієнти детермінації порівняти не можна. Як бути у такому разі? Для того, щоб порівняти коефіцієнти детермінації регресійних моделей (4.24) та (4.25), необхідно, по-перше, знайти $\ln \hat{y}_i$ за моделлю (4.24), обчислити антилогарифм; по-друге, розрахувати коефіцієнт детермінації між антилогарифмом та значенням y_i .

Знайдений таким чином коефіцієнт детермінації можна порівнювати з коефіцієнтом детермінації R^2 регресійної моделі (4.25).

Можна піти і зворотним шляхом: обчислити за моделлю (4.24) $\ln \hat{y}_i$, розрахувати \hat{y}_i , обчислити R^2 за формулою (4.13), а його в свою чергу можна порівнювати з R^2 , знайденим для моделі (4.25).

4.8. ANOVA-ДИСПЕРСІЙНИЙ АНАЛІЗ

Елементарна ANOVA-таблиця у випадку багатофакторної регресії має такий вигляд (табл. 4.1).

Таблиця 4.1

Джерело варіації	Суми квадратів	Ступені вільності	Середні квадрати
Модель	$SSR = \sum_{i=1}^n (\hat{y}_i - y)^2$	$k - 1$	$MSR = SSR / (k - 1)$
Помилка	$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$	$n - k$	$MSE = SSE / (n - k)$
Загальне	$SST = \sum_{i=1}^n (y_i - \bar{y})^2$	$n - 1$	

де k — кількість параметрів регресійної моделі, включаючи перетин.

З ANOVA-таблиці можна легко отримати вираз як для простого коефіцієнта детермінації (R^2), так і для оціненого.

Для цього спочатку ще раз пригадаємо формулу розподілу сум квадратів та вираз для коефіцієнта детермінації:

$$SST = SSR + SSE; \quad (4.28)$$

$$1 = \frac{SSR}{SST} + \frac{SSE}{SST} = R^2 + \frac{SSE}{SST}; \quad (4.29)$$

$$R^2 = 1 - \frac{SSE}{SST}. \quad (4.30)$$

З ANOVA-таблиці дисперсійного аналізу (табл. 4.1) видно, що у випадку багатофакторної регресії сума квадратів залишків має $(n - k)$ ступенів вільності, а загальна сума квадратів $(n - 1)$ ступінь вільності.

Оцінений, як ми вже показували раніше, коефіцієнт детермінації — це коефіцієнт, у якого відповідні суми квадратів скориговані на їхні ступені вільності, тобто:

$$\bar{R}^2 = 1 - \frac{SSE / (n - k)}{SST / (n - 1)} = 1 - (1 - R^2) \frac{n - 1}{n - k}. \quad (4.31)$$

4.9. ПЕРЕВІРКА МОДЕЛІ НА АДЕКВАТНІСТЬ ЗА F -КРИТЕРІЄМ ФІШЕРА

Для перевірки адекватності багатофакторної регресійної моделі, як і у випадку простої лінійної моделі, використовується F -критерій Фішера.

При цьому нуль-гіпотеза узагальнюється і має вигляд:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$$

проти альтернативної гіпотези H_1 : хоча б одне значення β_i відмінне від нуля.

Якщо нуль-гіпотеза H_0 не правильна, то тоді правильна гіпотеза H_1 , тобто не всі параметри незначною мірою відрізняються від нуля, що дає підставу вважати, що відібрані фактори пояснюють зміну залежної величини y . Для перевірки H_0 -гіпотези розраховується F -статистика Фішера з p та $(n - p - 1)$ ступенями вільності:

$$F_{p, n-p-1} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 / p}{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n - p - 1)}, \quad (4.32)$$

де p — кількість факторів, які увійшли в модель; n — загальна кількість спостережень.

За F -таблицями Фішера, як і у випадку простої регресії, знаходимо критичне значення $F_{\alpha, p}$ з p та $(n - p - 1)$ ступенями вільності, задавши попередньо рівень помилки $\alpha \cdot 100\%$ (або рівень довіри $(1 - \alpha) \cdot 100\%$).

Якщо $F > F_{\alpha, p}$, тоді нуль-гіпотеза відкидається, що свідчить про адекватність побудованої моделі. У протилежному випадку вона приймається і модель вважається неадекватною.

4.9.1. Зв'язок між коефіцієнтом детермінації (R^2) та F -відношенням Фішера

Покажемо, що між коефіцієнтом детермінації R^2 та F -відношенням Фішера є зв'язок. Розпишемо вираз (4.32), попередньо замінивши p на $(k - 1)$, а $(n - p - 1)$ на $(n - k)$, де k — кількість параметрів, включаючи перетин.

Зазначимо, що ступені вільності ми виразили через кількість параметрів моделі, що підлягають оцінці. Виходячи з (4.32), отримаємо:

$$X_{n \times (p+1)} = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix} \text{ — матриця розмірності } n \times (p+1) \text{ } n$$

спостережень за p змінними x_1, x_2, \dots, x_p , де перший стовпець вміщує значення 1 для отримання перетину. Матрицю X ще називають матрицею спостережень.

$$\beta_{((p+1) \times 1)} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \dots \\ \beta_p \end{pmatrix} \text{ — вектор розміру } ((p+1) \times 1) \text{ невідомих параметрів.}$$

$$\varepsilon_{(n \times 1)} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{pmatrix} \text{ — вектор розміру } (n \times 1) \text{ } n \text{ випадкових величин } \varepsilon_i.$$

Виходячи з введених позначень для (4.35), отримаємо:

$$\begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{21} & \dots & x_{p1} \\ 1 & x_{12} & x_{22} & \dots & x_{p2} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{1n} & x_{2n} & \dots & x_{pn} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \dots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{pmatrix}. \quad (4.36)$$

Вираз (4.36) зручно переписати у вигляді:

$$y = X\beta + \varepsilon. \quad (4.37)$$

Вираз (4.37) є записом простої лінійної багатofакторної регресії у матричному вигляді.

4.10.2. Припущення класичної лінійної багатofакторної регресії у матричному вигляді

Запишемо основні припущення у матричному вигляді

Припущення 1

Математичне сподівання i -го значення ($i = \overline{1, n}$) випадкової величини ε_i дорівнює нулеві, або вектор випадкових величин дорівнює нулеві $E(\varepsilon) = 0$:

$$E \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} E(\varepsilon_1) \\ E(\varepsilon_2) \\ \dots \\ E(\varepsilon_n) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix}. \quad (4.38)$$

Припущення 2

Випадкові величини незалежні між собою. У матричному вигляді це можна записати таким чином:

$$E(\varepsilon\varepsilon') = E \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{pmatrix} (\varepsilon_1\varepsilon_2 \dots \varepsilon_n), \quad (4.39)$$

де ε' — транспланований вектор-стовпець випадкових величин, тобто вектор-рядок.

Розписавши (4.39), отримаємо:

$$E(\varepsilon\varepsilon') = E \begin{pmatrix} \varepsilon_1^2 & \varepsilon_1\varepsilon_2 & \dots & \varepsilon_1\varepsilon_n \\ \varepsilon_2\varepsilon_1 & \varepsilon_2^2 & \dots & \varepsilon_2\varepsilon_n \\ \dots & \dots & \dots & \dots \\ \varepsilon_n\varepsilon_1 & \varepsilon_n\varepsilon_2 & \dots & \varepsilon_n^2 \end{pmatrix}. \quad (4.40)$$

Застосовуючи оператор математичного сподівання до кожного елемента матриці, з огляду на властивості гомоскедастичності та відсутність зв'язку між випадковими величинами отримаємо:

$$E(\varepsilon\varepsilon') = \begin{pmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma^2 \end{pmatrix} = \sigma^2 \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix} = \sigma^2 I, \quad (4.41)$$

що скорочено можна записати:

$$X'Xb = X'y. \quad (4.48)$$

Якщо зворотна матриця $(X'X)^{-1}$ до $(X'X)$ існує, то, помноживши обидві частини (4.48) на цю матрицю, отримаємо:

$$(X'X)^{-1}(X'X)b = (X'X)^{-1}X'y, \quad (4.49)$$

де $(X'X)^{-1}(X'X) = I$ — одинична матриця розміру $(p+1) \times (p+1)$, що дає:

$$\begin{aligned} Ib &= (X'X)^{-1}X'y; \\ b &= (X'X)^{-1}X'y. \end{aligned} \quad (4.50)$$

Рівняння (4.50) є фундаментальним результатом для визначення невідомих параметрів у матричному вигляді.

Вектор невідомих параметрів може бути отриманий з (4.50) або з (4.44) шляхом диференціювання за кожним параметром (b_0, \dots, b_p) та прирівнюванням часткових похідних до нуля.

4.10.4. Дисперсійно-коваріаційна матриця параметрів регресії

Застосування теорії матриць допомагає не тільки знайти дисперсії параметрів b , а й встановити коваріації між двома попарними їхніми значеннями, тобто між b_i та b_j ; $i \neq j$.

За означенням дисперсійно-коваріаційна матриця для b є:

$$\text{var-cov}(b) = E\{[(b - E(b))[(b - E(b))']]\},$$

що можна записати:

$$\text{var-cov}(b) = \begin{bmatrix} \text{var}(b_0) & \text{cov}(b_0, b_1) & \dots & \text{cov}(b_0, b_p) \\ \text{cov}(b_1, b_0) & \text{var}(b_1) & \dots & \text{cov}(b_1, b_p) \\ \dots & \dots & \dots & \dots \\ \text{cov}(b_p, b_0) & \text{cov}(b_p, b_1) & \dots & \text{var}(b_p) \end{bmatrix} \quad (4.51)$$

Можна показати (доведення ми не наводимо), що

$$\text{var-cov}(b) = \sigma^2 (X'X)^{-1}, \quad (4.52)$$

де σ^2 — дисперсія випадкової величини ε ; $(X'X)^{-1}$ — зворотна матриця до матриці $(X'X)$.

4.10.5. Оцінка дисперсії випадкової величини ε

Можна показати, що оцінена дисперсія у випадку багатофакторної регресії дорівнює:

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_{i=1}^n e_i^2}{n - k}; \quad (4.53)$$

або у матричному вигляді

$$\hat{\sigma}_\varepsilon^2 = \frac{e'e}{n - k}$$

де k — кількість параметрів, включаючи перетин; $(n - k)$ — ступені вільності

для $\sum_{i=1}^n e_i^2 = SSE$.

Ми можемо обчислити $e'e$. У випадку простої регресії, тобто для однієї незалежної змінної, маємо:

$$\sum_{i=1}^n e_i^2 = SST - SSR = \sum_{i=1}^n \bar{y}_i^2 - b_1^2 \sum_{i=1}^n \bar{x}_i^2. \quad (4.54)$$

Вираз (4.54) для двох незалежних змінних можна записати таким чином:

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n \bar{y}_i^2 - b_1 \sum_{i=1}^n \bar{y}_i \bar{x}_{1i} - b_2 \sum_{i=1}^n \bar{y}_i \bar{x}_{2i}. \quad (4.55)$$

Узагальнюючи (4.54) та (4.55) на випадок p незалежних змінних, отримаємо:

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n \bar{y}_i^2 - b_1 \sum_{i=1}^n \bar{y}_i \bar{x}_{1i} - \dots - b_p \sum_{i=1}^n \bar{y}_i \bar{x}_{pi}; \quad (4.56)$$

де $\bar{y}_i = (y_i - \bar{y})$; $\bar{x}_{ij} = (x_{ij} - \bar{x}_j)$.

У матричному вигляді загальну суму квадратів та суму квадратів, що пояснюють регресію, можна записати:

$$SST = \sum_{i=1}^n (y_i - \bar{y})^2 = y'y - n\bar{y}^2; \quad (4.57)$$

$$SSR = b_1 \sum_{i=1}^n \bar{y}_i \bar{x}_{1i} + \dots + b_p \sum_{i=1}^n \bar{y}_i \bar{x}_{pi} = b'X'y - n\bar{y}^2, \quad (4.58)$$

де $n\bar{y}^2$ відоме під назвою “корекція на середні значення”.

З (4.57), (4.58) отримаємо:

$$e'e = y'y - b'X'y. \quad (4.59)$$

Як тільки ми встановили значення $e'e$, легко обчислити оцінку дисперсії випадкової величини, а використовуючи (4.53), легко знайти оцінені значення дисперсійно-кореляційної матриці (4.51).

4.10.6. Перевірка гіпотез щодо параметрів багатofакторної регресії у матричному вигляді

Одним із важливих припущень у багатofакторному регресійному аналізі є припущення про нормальний закон розподілу випадкової величини з нульовим математичним сподіванням та постійною дисперсією σ^2 , яке в матричному вигляді записується:

$\varepsilon \sim N(0, \sigma^2 I)$, де ε та 0 є векторами-стовпчиками розмірності $(n \times 1)$, а I — одиничною матрицею розмірності $(n \times n)$.

Можна показати, що у випадку багатofакторної регресії кожний параметр також відповідає нормальному законові розподілу:

$$b \sim N\left[\beta, \sigma^2(X'X)^{-1}\right] \quad (4.60)$$

з математичним сподіванням, яке дорівнює значенню параметра узагальненої регресії β , та дисперсією, яка дорівнює дисперсії випадкової величини σ^2 , помноженої на відповідний діагональний елемент зворотної матриці. Як ми вже не раз підкреслювали, справжнє значення дисперсії випадкової величини невідоме, тому ми замінюємо його на оцінку дисперсії. Така заміна приводить до того, що кожний елемент вектора (4.60) відповідатиме вже t -розподілу Ст'юдента з $(n - k)$ ступенями вільності, що дає нам змогу обчислити t -статистику для кожного параметра:

$$t = \frac{b_i - \beta_i}{\hat{\sigma}_{b_i}} \quad \text{з } df = (n - k), \quad (4.61)$$

де k — кількість параметрів, включаючи перетин.

T -розподіл Ст'юдента дає змогу протестувати гіпотезу щодо значення кожного параметра та побудувати їхні інтервали довіри, як і у випадку простої лінійної регресії. Нагадаємо, як це робиться.

Для перевірки параметрів багатofакторної регресії необхідно задати $\alpha \cdot 100\%$ рівень значимості, а потім побудувати t -статистику для кожного параметра окремо:

$$t = \frac{b_i - \beta_i^*}{\hat{\sigma}_{b_i}} \quad \text{з } (n - k) \text{ ступенями вільності,} \quad (4.62)$$

де b_i — оцінка параметра β_i , отримана за методом найменших квадратів; β_i^* — гіпотетичне значення, яке має прийняти параметр β_i ; $\hat{\sigma}_{b_i}^2$ — оцінка дисперсії параметра (з регресії); n — кількість спостережень; k — загальна кількість оцінених параметрів.

В економетриці поширеною формою нуль-гіпотези є така:

$$H_0: \beta_i^* = 0$$

проти альтернативної:

$$H_1: \beta_i^* \neq 0.$$

У цьому випадку t -статистика для параметрів має вигляд:

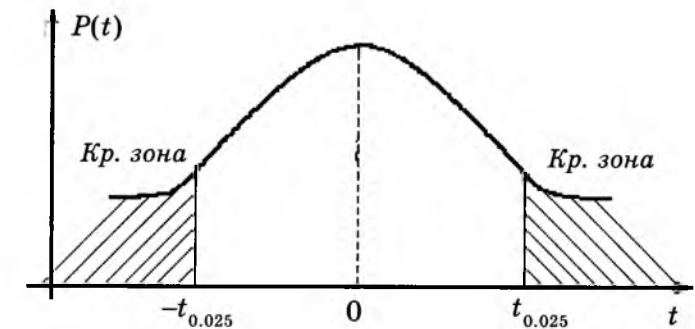
$$t^* = \frac{b_i}{\hat{\sigma}_{b_i}}. \quad (4.63)$$

Її значення порівнюємо з табличним значенням, яке дає змогу знайти критичну зону з $(n - k)$ ступенями вільності.

Якщо значення t^* потрапляє не у критичну зону, можемо зробити висновок, що з імовірністю $(1 - \alpha) \cdot 100\%$ оцінка є статистично незначимою (тобто ми приймаємо нуль-гіпотезу).

Двовимірний тест для нуль-гіпотези показано на мал. 4.2 ($\alpha = 0.05$).

З малюнка видно, якщо значення t^* потрапляє в критичну (заштриховану) зону, то ми нуль-гіпотезу відкидається.



T -статистика (4.63) є ніщо інше, як відношення b_i до оцінки свого стандартного відхилення або до свого середньоквадратичного відхилення.

У багатьох програмах поряд із значенням b_i видають значення оцінки стандартного відхилення та відношення між b_i і цією оцінкою. Це відношення називають t -значенням для відповідного параметра. Якщо воно перевищує критичне значення, яке ми знаходимо за таблицею, то приймаємо гіпотезу і вважаємо, що відповідний параметр статистично значимий.

4.10.7. Знаходження інтервалів довіри для параметрів

Для того, щоб визначити, як же знайдені оцінки параметрів багатofакторної регресії пов'язані з параметрами узагальненої регресії, потрібно побудувати інтервали довіри для параметрів.

Процедура побудови інтервалів довіри є аналогічною процедурі тестування нуль-гіпотези для параметрів за t -тестом Ст'юдента, який ми розглянули вище.

Спочатку необхідно розрахувати t -статистику для кожного з параметрів багатofакторної регресії:

$$t = \frac{b_i - \beta_i}{\hat{\sigma}_{b_i}}.$$

Потім ми обираємо рівень значимості (α або $\alpha \cdot 100\%$), відповідно рівень довіри дорівнюватиме: $(1-\alpha)$ або $(1-\alpha) \cdot 100\%$. За t -таблицею Ст'юдента знаходимо $t_{\alpha/2}$ з $(n-k)$ ступенями вільності. Тоді можемо записати:

$$P(-t_{\alpha/2} < t < t_{\alpha/2}) = (1-\alpha). \quad (4.64)$$

Проводячи заміну $t = (b_i - \beta_i) / \hat{\sigma}_{b_i}$, отримаємо:

$$P\left\{-t_{\alpha/2} < \frac{b_i - \beta_i}{\hat{\sigma}_{b_i}} < t_{\alpha/2}\right\} = (1-\alpha) \quad (4.65)$$

або

$$P\{b_i - t_{\alpha/2}\hat{\sigma}_{b_i} < \beta_i < b_i + t_{\alpha/2}\hat{\sigma}_{b_i}\}$$

з $(n-k)$ ступенями вільності,

або

$$\beta_i = b_i \pm t_{\alpha/2}\hat{\sigma}_{b_i}$$

з $(n-k)$ ступенями вільності.

4.10.8. Прогнозування за багатofакторною регресійною моделлю

Якщо побудована регресійна модель адекватна за F -критерієм Фішера, її можна використовувати для прогнозу залежної змінної. Припустимо, що нам відомі значення факторів в $(n+k)$ -період, тоді ми можемо отримати прогнозне значення \hat{y}_{n+k} за моделлю

$$\hat{y}_{n+k} = b_0 + b_1 x_{1,n+k} + \dots + b_p x_{p,n+k} \quad (4.64)$$

або в матричній формі:

$$\hat{y}_{n+k} = x'_{n+k} \times b, \quad (4.65)$$

де $x'_{n+k} = \{1, x_{1,n+k}, \dots, x_{p,n+k}\}$, $b' = \{b_0, b_1, \dots, b_p\}$.

Як уже відомо з розділу 2, теорія прогнозування дає змогу отримати точкові та інтервальні прогнозні значення. Точкові прогнозні значення знаходимо за (4.64), що аналогічно запису в матричній формі (4.65). Інтервальний прогноз, як і у випадку простої лінійної регресії, отримаємо для математичного сподівання залежної змінної y та для індивідуального значення y . Відмітимо, що дисперсії величини y будуть у цих випадках різними.

Побудова інтервалів довіри для математичного сподівання залежної змінної y

Для того, щоб отримати інтервальний прогноз математичного сподівання залежної змінної, розглянемо, чому дорівнює дисперсія цієї величини. Можна показати¹, що

$$\text{var}(E(y_j/x_j)) = \text{var}(\hat{y}_j/x_j) = \sigma_\varepsilon^2 x'_j (X'X)^{-1} x_j, \quad (4.65)$$

де σ_ε^2 — дисперсія випадкової величини ε ; x'_j — вектор значень з p факторів у період j . Виходячи з цього, інтервали довіри для $100(1-\alpha)$ рівня довіри математичного сподівання y дорівнюватимуть:

$$\hat{y}_j - t_{\alpha/2} \sqrt{\hat{\sigma}^2 x_j (X'X)^{-1} x_j} \leq E(y/x_j) \leq \hat{y}_j + t_{\alpha/2} \sqrt{\hat{\sigma}^2 x_j (X'X)^{-1} x_j}. \quad (4.66)$$

Побудова інтервалів довіри для індивідуального значення залежної змінної y

Формула для дисперсії індивідуального значення залежної змінної y буде такою:

$$\text{var}(y_j/x_j) = \sigma_\varepsilon^2 (1 + x'_j (X'X)^{-1} x_j). \quad (4.67)$$

Звідси дещо змінюється формула для обчислення інтервалу довіри, а саме:

$$\hat{y}_j - t_{\alpha/2} \sqrt{\hat{\sigma}^2 (1 + x_j (X'X)^{-1} x_j)} \leq y_j \leq \hat{y}_j + t_{\alpha/2} \sqrt{\hat{\sigma}^2 (1 + x_j (X'X)^{-1} x_j)}. \quad (4.68)$$

¹ Johnston J. Econometric Methods. — 3rd ed. — New York: McGraw-Hill Book Company, 1984. — P. 195—196.

4.11. МЕТОДИ ПОБУДОВИ БАГАТОФАКТОРНОЇ РЕГРЕСІЙНОЇ МОДЕЛІ

4.11.1. Вибір “найкращого” рівняння регресії

Нехай відібрано множину факторів x_1, x_2, \dots, x_p , що впливають на показник y , який досліджується. Є два протилежних критерії для вибору кінцевої моделі регресійного аналізу.

1. Якщо ми хочемо зробити модель корисною для прогнозу, маємо включити якомога більше факторів для того, щоб визначення величин, які прогнозуються, було надійнішим.

2. Оскільки отримання інформації з послідовним контролем при великій кількості змінних x потребує великих витрат, слід прагнути, щоб модель включала якомога менше факторів x .

Компромісом між цими крайнощами і є те, що називають вибором “найкращого” рівняння регресії. Для реалізації такого вибору немає єдиної статистичної процедури. Взагалі ж існує досить велика кількість методів побудови регресійної моделі, найвідомішими з яких є:

- 1) метод усіх можливих регресій;
- 2) метод виключень;
- 3) кроковий регресійний аналіз;
- 4) деякі модифікації попередніх методів;
- 5) гребенева регресія;
- 6) ПРЕС-регресія;
- 7) регресія на головні компоненти;
- 8) ступеневий регресійний аналіз та інші.

Ми розглянемо лише найбільш поширені на практиці методи побудови лінійних регресійних моделей.

4.11.2. Метод усіх можливих регресій

Метод усіх можливих регресій був історично першим методом побудови регресійної моделі. Він дуже громіздкий і може бути реалізований лише на ЕОМ.

Метод потребує побудови кожного з усіх можливих регресійних рівнянь, які обов'язково включають член b_0 . Оскільки для кожного фактора x_i є дві можливості — входити або не входити в регресію, то всього буде 2^p рівнянь (де p — кількість факторів x_i , $i = \overline{1, p}$). Кожне з рівнянь потім оцінюється за допомогою 3 критеріїв: R^2 , MSE та C_p — статистики (критерій Малоуза).

Розглянемо ідею цього методу на прикладі лінійної регресійної моделі з 4 факторами: x_1, x_2, x_3, x_4 . При цьому матимемо $2^4=16$ усіх можливих рівнянь, які розіб'ємо на 5 серій:

I серія моделей включає тільки один випадок:

$$y = b_0 + e.$$

II серія — всі можливі однофакторні рівняння, в нашому випадку їх чотири:

$$y = b_0 + b_i x_i + e, \quad i = \overline{1, 4}.$$

III серія — всі можливі двофакторні моделі.

IV серія — всі трифакторні моделі.

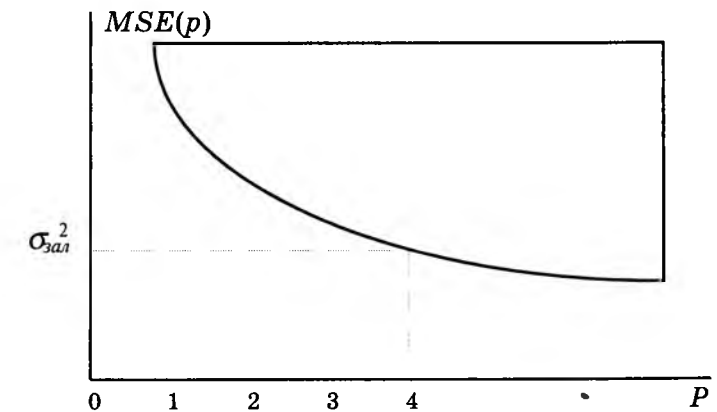
V серія — всі чотирифакторні. Це буде, як і в I серії, одна модель:

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + e.$$

Після того, як ми розбили всі моделі за серіями, проранжуємо їх всередині кожної серії за значенням R^2 (обмежимося розглядом тільки критерію R^2 та критерію MSE ; критерій Малоуза розглядати не будемо). Виберемо моделі, які мають найбільше значення коефіцієнта детермінації R^2 у кожній з серій, та проаналізуємо, чи є якась закономірність у змінних, які входять у кожне з “найкращих” рівнянь. Вибір остаточного рівняння — це деякою мірою суб'єктивна оцінка дослідника. У випадку, коли важко зробити такий вибір, можна розглянути додатковий критерій — середній квадрат залишків: MSE .

Якщо для певної задачі побудовано всі регресійні рівняння, то, розглядаючи залежність величини середнього квадрата залишків (MSE) від числа змінних (p), іноді можна найкращим способом обрати кількість змінних, які необхідно зберегти в регресійній моделі. Якщо ми до такої моделі додаватимемо все нові й нові фактори, середній квадрат залишків буде стабілізуватися та наближатися до дисперсії залишків σ_e^2 (за умови, що найважливіші змінні увійшли у модель, а кількість факторів у 5 — 6 разів перевищує кількість спостережень (мал. 4.3).

Аналіз усіх можливих рівнянь регресії — дуже громіздка і ненадійна процедура. З одного боку, метод дає можливість розглянути і дослідити усі можливі рівняння, але, з іншого боку, при великій кількості факторів це призводить до великих витрат машинного часу, збільшення тривалості



Малюнок 4.3. Залежність між значенням MSE та кількістю факторів

аналізу, можливих помилок і т.д. Виходячи з цього, метод усіх можливих рівнянь краще використовувати при невеликій кількості факторів, які входять до моделі.

4.11.3. Метод виключень

Цей метод економічніший, ніж метод усіх регресій. Загальний алгоритм методу складається з 5 етапів.

1. На першому етапі розраховується регресійне рівняння, яке включає всі фактори, що входять до моделі. Якщо було відібрано p факторів, то базове регресійне рівняння має вигляд:

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_px_p + e.$$

2. На другому етапі обчислюється величина часткового F -критерію для кожного фактора (частковий F -критерій пояснимо дещо пізніше).

3. Найменше значення часткового F -критерію, яке порівнюється із задалегідь обраним критичним значенням $F_{кр}$, позначається F_l .

4. Якщо $F_l < F_{кр}$, то відповідний фактор виключається з рівняння. Проводиться новий розрахунок регресійного рівняння вже без цього виключеного фактора і знову переходять до етапу 2.

5. Якщо $F_l > F_{кр}$, то регресійне рівняння залишають без змін.

У статистиці метод виключень є досить поширеним, бо дозволяє відразу побачити всі фактори в моделі.

У деяких комп'ютерних програмах замість часткового F -критерію використовується t -критерій, який є коренем квадратним від значення часткового F -критерію. Крім того, в них інколи використовується термін “ F -критерій для виключення”, ідентичний терміну “частковий F -критерій”.

4.11.4. Кроковий регресійний метод

Кроковий регресійний метод діє в зворотному порядку порівняно з методом виключень. Фактори по черзі включаються в модель доти, поки вона не стане задовільною. Порядок включення вибирається за допомогою коефіцієнта кореляції як міри важливості факторів (незалежних змінних), які ще не включені в модель. Кроковий регресійний метод також зручно подати в вигляді узагальненого алгоритму, який умовно розіб'ємо на три етапи.

Алгоритм методу

1. Спочатку обирається фактор x_i ($i = \overline{1, p}$), який має найбільший коефіцієнт кореляції з y (нехай це буде змінна x_1). Будується регресійне рівняння з однією незалежною змінною $y = b_0 + b_1x_1$. Після цього перевіряємо, чи значима ця змінна за частковим F -критерієм. Якщо ні, то приймаємо $y = \bar{y}$ і припиняємо процес. Якщо так, то шукаємо другу змінну,

що має найменший коефіцієнт кореляції з y , яку варто включити в модель. Нехай це буде змінна x_2 .

2. Будується нове рівняння регресії:

$$y = b_0 + b_1x_1 + b_2x_2.$$

Аналізується зміна коефіцієнта кореляції R^2 та розраховуються часткові F -критерії для кожного фактора. Серед них обирається найменше його значення і порівнюється із задалегідь обраним критичним значенням F -розподілу. Відповідно до результатів перевірки змінна або залишається в моделі, або відкидається.

Може так статися, що фактор, який на якомусь з етапів був найкращим для включення, потім з моделі вилучається.

3. Після цього модель перерозраховується залежно від факторів, які залишилися.

Варто зазначити, що на першому етапі, крім перевірки за коефіцієнтом кореляції, фактор — претендент на включення, може перевірятись також за частковим F -критерієм. Якщо він відповідає йому, то фактор включають у модель, якщо ні, то переходять до аналізу наступного фактора.

Процес побудови моделі припиняється, якщо жодний фактор, що знаходиться в поточному рівнянні, не вдається виключити, а новий претендент на включення не відповідає частковому F -критерію.

Кроковий метод є найпоширенішим на практиці.

У розглянутих вище методах ми використовували поняття часткового F -критерію Фішера. Для того, щоб пояснити, що таке частковий F -критерій Фішера, розглянемо деякі додаткові поняття, а саме: принцип “додаткової суми квадратів”.

4.11.5. Принцип “додаткової суми квадратів”.

Обчислення додаткової суми квадратів

Цей принцип пов'язаний з питанням, чи слід включати в модель певні фактори, чи ні. Відповіді на ці запитання можливо, досліджуючи додаткову частину суми квадратів, що пояснює регресію, пов'язану саме з додатковим включенням у модель факторів, які розглядаються.

Середній квадрат цієї додаткової суми можна порівняти з оціненою дисперсією випадкової величини ($\hat{\sigma}_e^2$). Якщо розрахований нами середній квадрат значно перевищує $\hat{\sigma}_e^2$, то відповідні члени слід включати в модель. У протилежному випадку їх можна розглядати як зайві та вилучати з моделі.

Обчислення додаткової суми квадратів

Розглянемо загальну процедуру обчислення додаткової суми квадратів.

Нехай ми маємо x_1, x_2, \dots, x_p факторів. Дослідимо дві моделі: перша містить всі p факторів, а друга — тільки q факторів, де $q < p$.

Перша модель має вигляд:

$$1. y = b_0 + b_1x_1 + b_2x_2 + \dots + b_px_p + e.$$

Знайдемо параметри $b_0(1); b_1(1); b_2(1), \dots, b_p(1)$ методом найменших квадратів і припустимо, що модель є адекватною.

Введемо позначення: $SSR(b_0(1), b_1(1), \dots, b_p(1)) = S_1$; нехай $\hat{\sigma}_\varepsilon^2(1)$ — оцінена дисперсія випадкової величини першої моделі.

Друга модель відповідно має вигляд:

$$2. y = b_0 + b_1x_1 + b_2x_2 + \dots + b_qx_q + e, \quad q < p.$$

Значення факторів з однаковими індексами збігаються для першої та другої моделі. Розрахуємо невідомі параметри за методом найменших квадратів і отримаємо: $b_0(2); b_1(2), \dots, b_q(2)$.

Позначимо $SSR(b_0(2), b_1(2), \dots, b_q(2)) = S_2$.

Тоді різниця $S_1 - S_2$ дорівнюватиме додатковій сумі квадратів, що пояснюють регресію, яка пов'язана з включенням у першу модель членів $b_{q+1}x_{q+1} + \dots + b_px_p$. Сума квадратів S_1 має p ступенів свободи, а S_2 — q ступенів свободи, відповідно величина $(S_1 - S_2)$ має $(p - q)$ ступенів свободи.

З математичної статистики відомо, якщо

$$\beta_{q+1} = \beta_{q+2} = \dots = \beta_p = 0,$$

то

$$E\{(S_1 - S_2) / (p - q)\} = \sigma_\varepsilon^2.$$

Крім того відомо, що коли помилки розподілені нормально, то величина $(S_1 - S_2)$ розподілиться як $\sigma_\varepsilon^2 \chi_{p-q}^2$ і буде незалежною від $\hat{\sigma}_\varepsilon^2$. Це означає, що ми можемо порівняти величину $(S_1 - S_2) / (p - q)$ з $\hat{\sigma}_\varepsilon^2$ за допомогою $F(p - q, \nu)$ -критерію Фішера, де ν — кількість ступенів свободи, яка пов'язана з оцінкою дисперсії випадкової величини $\hat{\sigma}_\varepsilon^2$, та використати цю властивість для перевірки гіпотези.

$$H_0: \beta_{q+1} = \beta_{q+2} = \dots = \beta_p = 0.$$

Для зручності $(S_1 - S_2)$ можна записати, як $SS(b_{q+1}, \dots, b_p | b_0, b_1, b_q)$ — сума квадратів, пов'язана з b_{q+1}, \dots, b_p , за умови, що коефіцієнти b_0, b_1, \dots, b_q входять у модель.

Використовуючи такий запис, можна послідовно отримати:

$$SS(b_0), SS(b_1 | b_0), SS(b_2 | b_0, b_1), \dots, SS(b_p | b_0, b_1, \dots, b_{p-1}).$$

Усі ці суми квадратів статистично незалежні від $\hat{\sigma}_\varepsilon^2$ та дорівнюють своїм середнім квадратам, тому що мають по одному ступеню вільності. Зазначимо, що кількість ступенів вільності для подібних сум квадратів дорівнюватиме кількості параметрів, вказаних у дужках до вертикальної

лінії. Ці додаткові суми розподілені незалежно від $\hat{\sigma}_\varepsilon^2$. Тому відповідні середні квадрати (які дорівнюють відношенню суми квадратів до числа ступенів свободи) можна поділити на $\hat{\sigma}_\varepsilon^2$ та отримати F -відношення для перевірки гіпотези про те, що правильні значення параметрів узагальненої моделі, оцінки яких створюють додаткову суму квадратів, дорівнюють нулю.

Принцип додаткової суми квадратів фактично зводиться до особливо-го випадку перевірки адекватності лінійної моделі.

При більш загальному підході додаткова сума квадратів обчислюється, виходячи із сум квадратів залишків, а не з суми квадратів, обумовлених регресією.

4.11.6. Часткові і послідовні F -критерії

Якщо в регресійній моделі міститься кілька факторів, то можемо вважати, що вони вводяться в рівняння у будь-якій бажаній послідовності.

Знайдемо

$$SS(b_i | b_0, b_1, \dots, b_{i-1}, b_{i+1}, \dots, b_k), \quad i = \overline{1, k}. \quad (4.58)$$

Матимемо один ступінь вільності для суми квадратів, яка вимірює внесок кожного коефіцієнта b_i в суму квадратів, обумовлену регресією за умови, що всі фактори, крім i -го, вже входять у модель.

Іншими словами, матимемо міру для оцінки включення b_i в модель, яка спочатку не містила такого члена. Тобто ми знатимемо міру важливості параметра, якби він був доданий у модель останнім.

Відповідний середній квадрат, який дорівнює сумі квадратів (4.58), поділений на кількість ступенів вільності (нагадаємо, що (4.58) має один ступінь вільності), може порівнюватися з величиною $\hat{\sigma}_\varepsilon^2 = S^2$ за допомогою F -критерію. Такий варіант F -критерію називають **частковим F -критерієм для b_i** .

При побудові відповідної моделі частковий F -критерій — це корисний критерій для вирішення питання про додавання чи вилучення факторів з моделі.

Якщо фактори додаються до регресійної моделі послідовно один за одним, то ми говоримо про послідовний F -критерій. Це якраз і є частковий F -критерій щодо фактора, який вводиться на даному етапі.

У деяких статистичних пакетах STATGRAPH, BMDP, SPSS, SAS частковим F -критерієм називають “ F -критерій для вилучення”; а послідовний F -критерій — “ F -критерієм для включення”.

Можна також показати, що $(t = F^{1/2})$. Тому частковий F -критерій з кількістю ступенів вільності 1 та ν для перевірки гіпотез $H_0: \beta_i = 0, H_1: \beta_i \neq 0$ дорівнює квадрату t -статистики з ν ступенями вільності, яка підраховується за формулою: $t = b_i / \hat{\sigma}_{b_i}$. Отже, перевірка гіпотези про значимість i -го

параметра може виконуватись або з використанням часткового F -критерію, або t -статистики.

Аналізуючи відповідні статистичні таблиці F -розподілу Фішера та t -розподілу Ст'юдента, можна переконатися, що $F(1, v, \alpha) = \{t(v, \alpha / 2)\}^2$ за будь-яких значень v та α .

У багатьох програмах використовуються обидві статистики, тому що через помилки округлення між ними немає рівності.

ПІДСУМУЄМО ВИВЧЕНЕ

Після вивчення цього розділу ви зможете

1. Будувати регресійні моделі з більш ніж однією незалежною змінною.
2. Оцінювати невідомі параметри багатофакторної регресії.
3. Проводити ANOVA-дисперсійний аналіз.
4. Перевіряти модель на адекватність за F -критерієм Фішера.
5. Розраховувати множинний коефіцієнт кореляції, простий та оцінений коефіцієнти детермінації.
5. Перевіряти незалежні змінні на наявність мультиколінеарності.
6. Перевіряти гіпотези про значимість параметрів багатофакторної регресії.
7. Будувати інтервали довіри для параметрів.
8. Будувати інтервали довіри для прогнозних величин.

Короткий огляд розділу

1. *Багатофакторна лінійна регресія є звичайним продовженням простої лінійної регресії, яка розглядалась у попередніх розділах. У багатофакторній регресії ми припускаємо, що на залежну змінну y може впливати більше, ніж один фактор. Наприклад, обсяг продажу компанії може залежати як від ціни товару, так і від кількості грошей, затрачених на рекламу і т.ін.*

2. Основні припущення у випадку багатофакторної лінійної регресії:

- а) випадкова величина ε є нормально розподіленою з математичним сподіванням нуль та постійною дисперсією;
- б) значення випадкової величини не залежать від значень p незалежних змінних, які входять у модель;
- в) випадкові величини ε є незалежними одна від одної;
- г) фактори мають бути лінійно незалежними між собою. Якщо вони лінійно залежні, ми маємо *мультиколінеарність*.

3. У багатофакторній регресії, як і в простій лінійній регресії, невідомі параметри розраховуються за методом найменших квадратів, який полягає в мінімізації суми квадратів помилок.

4. Ефект кожного фактора або кожної незалежної змінної на залежну змінну обчислюється *частковими регресійними коефіцієнтами (параме-*

трами). Частковий регресійний коефіцієнт показує, наскільки зміниться значення залежної змінної при зміні на одиницю значення відповідного фактора за умови, що значення інших факторів залишаються постійними.

6. За допомогою F -тесту можна перевірити значимість усіх параметрів одночасно, а за допомогою t -тесту — кожного параметра окремо. Пригадайте, що F - та t -тести були еквівалентними в моделі простої лінійної регресії, але якщо є більш ніж одна незалежна змінна, ці тести не є еквівалентними.

Основні формули

Узагальнена багатофакторна регресійна модель:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + \varepsilon_i.$$

Вибіркова багатофакторна регресійна модель:

$$y_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + \dots + b_p x_{pi} + e_i.$$

Оцінки параметрів у багатофакторній регресії тільки з двома факторами визначаються за формулою

$$b_1 = \frac{\left(\sum_{i=1}^n x_{i2}^2 - n\bar{x}_2^2 \right) \left[\sum_{i=1}^n x_{i1} y_i - n\bar{x}_1 \bar{y} \right] - \left(\sum_{i=1}^n x_{i1} x_{i2} - n\bar{x}_1 \bar{x}_2 \right) \left(\sum_{i=1}^n x_{i2} y_i - n\bar{x}_2 \bar{y} \right)}{\left(\sum_{i=1}^n x_{i1}^2 - n\bar{x}_1^2 \right) \left(\sum_{i=1}^n x_{i2}^2 - n\bar{x}_2^2 \right) - \left(\sum_{i=1}^n x_{i1} x_{i2} - n\bar{x}_1 \bar{x}_2 \right)^2};$$

$$b_2 = \frac{\left(\sum_{i=1}^n x_{i1}^2 - n\bar{x}_1^2 \right) \left(\sum_{i=1}^n x_{i2} y_i - n\bar{x}_2 \bar{y} \right) - \left(\sum_{i=1}^n x_{i1} x_{i2} - n\bar{x}_1 \bar{x}_2 \right) \left(\sum_{i=1}^n x_{i1} y_i - n\bar{x}_1 \bar{y} \right)}{\left(\sum_{i=1}^n x_{i1}^2 - n\bar{x}_1^2 \right) \left(\sum_{i=1}^n x_{i2}^2 - n\bar{x}_2^2 \right) - \left(\sum_{i=1}^n x_{i1} x_{i2} - n\bar{x}_1 \bar{x}_2 \right)^2};$$

$$b_0 = \bar{y} - b_1 \bar{x}_1 - b_2 \bar{x}_2.$$

Коефіцієнт детермінації:

$$R^2 = 1 - \frac{SSE}{SST}.$$

Оцінений коефіцієнт детермінації:

$$\bar{R}^2 = 1 - \frac{SSE / (n - p - 1)}{SST / (n - 1)} = 1 - (1 - R^2) \frac{(n - 1)}{(n - p - 1)}.$$

F-відношення:

$$F_{p, n-p-1} = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \frac{n - p - 1}{p}.$$

ВПРАВИ І КОМЕНТАРІ

Вправа 4.1. Багатофакторна регресія

Є такі дані (табл. 4.2):

Таблиця 4.2

Відстань y	Вага спортсмена x_1	Час x_2
925	210	4.8
850	185	4.7
1622	225	4.7
1121	215	4.6
658	180	4.9
977	212	4.6
574	195	5.0

Оцініть параметри регресійної моделі.

Розв'язок

$$\bar{y} = 961; \bar{x}_1 = 203.14; \bar{x}_2 = 4.66;$$

$$\sum (x_{i1} - \bar{x}_1)^2 \sum (x_{i1} - \bar{x}_2)^2 = 0.4171.$$

$$\sum (y - \bar{y})(x_{i1} - \bar{x}_1) = 28417; \sum (y - \bar{y})(x_{i2} - \bar{x}_2) = -495.9;$$

$$\sum (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) = -20.46.$$

$$b_1 = \frac{(28417)(0.4171) - (-495.9)(-20.46)}{(1674.86)(0.4171) - (-20.46)^2} = 6.101;$$

$$b_2 = \frac{(-495.9)(1674.86) - (28417)(-20.46)}{(1674.86)(0.4171) - (-20.46)^2} = -889.602;$$

$$b_0 = \bar{y} - b_1 x_1 - b_2 x_2 = 961 - 6.101(203.143) - (-889.602)(4.66) = 3864.63.$$

$$\hat{y} = 3864.63 + 6.101x_1 - 889.602x_2.$$

Вправа 4.2. Значимість параметрів регресії

Використайте дані та ваші результати з вправи 4.1 для перевірки значимості оцінених параметрів b_1 та b_2 з рівнем значимості 5%.

Розв'язок

$$SSE = \sum (y_i - \hat{y}_i)^2 = 103446.497;$$

$$\hat{\sigma}_\varepsilon^2 = SSE / (n - p - 1) = 103446.37 / (7 - 2 - 1) = 25861.59;$$

$$\hat{\sigma}_{b_1}^2 = \frac{\hat{\sigma}_\varepsilon^2 \sum (x_2 - \bar{x}_2)^2}{\sum (x_{1i} - \bar{x}_1)^2 \sum (x_{2i} - \bar{x}_2)^2 - [\sum (x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2)]^2} =$$

$$= \frac{25861.59(0.4171)}{1647.86(0.4171) - 418.94^2} = 38.51;$$

$$\hat{\sigma}_{b_2}^2 = \frac{\hat{\sigma}_\varepsilon^2 \sum (x_1 - \bar{x}_1)^2}{\sum (x_{1i} - \bar{x}_1)^2 \sum (x_{2i} - \bar{x}_2)^2 - [\sum (x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2)]^2} =$$

$$= \frac{25861(1674.86)}{1647.86(0.4171) - 418.94^2} = 154606.24.$$

Критичне значення для $t_{0.025,4} = 2.776$, отже, жодна із змінних не є статистично значимою. Але параметр b_2 є майже значимим, і був би значимим, якщо б ми вибрали рівень значимості 10%, а не 5%.

Вправа 4.3. Спільний тест значимості, або перевірка моделі на адекватність

Використайте дані результатів із вправ 4.1 та 4.2 для перевірки значимості b_1 та b_2 при 5%-ному рівні значимості.

Розв'язок. Спільний тест означає, що ми перевіряємо $H_0: \beta_1 = \beta_2 = 0$ з використанням F-тесту.

$$F_{p, n-p-1} = (SSR / p) / (SSE / (n - p - 1));$$

$$SST = SSE + SSR;$$

$$SST = \sum (y_i - \bar{y})^2 = 717972;$$

$$SSE = \sum (y_i - \hat{y}_i)^2 = 103446.497;$$

$$SSR = 717972 - 103446.497 = 614525.503;$$

$$F = [614525.503/2]/[103446.497/4] = 11.88.$$

Критичне значення для $F_{2,4,5\%} = 19.2$, отже, ми не можемо відкинути нульову гіпотезу ($F < F_{кр}$).

Вправа 4.4. Інтерпретація оцінок параметрів

Припустимо, що цікавить зв'язок між досвідом, навчанням і доходом. На підставі певної статистики ви одержали таку модель:

$$y = 180 + 12x + 32.9Z,$$

де y — доход людини в гривнях; x — досвід, років; Z — роки навчання.

Проінтерпретуйте коефіцієнти регресії для x та Z .

Розв'язок

Регресія показує, що людина отримуватиме 180 гривень незалежно від наявності досвіду чи освіти. За кожен додатковий рік досвіду очікується, що його/її доход зросте на 12 гривень. За кожен додатковий рік навчання доход зросте на 32.9 гривні.

Вправа 4.5. Коефіцієнт детермінації

Ви оцінюєте таку модель:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i,$$

що базується на 20 спостереженнях, і отримали

$$\sum (y_i - \hat{y}_i)^2 = 285, \quad \sum (y_i - \bar{y})^2 = 425.$$

Підрахуйте коефіцієнт детермінації (R^2). Чи “добре” модель пояснює виявлену закономірність?

Розв'язок

$$SSE = \sum (y_i - \hat{y}_i)^2;$$

$$SST = \sum (y_i - \bar{y})^2;$$

$$R^2 = 1 - SSE/SST = 1 - 285/425 = 0.3294.$$

$R^2 = 0.3294$ означає, що незалежні змінні x_1 та x_2 пояснюють 32.94% варіації залежної змінної y . Чи цього досить для пояснення закономірності залежить від того, яку залежну змінну ми маємо. Наприклад, пояснити 32% змін біржових цін було б дуже непогано, але пояснення лише 32% змін, що відбуваються в реалізації товарів фірми на ринку, недостатньо.

ВИКОНАЙТЕ САМОСТІЙНО

Вправи. Багатофакторна лінійна регресія

Вправа 4.6. У табл. 4.3 наведено дані про реальний валовий продукт, кількість робочої сили та кількість витрат на капітал в українському промисловому секторі (дані умовні).

Таблиця 4.3

Рік	Реальний валовий продукт, млн. грн. y	Кількість робочої сили, на 1000 осіб, x_2	Кількість витрат на капітал, млн. грн., x_3
1977	8911.4	281.5	120753
1978	10873.2	284.4	122242
1979	11132.5	289.0	125263
1980	12086.5	375.8	128539
1981	12767.5	375.2	131427
1982	16347.1	402.5	134267
1983	19542.7	478.0	139038
1984	21075.9	553.4	146450
1985	23052.0	616.7	153714
1986	26128.2	695.7	164783
1987	29563.7	790.3	176864
1988	33376.6	816.0	188146
1989	38354.3	848.4	205841
1990	46868.3	873.1	221748
1991	54308.0	999.2	239715

1. Підставте наведені дані у такі дві моделі:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon_t,$$

$$\ln y_t = \alpha_0 + \alpha_1 \ln x_{1t} + \alpha_2 \ln x_{2t} + \varepsilon_t.$$

2. Яка модель більш прийнятна і чому?

3. Для log-лінійної моделі α_1 та α_2 показують еластичність пропозиції для праці та капіталу відповідно. Як би ви обчислили ту саму еластичність для лінійної моделі?

4. Як би ви порівняли коефіцієнти R^2 обох моделей? (Зробіть розрахунки).

Вправа 4.7. У таблиці 4.4. подано щоквартальні дані про:

y — кількість проданих гербер у м. Києві; x_1 — середню роздрібну ціну гербер на ринку (грн.); x_2 — середню роздрібну ціну троянд на ринку (грн.); x_3 — середній місячний сімейний доступний доход у м. Києві; x_4 — змінну, що нумерує періоди з III кв. 1993 р. по II кв. 1997 р. (дані умовні).

Таблиця 4.4

Рік та квартал	y	x_1	x_2	x_3	x_4
1993-III	6482	1.26	2.08	121.11	1
-IV	5348	1.44	2.05	123.36	2
1994-I	5429	1.57	2.06	125.26	3
-II	6079	1.51	2.24	132.92	4
-III	4924	1.53	2.21	138.46	5
-IV	5862	1.47	2.06	148.62	6
1995-I	4216	1.59	2.16	156.28	7
-II	4253	1.23	2.02	158.98	8
-III	4038	1.60	2.00	150.49	9
-IV	3476	1.49	2.10	153.33	10
1996-I	2911	1.77	2.25	151.87	11
-II	3950	1.44	2.06	155.00	12
-III	4134	1.62	2.34	154.00	13
-IV	2868	1.46	2.00	158.20	14
1997-I	1160	1.24	2.28	155.67	15
-II	3872	1.09	2.13	158.00	16

У запропонованих функціях попиту

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x_{4t} \varepsilon_t;$$

$$y_t = \alpha_1 + \alpha_2 x_{2t} + \alpha_3 x_{3t} + \alpha_4 x_{4t} + \alpha_5 x_{5t} + u_t;$$

$$\ln y_t = \beta_1 + \beta_2 \ln x_{2t} + \beta_3 \ln x_{3t} + \beta_4 \ln x_{4t} + \beta_5 \ln x_{5t} + u_t$$

1. Підрахуйте параметри лінійної моделі та проінтерпретуйте результати.
2. Визначте параметри log-лінійної моделі та проінтерпретуйте результати.
3. β_2 , β_3 та β_4 показують особисту ціну, перехресну ціну та коефіцієнт еластичності попиту за доходом. Чи збігаються результати з попередніми сподіваннями?
4. Як би ви обчислили особисту ціну, перехресну ціну та коефіцієнт еластичності попиту за доходом для лінійної моделі?
5. Грунтуючись на вашому аналізі, яку модель ви б вибрали (якщо б вибрали) і чому?

Вправа 4.8. Нехай відома статистика витрат оборонного бюджету умовної країни за 1972 — 1991 рр. Поясніть оборонний бюджет даної країни, розглянувши таку модель:

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x_{4t} + \beta_5 x_{5t} + \varepsilon_t;$$

де y_t — щорічні витрати оборонного бюджету t (млрд. умовн. од.);
 x_{2t} — ВВП за рік t (млрд. умовн. од.);

x_{3t} — щорічні надходження коштів від військових продаж цієї країни t (млрд. умовн. од.);

x_{4t} — продажі аерокосмічної промисловості (млрд. умовн. од.);

x_{5t} — військові конфлікти із залученням більше ніж 100 військових частин. Ця змінна приймає значення 1, коли залучено 100 чи більше військових частин, та рівна нулю, коли це число менше за 100.

Для тестування цієї моделі ви маєте дані табл. 4.5.

Таблиця 4.5

Рік	Витрати оборонного бюджету y	ВВП x_2	Надходження від військових продаж x_3	Продажі аерокосмічної промисловості x_4	Конфлікти $100 + x_5$
1972	51.100	560.300	0.600	16.000	0.0
1973	52.300	590.500	0.900	16.400	0.0
1974	53.600	632.400	1.100	16.700	0.0
1975	49.600	684.900	1.400	17.000	1.0
1976	56.800	749.900	1.600	20.200	1.0
1977	70.100	793.900	1.000	23.400	1.0
1978	80.500	865.000	0.800	25.600	1.0
1979	81.200	931.400	1.500	24.600	1.0
1980	80.300	992.700	1.000	24.800	1.0
1981	77.700	1077.600	1.500	21.700	1.0
1982	78.300	1185.900	2.950	21.500	1.0
1983	74.500	1326.400	4.800	24.300	0.0
1984	77.800	1434.200	10.300	26.800	0.0
1985	85.600	1549.200	16.000	29.500	0.0
1986	89.400	1718.000	14.700	30.400	0.0
1987	97.500	1918.300	8.300	33.300	0.0
1988	105.200	2163.900	11.000	38.000	0.0
1989	117.700	2417.800	13.000	46.200	0.0
1990	135.900	2633.100	15.300	57.600	0.0
1991	162.100	2937.700	18.000	68.900	0.0

1. Обчисліть параметри цієї моделі та їхні стандартні помилки; розрахуйте R^2 та \bar{R}^2 .

2. Поясніть результати, беручи до уваги будь-які попередні сподівання щодо взаємозв'язку між y та різними змінними x .

3. Які інші змінні ви, можливо, хочете включити в модель і чому?

ТЕСТИ

Вибрати правильну відповідь на запитання

1. У багатофакторній регресії:

- а) більш ніж одна залежна змінна і тільки одна незалежна змінна;
- б) більш ніж одна незалежна змінна і тільки одна залежна змінна;
- в) більш ніж одна залежна змінна і більш ніж одна незалежна змінна;
- г) тільки одна залежна змінна і тільки одна незалежна змінна;
- д) більш ніж дві залежні змінні і більш ніж одна незалежна змінна.

2. При геометричній інтерпретації регресійної моделі з двома незалежними змінними ми будемо:

- а) пряму лінію, щоб показати зв'язок між залежною змінною та незалежними змінними;
- б) трикутник, щоб показати зв'язок між залежною змінною та незалежними змінними;
- в) площину, щоб показати зв'язок між залежною змінною та незалежними змінними;
- г) коло, щоб показати зв'язок між залежною змінною та незалежними змінними;
- д) еліпс, щоб показати зв'язок між залежною змінною та незалежними змінними.

3. У множинній регресії кожен параметр показує:

- а) загальний вплив усіх незалежних змінних на залежну змінну;
- б) вплив незалежної змінної на залежну за умови, що всі інші незалежні змінні залишаються незмінними;
- в) де площина регресії перетинає вісь y ;
- г) як частковий, так і загальний вплив незалежних змінних;
- д) точку, що дорівнює значенню перетину.

4. Мультиколінеарність виникає тоді, коли:

- а) помилка не має нульового середнього значення;
- б) помилка залежить від незалежної змінної;
- в) дві помилки корелюють між собою;
- г) незалежні змінні корелюють між собою;
- д) дисперсія помилок не є постійною.

5. Щоб перевірити значимість окремого параметра, використовують:

- а) F -тест;
- б) t -тест;

- в) x^2 -тест;
- г) біноміальний розподіл;
- д) експоненційний розподіл.

6. Для перевірки значимості одночасно всіх параметрів використовується:

- а) F -тест;
- б) t -тест;
- в) x^2 -тест;
- г) біноміальний розподіл;
- д) експоненційний розподіл.

7. За інших рівних умов, якщо ми збільшуємо кількість незалежних змінних у регресії:

- а) R^2 збільшується;
- б) R^2 зменшується;
- в) R^2 може або збільшитись, або зменшитись;
- г) немає ніякого ефекту на R^2 ;
- д) імовірність мультиколінеарності зменшується.

8. Зв'язок між R^2 та оціненим R^2 є:

- а) оцінене $R^2 = R^2$;
- б) оцінене $R^2 = R^2(n-1)/(n-p-1)$;
- в) оцінене $R^2 = [1-(1-R^2)](n-1)/(n-p-1)$;
- г) оцінене $R^2 = 1-R^2$;
- д) оцінене $R^2 = 1+R^2$.

9. Для регресії з n спостережень та p незалежних змінних зв'язок між R^2 та F є:

- а) $F_{p, n-p-1} = [(n-p-1)/p][R^2/(1-R^2)]$;
- б) $F_{p, n-p-1} = R^2/(1-R^2)$;
- в) $F_{p, n-p-1} = [n/p][R^2/(1-R^2)]$;
- г) $F_{p, n-p-1} = [(n-p-1)/p]R^2$;
- д) $F_{p, n-p-1} = R^2/(1+R^2)$.

10. Якщо ви оцінюєте рівняння доходів для 100 працівників, приймаючи кількість років навчання як одну незалежну змінну і кількість місяців навчання як іншу незалежну змінну, то ви можете:

- а) отримати оцінки, які не будуть BLUE;
- б) мати неоднакові дисперсії помилок;
- в) мати мультиколінеарність;
- г) б та в.

11. Ступені вільності для t -статистики для перевірки значимості параметрів регресії, що складається з 35 спостережень та 3 незалежних змінних, такі:

- а) 35; з) 33;
 б) 3; д) 31.
 в) 32;

12. Ступені вільності чисельника F -статистики в регресії, що складається з 50 спостережень та 4 незалежних змінних, такі:

- а) 50; з) 46;
 б) 4; д) 45.
 в) 3;

13. Ступені вільності знаменника F -статистики, що складається з 50 спостережень та 4 незалежних змінних, такі:

- а) 50; з) 46;
 б) 4; д) 45.
 в) 3;

14. Однією з проблем, що може виникнути у багатофакторній регресії і ніколи не буває в простій регресії, є:

- а) кореляція між величинами помилок;
 б) нерівна дисперсія помилок;
 в) кореляція між помилками та незалежними змінними;
 г) кореляція між незалежними змінними;
 д) помилка без нульової середньої.

ВІДПОВІСТІ НА ЗАПИТАННЯ: ТАК/НІ; ЯКЩО НІ, ПОЯСНІТЬ ЧОМУ

- Чи можливо мати незначимі t -величини для параметрів та значиму F -величину для регресії?
- У багатофакторній регресії R^2 є кращим показником для виміру якості побудованої регресії, ніж оцінений R^2 .
- Чи може бути t -статистика від'ємною?
- Негативна t -статистика одного з параметрів свідчить, що він для регресії не є важливим.
- Проста регресія — це особливий випадок багатофакторної регресії.
- Геометрична інтерпретація регресії, що складається з двох незалежних змінних, є прямою лінією.
- Мультиколінеарність виникає, коли незалежна змінна надто корельована із залежною змінною.

8. Оцінений R^2 завжди більший за R^2 .

9. T -тест використовується для перевірки спільної значимості параметрів у багатофакторній регресії.

10. F -тест використовується для перевірки значимості одного окремого параметра.

11. $SST = SSR + SSE$.

12. SST не може бути від'ємним.

ЗАПИТАННЯ І ЗАДАЧІ

1. Уявіть собі, що ви економіст з міністерства праці і цікавитесь оцінкою зв'язку між доходами та іншими факторами, такими як вік, освіта та досвід роботи:

а) яка з цих змінних повинна бути залежною змінною? Які змінні можуть бути незалежними змінними?

б) чи можете ви зіткнутись з якимись проблемами в оцінці вашої регресії?

2. Ви оцінюєте регресію, використовуючи три незалежні змінні, які мають такий зв'язок:

$$x_{3i} = 2x_{1i} + x_{2i}.$$

а) які проблеми можуть виникнути при оцінці регресії з використанням усіх трьох змінних?

б) які є шляхи розв'язання цих проблем?

3. Ви оцінюєте регресію, використовуючи 50 спостережень та 4 незалежні змінні. R^2 для регресії становить 0.64. Визначте оцінений коефіцієнт детермінації R^2 для регресії.

4. Ви оцінюєте таку регресію, використовуючи 30 спостережень:

$$y_i = 122 + 32.4x_{1i} - 0.78x_{2i};$$

$$r_{x_1, x_2} = 0.59;$$

$$\sum (x_{1i} - \bar{x}_1)^2 = 23.6;$$

$$\sum (x_{2i} - \bar{x}_2)^2 = 12.3;$$

$$\hat{\sigma}_\epsilon^2 = 41.7.$$

Перевірте значимість параметрів при рівні значимості 5%.

ВІДПОВІДІ

Тести

- | | | |
|-------|--------|--------|
| 1) б; | 6) а; | 11) д; |
| 2) в; | 7) а; | 12) б; |
| 3) б; | 8) в; | 13) д; |
| 4) г; | 9) а; | 14) г. |
| 5) б; | 10) в; | |

Так / Ні

1. Так.
2. Ні. Оцінений коефіцієнт детермінації R^2 придатніший для багатофакторної регресії, оскільки він скоригований щодо кількості незалежних змінних.
3. Так, t -значення буде від'ємним, якщо знак параметра буде від'ємним.
4. Ні. Це свідчить про те, що знак параметра є негативним.
5. Так.
6. Ні. Геометричною інтерпретацією є площа.
7. Ні. Мультиколінеарність виникає, коли незалежні змінні надто корельовані між собою.
8. Ні. Оцінений R^2 завжди менший чи дорівнює R^2 .
9. Ні. F -тест використовується для перевірки спільної значимості параметрів.
10. Ні, t -тест використовується для перевірки значимості окремих параметрів.
11. Так.
12. Так.

Запитання та проблеми

1. А. Доходи будуть залежною змінною, вік, трудовий досвід та освіта будуть незалежними змінними.
Б. Тому що вік та трудовий досвід є дуже корельованими, можлива проблема мультиколінеарності.
2. А. Оскільки x_3 є лінійною комбінацією x_1 та x_2 , то виникне проблема мультиколінеарності.
Б. Одним з методів вирішення цієї проблеми є вилучення залежних змінних з регресії.
3. Оцінений коефіцієнт детермінації:
 $\bar{R}^2 = 1 - (1 - R^2)[(n-1)/(n-p-1)] = 1 - (1 - 0.64)[(50-1)/(50-4-1)] = 0.608$.

$$4. \hat{\sigma}_{b1}^2 = \frac{\hat{\sigma}_\varepsilon^2}{(1 - r_{x1,x2}^2) \sum (x_{1i} - x_1)^2} = \frac{41.7}{(1 - 0.59)(23.6)} = 4.31;$$

$$\hat{\sigma}_{b1} = \sqrt{4.31} = 2.08;$$

$$\hat{\sigma}_{b2}^2 = \frac{\hat{\sigma}_\varepsilon^2}{(1 - r_{x1,x2}^2) \sum (x_{2i} - x_2)^2} = \frac{41.7}{(1 - 0.59)(12.3)} = 8.27;$$

$$\hat{\sigma}_{b2} = \sqrt{8.27} = 2.88;$$

$$t_{b1} = 32.4/2.08 = 15.58;$$

$$t_{b2} = -0.78/2.88 = -2.71.$$

$t_{0.025;27} = \pm 2.052$, отже, ми відкидаємо нульову гіпотезу, що $\beta_1 = 0$, але приймаємо нульову гіпотезу, що $\beta_2 = 0$.

ЗАСТОСУВАННЯ ЕКОНОМІЧНИХ ЗНАТЬ

Приклади практичного застосування багатофакторної регресії

Економетрична модель витрат виробництва

Поліноміальні регресійні моделі досить поширені на практиці, особливо для встановлення зв'язку між обсягами продукції, що випускається, та типами витрат на її виробництво.

У загальному випадку поліноміальна модель має такий вигляд:

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \dots + \beta_p x_i^p + \varepsilon_i. \quad (4.59)$$

Як бачимо, модель (1) має тільки одну незалежну змінну (x) та одну залежну змінну (y). Але те, що незалежна змінна входить у різних степенях, переводить модель (1) у клас багатофакторних регресійних моделей, а саме, вводячи заміну змінних $x^p = Z_p$, отримаємо:

$$y_i = \beta_0 + \beta_1 z_{1i} + \beta_2 z_{2i} + \dots + \beta_p z_{pi} + \varepsilon_i. \quad (4.60)$$

Отже, p -ступенева поліноміальна модель лінійна за своїми параметрами, які можуть бути знайдені методом найменших квадратів.

У мікроекономіці найпоширенішим є застосування поліноміальних моделей для оцінки загальних, середніх та граничних витрат у виробництві продукції.

Економетрична модель загальних витрат

Як приклад поліноміальної регресійної моделі розглянемо зв'язок між випуском продукції та загальними витратами на її виробництво (табл. 4.6).

Якою буде форма такого зв'язку? Щоб відповісти на це запитання, вдаємося до графічного зображення даних (мал. 4.4.)

Як бачимо з малюнка, залежність між загальними витратами та обсягами виробництва нагадує S-криву. Така S-крива загальних витрат може відповідати такій узагальненій регресійній моделі:

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \varepsilon_i, \quad (4.61)$$

де y — загальні витрати; x — обсяг виробництва. Невідомі параметри моделі (4.61) можуть бути знайдені методом найменших квадратів. Але економічна теорія накладає деякі умови на параметри кубічної функції загальних витрат. Які саме?

Щоб відповісти на це запитання, необхідно згадати залежність між загальними витратами, середніми та граничними витратами. Мікроекономічна теорія показує, що крива середніх та граничних витрат має U-подібну форму. Коли випуск зростає, то обидві криві спочатку спадають, а потім зростають. Крім того, ці криві обов'язково перетинаються, причому до точки перетину значення середніх витрат вище, ніж граничних; а післяграничні витрати вищі за середні.

Справді, це можна побачити з виразу середніх та граничних витрат через загальні витрати (TC): $TC = y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \varepsilon$:

$$\text{середні витрати (AC): } AC = \frac{TC}{x} = y = \frac{\beta_0}{x} + (\beta_1 + \beta_2 x + \beta_3 x^2) + \varepsilon;$$

$$\text{граничні витрати (MC): } MC = \frac{d(TC)}{dx} = \beta_1 + 2\beta_2 x + 3\beta_3 x^2 + \varepsilon.$$

Як бачимо, середні та граничні витрати справді мають U-подібну форму.

Крім того, щоб були правильними класичні мікроекономічні співвідношення між усіма типами витрат, необхідно, щоб оцінки параметрів моделі (4.61) відповідали таким співвідношенням.

$$1. \beta_0, \beta_1, \beta_3 > 0.$$

$$2. \beta_2 < 0.$$

$$3. \beta_2^2 < 3\beta_1\beta_3.$$

Ці співвідношення можуть бути корисними, коли ми досліджуємо вибірку модель.

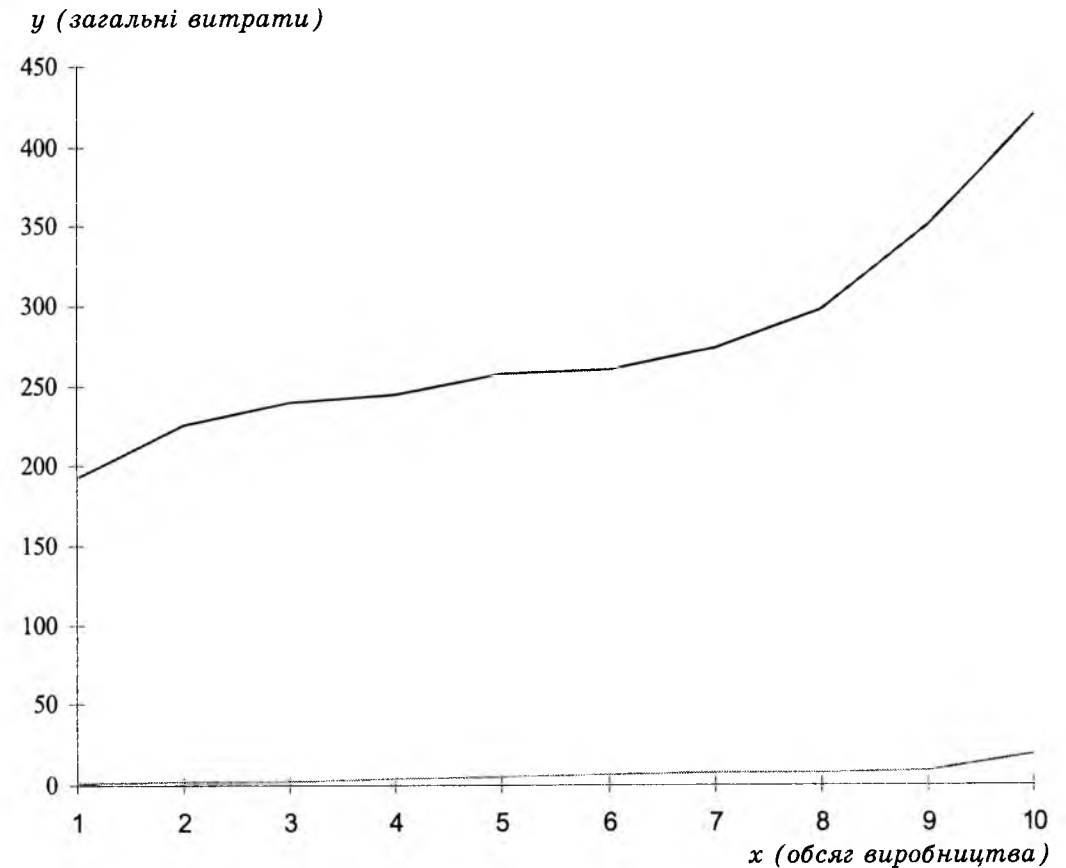
Проведемо розрахунки за моделлю (4.61) за табл.4.3. Отримаємо:

$$\hat{y}_i = 141,776 + 63,48x_i - 12,96x_i^2 + 0,94x_i^3; \quad (4.62)$$

$$R^2 = 0,9983.$$

Таблиця 4.6

Загальні витрати, млн.крб. y	Обсяг виробництва x
193	1
226	2
240	3
244	4
257	5
260	6
274	7
297	8
350	9
420	10



Малюнок 4.4. Залежність між загальними витратами та обсягом виробництва

Приклад 2. Виробнича функція Кобба — Дугласа

Відома виробнича функція Кобба — Дугласа має такий вигляд:

$$y = \beta_3 x_1^{\beta_1} x_2^{\beta_2} e^{\varepsilon}, \quad (4.63)$$

де y — випуск продукції;

x_1 — витрати праці;

x_2 — витрати капіталу;

ε — випадкова величина;

e — основа натурального логарифма.

Модель (4.63) є нелінійною, але, як уже розглядалося раніше, шляхом логарифмування її легко звести до лінійної регресійної моделі:

$$\ln y_i = \beta_0 + \beta_1 \ln x_{i1} + \beta_2 \ln x_{i2} + \dots + \beta_p \ln x_{ip} + \varepsilon_i. \quad (4.64)$$

Модель (4.64) лінійна за своїми параметрами і має дві незалежні змінні, отже, вона відноситься до класу багатофакторних регресійних моделей.

Властивості функції Кобба — Дугласа широко відомі.

1. Параметр β_1 є еластичністю випуску від затрат праці. Він показує, на скільки відсотків зміняться витрати, коли затрати праці змінюються на один відсоток за припущення, що витрати капіталу залишаються незмінними.

2. Параметр β_2 є коефіцієнтом еластичності випуску від витрат капіталу за припущення, що витрати праці залишаються незмінними.

3. Сума параметрів $(\beta_1 + \beta_2)$ описує масштаб виробництва.

Якщо ця сума дорівнює 1, то маємо постійний масштаб виробництва; коли він менше 1, то маємо спадний масштаб виробництва, за якого зростання значень факторів у певну кількість разів призводить до спаду випуску виробництва. Коли $(\beta_1 + \beta_2) > 1$, навпаки, спостерігається зростаючий масштаб виробництва; збільшення значень факторів, наприклад у α разів, спричиняє збільшення випуску виробництва більше ніж у α разів.

Розглянемо конкретний приклад.

Є дані про деревообробний сектор України (табл.4.7).

Таблиця 4.7

Рік	Обсяг продукції, млн. ум. грош. од. y	Витрати праці, млн.днів x_1	Витрати капіталу, млн.ум. грош. од. x_2
1972	12767.5	375.2	131427
1973	16347.1	402.5	134267
1974	19542.7	478	139038
1975	21075.9	553.4	146450
1976	23052.0	616.7	153714
1977	26128.2	695.7	164783
1978	29563.7	790.3	176864
1979	33376.6	816.0	188146
1980	38354.3	848.8	205841
1981	46868.3	873.1	221748
1982	54308.0	999.2	239715

Побудуємо функцію Кобба — Дугласа, яка описує залежність між випуском продукції деревообробної галузі (млн. ум. грош. од.), затратами праці (млн. днів) та витратами капіталу (млн. ум. грош. од.). Розрахувавши її параметри методом найменших квадратів, отримаємо:

$$\ln \hat{y}_i = -3.338 + 1.498 \ln x_1 + 0.489 \ln x_2; \quad (2.449) \quad (0.539) \quad (0.102) \quad (4.65)$$

$$R^2 = 0.889; \quad \bar{R}^2 = 0.87.$$

З моделі (4.65) бачимо, що у деревообробному секторі коефіцієнти еластичності випуску продукції від праці та капіталу відповідно дорівнювали

1.498 та 0.489. Тобто, коли витрати праці за незмінного значення капіталу збільшувались на 1 %, випуск продукції збільшувався на 1.498 %. Збільшення на 1 % витрат капіталу, коли витрати праці залишаються незмінними, спричиняє збільшення на 0.5 % випуску. Сума параметрів $(\beta_1 + \beta_2) = 1.987$ свідчить про розширення масштабів деревообробного виробництва.

Моделювання процесу малої приватизації в Україні

Приватизація в Україні викликає велику зацікавленість економістів, практиків, фахівців різних напрямків, іноземних інвесторів та всього населення України. Це те питання, яке завжди є індикатором змін у суспільстві, а для таких постсоціалістичних країн, як Україна, ще й питанням життєздатності реформ та процесів формування ринкової економіки. Спробуємо розглянути питання приватизації лише під вузьким кутом, а саме: дослідити вплив певних факторів на процес малої приватизації в Україні за допомогою багатофакторної економетричної моделі. Підкреслимо, що наведений матеріал не є науковим дослідженням. Він розроблений у процесі написання творчих робіт студентами НаУКМА і має сприйматись як ілюстрація до практичного застосування економетричних моделей у різноманітних економічних дослідженнях.

Процес приватизації в Україні навіть сьогодні має певні особливості, на його успішність впливає велика кількість факторів, деякі з них носять суто суб'єктивний характер і тому не можуть бути виміряні кількісно. Про деякі з них важко знайти інформацію, тому відберемо невелику кількість факторів, які, на нашу думку, мали б впливати на процес малої приватизації в Україні.

По-перше, це кількість громадян, що скористалися приватизаційними сертифікатами (x_1). Вважаємо, що доречно брати саме кількість тих громадян, які скористалися своїм правом на частку державного майна, бо кількість отриманих сертифікатів не повністю відображає попит на державне майно, яке приватизується. Таким чином, відпадає потреба враховувати відсоток і середню швидкість використання приватизаційних сертифікатів, що взагалі полегшує дослідження впливу цього фактора.

По-друге, це зважена середня позичкова ставка комерційних банків (x_2). Приватизовані підприємства дуже часто потребують нових капіталовкладень, тому доступність кредитів має суттєво впливати на попит об'єктів приватизації.

По-третє, кредити в валюті (x_3). Слід визначити, що сьогодні немає, як не дивно, залежності між процентною ставкою в Україні і закордонними кредитами.

Четвертим фактором, який, на наш погляд, впливає на процес малої приватизації, є індекс реальної продукції (за 100 взятий рівень 1990 року) (x_4). При цьому ми виходимо з того, що спад виробництва має наводити людей на думку, що негайно треба щось змінити, і, таким чином, прискорювати процес приватизації.

П'ятим відібраним фактором є інфляція оптових цін (x_5). Ми вибрали саме інфляцію оптових цін, бо вважаємо, що вона має впливати на вартість виробництва, і, тим самим, на привабливість придбання того чи іншого підприєм-

ства.

Тепер ми маємо дослідити вплив відібраних факторів на залежну змінну y — кількість приватизованих підприємств.

Підкреслимо, що при побудові багатофакторної регресійної моделі ми обмежимося лише дослідженням малої приватизації, бо, якщо не брати до уваги приватизацію житла, саме у цій галузі відбувається найбільше подій, і вони піддаються певному аналізу.

Дані, які використовувались для розрахунку моделі, наведені в табл. 4.8. Вони були зібрані за період 1993 (за винятком квітня), 1994 та 6 місяців 1995

Таблиця 4.8

Дата	Кількість приватизов. підприємств y	Кількість громадян, тис. чол. x_1	Позичкова ставка x_2	Кредити у валюті, тис. дол. x_3	Індекс реальної продукції x_4	ІОЦ x_5
01/93	30	5	6.59		74.7	118.2
02/93	140	65	7.35		75.0	20.8
03/93	430	115	7.26		65.4	11.6
05/93	570	176	11.88		62.1	52.2
06/93	830	275	19.25		59.3	90.1
07/93	1165	335	18.50		56.3	31.0
08/93	1530	370	16.88		54.0	33.0
09/93	1685	395	18.79	980	53.6	76.5
10/93	1728	425	20.49	1065	51.8	34.2
11/93	2825	455	23.27	1200	52.0	33.1
12/93	3585	728	24.59	1315	47.8	76.9
01/94	4300	1326	25.85	1447	40.9	23.8
02/94	4967	1340	32.06	1755	37.8	16.0
03/94	5442	1503	32.48	2124	39.1	6.9
04/94	6438	1980	27.33	2513	37.6	6.0
05/94	7557	2200	25.26	2827	38.2	2.3
06/94	8402	2873	21.23	3387	39.1	2.7
07/94	8910	3100	16.04	3920	38.1	4.1
07/94	9450	3900	12.61	4842	38.0	6.5
08/94	10214	4883	11.80	5717	41.5	10.8
09/94	10910	4902	11.78	6802	43.7	17.2
10/94	11100	5804	16.78	8210	39.8	101.0
11/94	11552	7106	17.06	8176	36.3	32.7
12/94	12375	8100	18.39	9043	33.0	29.2
01/95	12450	11186	17.74	10232	32.0	11.4
02/95	12802	12616	15.80	10556	35.3	9.3
03/95	12975	13180	12.73	11738	32.2	5.1
04/95	13100	14800	10.18	12739	31.6	7.1
05/95	14957	15600	7.66	14272	31.5	8.6

Таблиця 4.9

Незалежна змінна	Коефіцієнт	Стандартна помилка	t -статистика	Рівень значимості
Const.	15744.242360	3837.223341	4.1030	0.0008
x_1	- 0.883249	0.200383	- 4.4078	0.0004
x_2	- 97.491417	49.980252	- 1.9506	0.0688
x_3	1.515461	0.297273	5.0979	0.0001
x_4	- 223.031321	63.132450	- 3.5328	0.0028
x_5	- 15.444591	9.177465	- 1.6829	0.1118

року. Зібрати повну інформацію за пізніший період, на жаль, не вдалося.

Статистичні дані взято із “Звіту Фонду державного майна за 1994 рік”, “Тенденції української економіки” — бюлетеня Європейського Центру макроекономічного аналізу України та економічного бюлетеня “Україна в числах”.

У випадках, коли певні дані не збігалися, брались не офіційні дані, а дані Євроцентру як більш коректні. Деякі дані взято з графіків (приватизація за 1993 рік; кількість громадян, що використали приватизаційні сертифікати), а оскільки індекс реальної продукції та дані про використання сертифікатів за 1994/95 роки були надані поквартально, то повна точність даних не збереглася, хоча загальні тенденції відображені повністю.

Було розраховано багатофакторну регресійну модель малої приватизації з використанням пакету STATGRAPH та отримано результати, які наведено в табл. 4.9.

Модель має вигляд:

$$\hat{y} = 15744.2 - 0.883x_1 - 97.49x_2 + 1.52x_3 - 223.03x_4 - 15.44x_5.$$

Оцінений коефіцієнт детермінації дорівнює 96,76%. $DW=1.73234$.

Перевіримо статистичну значимість параметрів регресії. При заданому рівні значимості 5% (табл. 4.9) можна побачити, що параметри при факторах x_2 та x_5 є статистично незначимими, тому ці фактори можна з моделі виключити. Перевірка за тестом Дарбіна — Уотсона свідчить про відсутність кореляції між залишками.

Розглянемо ANOVA-таблицю (табл. 4.10).

Таблиця 4.10

Таблиця ANOVA-аналізу

Джерело варіації	Сума квадратів	DF	MS	F-Ratio
Модель	342441148	5	68488230	
Помилка	8657129	16	541071	126579
Загальне	351098277	21		

За F -критерієм Фішера визначаємо, що модель адекватна.

Вилучимо з моделі фактори x_2 та x_5 , коефіцієнти при яких були статис-

тично незначимими, та знову проведемо розрахунки.

Нова модель має вигляд:

$$y = -0.082426x_1 + 1.59262x_3 - 215.471x_4.$$

Слід зазначити, що такі ж результати ми б отримали, якби відразу використали метод покрокової регресії. При покроковому виборі факторів, фактори x_1 , x_3 та x_4 входять у модель ($F\text{-часткові} > F_{кр}$, $F_{кр} = 2.68$).

Спробуємо проінтерпретувати за остаточною моделлю вплив факторів на приватизацію малих підприємств. Дуже цікавий зв'язок спостерігається щодо приватизаційних паперів. Незвичайну зворотну залежність можна пояснити загальним зменшенням об'єктів приватизації та штучними урядовими обмеженнями. Адміністративні заходи для виправлення цієї ситуації було вжито з прийняттям Указу Президента України "Про заходи щодо прискорення процесу малої приватизації в Україні" від 30 грудня 1994 року, нової редакції Закону про оренду (особливо у ст. 25) та інших документів, але, нажаль, це не дає бажаних результатів. У січні 1996 року розрив між обсягом "паперового" попиту та "майнової" пропозиції становив 1061 трлн. крб. Таким чином, при збільшенні ціни об'єктів приватизації на кожен такий об'єкт витрачається більша кількість приватизаційних паперів.

Зрозумілим є прямий зв'язок між кредитами та приватизацією: більше кредитів — більше інвестицій, у тому числі в об'єкти приватизації.

При загальному спаді виробництва та зменшенні індексу загального виробництва зростає зацікавленість у приватизації. Налагоджене ефективне виробництво на приватизованих підприємствах на фоні загального спаду буде конкурентоздатним і приносить значний прибуток.

Звичайно, розроблена модель не є досить вдалою, тому що не відображає впливу суб'єктивних факторів, які в процесі приватизації відіграють визначальну роль, наприклад, настроїв та політики уряду (простим прикладом цього може слугувати адміністративне регулювання попиту та пропозиції на приватизацію.)

Але навіть на основі такої спрощеної моделі можна аргументувати деякі пропозиції щодо поліпшення процесу приватизації в Україні (звичайно, це стосується ситуації 1996 року):

1. Треба зняти мораторій на приватизацію тих підприємств, на які Верховною Радою було накладено вето. Це прискорить "паперову" приватизацію, а разом із застосуванням приватизаційних аукціонів ціна об'єкта визначатиметься у більшій відповідності до ринкових законів, що не раз намагалися зробити у Фонді державного майна України.

2. Необхідно визначити роль іноземних партнерів і залучати іноземних інвесторів, хоча, як показують деякі дослідження (зокрема проведені у виданнях "Фінансова Україна", "Фінансовий Київ", "Економіка України"), вже з'являються внутрішні інвестори, які здатні на масштабні капіталовкладення. Разом з тим не треба нехтувати й інституційними інвесторами (інвестами, трастами та ін.), які накопичили гігантську кількість приватизаційних па-

перів, але по-справжньому ще не почали їх вкладати.

3. Необхідно чіткіше визначити власника, який буде управляти підприємством.

4. Чисто адміністративними заходами можна домогтися лише кількісних результатів у приватизації, а не якісних. Приватизація не може іти окремо або у відповідності з кардинальними реформами економіки в цілому. Тому важливим залишається загальний процес реформування в Україні.

Аналіз впливу соціальних факторів на рівень злочинності в Україні

Проаналізуємо, як впливали різні соціальні фактори, а саме: забезпеченість житлом, рівень освіти, продаж алкогольних напоїв та інші — на кількість зареєстрованих в Україні злочинів протягом останнього десятиріччя. Для цього побудуємо багатофакторну регресійну модель, використовуючи дані, наведені в табл. 4.11. Залежною змінною є кількість зареєстрованих злочинів (y). Для остаточного аналізу обрано такі фактори: загальна кількість населення України, кількість міського населення України, забезпеченість населення житлом, продаж алкогольних напоїв у розрахунку на душу населення та кількість людей, що мають вищу та середню освіту, в розрахунку на 1000 осіб.

Таблиця 4.11

Рік	Кількість зареєстрованих злочинів, тис. y	Кількість населення, млн. осіб x_1	Забезпеченість житлом, м ² на особу x_2	Рівень освіти x_3	Кількість міського населення, млн. осіб x_4	Продаж алкогольних напоїв, л на особу x_5
1984	250.0	50.7	16.0	877	33.2	5.3
1985	249.6	50.8	16.3	884	31.4	5.2
1986	248.7	51.0	16.8	894	33.7	3.4
1987	237.8	51.2	17.1	903	34.2	2.9
1988	243.0	51.4	17.3	912	34.7	3.2
1989	322.3	51.7	17.6	928	34.6	3.8
1990	369.8	51.8	17.8	939	34.8	4.1
1991	405.5	51.9	18.0	951	35.1	4.1
1992	480.5	52.1	18.2	962	35.3	3.6
1993	539.3	52.2	18.5	974	35.4	3.3

Джерело: "Статистичний щорічник України" (1994)

У формалізованому записі модель має такий вигляд:

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \varepsilon.$$

Розрахунки невідомих параметрів моделі здійснювались за допомогою пакета STATGRAPH. Для побудови моделі (без константи) використовувався метод покрової регресії.

Кінцева модель містить тільки три з відібраних п'яти факторів: кількість населення України, забезпеченість житлом та освіченість:

$$y = -91.89x_1 - 189.38x_2 + 9.06x_3.$$

Відповідно до моделі коефіцієнт детермінації та оцінений коефіцієнт детермінації дорівнюють: $R^2 = 0.9984$, $\bar{R}^2 = 0.9980$. Це свідчить про те, що зміни значення залежної змінної великою мірою пояснюються саме змінами у відібраних факторах.

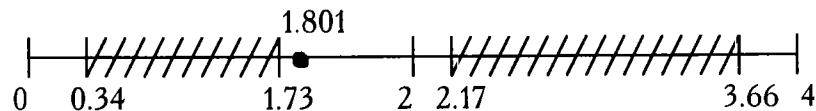
Перевіримо розроблену модель на адекватність за F -критерієм Фішера. Розрахункове F -відношення для даної регресії становить 1480.97. Задамо рівень значимості 5% та за таблицею F -розподілу Фішера при заданому рівні значимості і ступенях вільності відповідно 3 (нагадаємо, що перший ступінь вільності дорівнює кількості факторів у моделі) та 7 (другий ступінь вільності дорівнює кількості спостережень мінус кількість оцінених параметрів: $10 - 3 = 7$) знайдемо критичне значення, яке дорівнює: $F_{3,7,95\%} = 4.35$. Як бачимо, розраховане F -відношення значно перевищує $F_{кр}$, тобто модель є адекватною.

Перевіримо розраховані параметри на значимість. Отримані за допомогою пакета розрахункові значення t -статистики для кожного параметра відповідно дорівнюють $t_1 = -15.77$; $t_2 = -5.59$; $t_3 = 10.09$. Тепер за таблицею t -розподілу Ст'юдента, задавши рівень значимості 5%, знайдемо відповідні критичні значення при ступенях вільності 7 (нагадаємо, що в даному випадку ступінь вільності дорівнюють кількості спостережень мінус кількість оцінених параметрів). За таблицею $t_{кр} = 2.365$. Отже, всі параметри статистично значимі.

Для розробленої моделі справджується припущення про відсутність лінійного зв'язку між факторами, тобто відібрані фактори не є мультиколінеарними.

Перевіримо припущення про відсутність залежності між випадковими величинами за критерієм Дарбіна—Уотсона. Розрахована статистика Дарбіна—Уотсона дорівнює: $DW = 1.801$. За таблицею Дарбіна—Уотсона при рівні значимості 5% та відповідно кількості факторів $p=3$ і кількості спостережень $n = 10$ знаходимо два критичних значення: $d_L = 0.34$, $d_U = 1.73$.

Будуємо відповідні зони; зона значень від 0 до 0.34 є зоною позитивної кореляції, зони від 0.34 до 1.73 та від 2.17 та 3.66 є зонами невизначеності; зона від 1.73 до 2.17 є зоною відсутності кореляції між випадковими величинами, зона від 3.66 до 4 є зоною негативної кореляції (див. мал. 4.5).



Малюнок 4.5. Перевірка залежності між випадковими величинами за критерієм Дарбіна—Уотсона

Як бачимо з мал.4.5, розраховане значення статистики DW потрапляє в зону відсутності кореляції між залишками, що свідчить про виконання припущення незалежності випадкових величин.

Використовуючи побудовану модель, спробуємо спрогнозувати кількість злочинів у 1994 році. Розглядаючи кожен з факторів як лінійно залежний від часу ($x = a_0 + a_1 t$), розрахуємо прогнозне значення кожного фактора на 1994 р. Отримаємо відповідно:

$$\begin{array}{ll} x_1 = 50.51 + 0.18t; & \text{Прогноз: } x_1 = 52.45; \\ x_2 = 15.89 + 0.27t; & x_2 = 18.83; \\ x_3 = 861.48 + 11.08t; & x_3 = 983.33. \end{array}$$

Задавши прогнозні значення факторів у моделі за 1994 р., отримуємо прогнозне значення для y , а саме: за даною моделлю кількість зареєстрованих в Україні злочинів за 1994 р. мала становити 518 тис., причому

$$P(470.38 < y_{n+1} < 565.57) = 0.95.$$

Зауважимо, що створена модель є лише ілюстративною. Її основне призначення полягало у відображенні принципу побудови багатофакторних регресійних моделей на практиці та аналізі отриманих результатів. Звичайно, модель повністю не відображає реальної ситуації в країні. На злочинність в Україні впливає значно більше факторів, ніж зазначено у даній моделі, усі їх врахувати неможливо через ряд причин. Деякі з них взагалі не вимірюються статистично, а частина не вимірювалась на початку досліджуваного періоду.

Як відомо, якість багатофакторної регресійної моделі залежить також і від різниці між кількістю спостережень та кількістю врахованих факторів. У нашому випадку збільшити кількість спостережень неможливо через відсутність статистичних даних про деякі фактори в Україні за роки, що передували 1984 р.

РОЗДІЛ 5. ОСОБЛИВІ ВИПАДКИ У БАГАТОФАКТОРНОМУ РЕГРЕСІЙНОМУ АНАЛІЗІ

5.1. МУЛЬТИКОЛІНЕАРНІСТЬ ТА ЇЇ НАСЛІДКИ. ЗНАХОДЖЕННЯ МУЛЬТИКОЛІНЕАРНОСТІ ТА ЇЇ ВИЛУЧЕННЯ

5.1.1. Визначення мультиколінеарності та її природа

У розділі 4 ми вже стикалися з проблемою мультиколінеарності. Тепер розглянемо її детальніше. Перш за все потрібно зрозуміти природу мультиколінеарності та з'ясувати, чи справді наявність мультиколінеарності є проблемою при розробці моделей. І якщо це так, то в чому саме вона полягає і як її подлати? Насамперед згадаємо, що таке мультиколінеарність.

Термін “мультиколінеарність” означає, що в багатофакторній регресійній моделі дві або більше незалежних змінних (факторів) пов'язані між собою лінійною залежністю або, іншими словами, мають високий ступінь кореляції ($r_{x_i x_j} \rightarrow 1, i \neq j$).

Наприклад, мультиколінеарність може бути проблемою, коли ми вивчаємо залежність між ціною акції, дивідендами на акцію та заробленим прибутком на акцію, оскільки дивіденди та зароблений прибуток на одну акцію мають високий ступінь кореляції.

Крім того, у розділі 4 ми вже показували, що якщо два колінеарні фактори змінюються в одному напрямі, то майже неможливо оцінити окремий вплив кожного з них на досліджуваний показник y .

Мультиколінеарність може виникати за різних умов.

По-перше, є глобальна тенденція одночасної зміни економічних показників. На макроекономічні показники впливають однакові фактори. Це приводить до того, що вони відображають широкий спектр моделей однакової економічної ситуації. Наприклад, у період бумів або швидкого економічного зростання базові економічні показники також зростають, звичайно, з деяким лагом. Такі показники, як доход, споживання, накопичення, інвестиції, ціни, зайнятість мають тенденцію до зростання в період економічної експансії і до спаду в період рецесії. Сама наявність трендів у динамічних рядах є причиною мультиколінеарності.

По-друге, широке використання в економетричних моделях лагових значень однієї змінної також призводить до виникнення мультиколінеарності. Наприклад, добре відомі інвестиційні функції, в яких лагові значення минулого рівня економічної активності вводяться як окремі змінні. У

функціях споживання витрати на споживання у попередньому періоді вводяться в модель поряд з величиною поточного рівня доходу.

Таким чином, зрозуміло, що в економіці взагалі важко уникнути певного рівня залежності між показниками, тобто певного рівня колінеарності.

До яких же наслідків може привести мультиколінеарність? Це одне з найважливіших питань, яке потрібно зрозуміти при розробці економетричних моделей. Перш за все потрібно відрізнити досконалу кореляцію від недосконалої.

Спочатку мультиколінеарність розглядали як наявність досконалої кореляції ($r_{x_i x_j} = 1$). Сам термін “мультиколінеарність” вперше було введено Р.Фрішем¹.

Якщо ми розглянемо багатофакторну модель з p факторами, то досконалий лінійний зв'язок між x_1, x_2, \dots, x_p має місце, коли виконується така умова:

$$b_1 x_1 + b_2 x_2 + \dots + b_p x_p = 0, \quad (5.1)$$

де не всі $b_i (i = 1, \overline{p})$ одночасно дорівнюють нулю.

Проте сьогодні “мультиколінеарність” має ширше значення і включає випадок не тільки досконалої кореляції, а й випадок, коли фактори $x_i, i = 1, \overline{p}$ корелюють між собою (проте недосконало). Тобто є залежність:

$$b_1 x_1 + b_2 x_2 + \dots + b_p x_p + e = 0, \quad (5.2)$$

де e — випадкова величина (помилка).

Різницю між досконалою і недосконалою мультиколінеарністю розглянемо для випадку, коли, наприклад, $b_2 \neq 0$.

Вираз (5.1) можна тоді переписати таким чином:

$$x_{2i} = -\frac{b_1}{b_2} x_{1i} - \frac{b_3}{b_2} x_{3i} - \dots - \frac{b_p}{b_2} x_{pi} \quad (i = 1, n). \quad (5.3)$$

Фактор x_2 є лінійною комбінацією інших факторів.

Аналогічно, вираз (5.2) можна записати:

$$x_{2i} = -\frac{b_1}{b_2} x_{1i} - \frac{b_3}{b_4} x_{3i} - \dots - \frac{b_p}{b_2} x_{pi} - \frac{1}{b_2} e_2. \quad (5.4)$$

¹ Frisch R. Statistical Confluence Analysis by Means of Economics. — Oslo University, 1934.

Із (5.4) зрозуміло, що змінна x_2 не є точною лінійною комбінацією інших факторів, бо залежить також від випадкової величини e_1 .

Чому для класичної лінійної моделі вимогою одного з припущень є відсутність мультиколінеарності між її факторами? Тому, що у випадку досконалої мультиколінеарності параметри регресії стають невизначеними, а їхні середні квадратичні відхилення прямують до нескінченності.

Покажемо це для двофакторної регресійної моделі.

Властивість 1. Припустимо, що ми розглядаємо залежність національного доходу (y) від випуску промислових товарів (x_1) і товарів тривалого користування (x_2) за період часу t .

Нехай економетрична модель має такий вигляд:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon, \quad (5.5)$$

де x_1 і x_2 пов'язані між собою залежністю $x_2 = \alpha x_1$.

Формули для оцінок параметрів мають такий вигляд (див. додаток 5.1):

$$b_1 = \frac{\sum_{i=1}^n ((x_{1i} - \bar{x}_1)(y_i - \bar{y})) \sum_{i=1}^n (x_{2i} - \bar{x}_2)^2 - \sum_{i=1}^n ((x_{2i} - \bar{x}_2)(y_i - \bar{y})) \sum_{i=1}^n (x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2)}{\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2 \sum_{i=1}^n (x_{2i} - \bar{x}_2)^2 - (\sum_{i=1}^n (x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2))^2}; \quad (5.6)$$

$$b_2 = \frac{\sum_{i=1}^n ((x_{2i} - \bar{x}_2)(y_i - \bar{y})) \sum_{i=1}^n (x_{1i} - \bar{x}_1)^2 - \sum_{i=1}^n ((x_{1i} - \bar{x}_1)(y_i - \bar{y})) \sum_{i=1}^n (x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2)}{\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2 \sum_{i=1}^n (x_{2i} - \bar{x}_2)^2 - (\sum_{i=1}^n (x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2))^2}; \quad (5.7)$$

Замінивши x_2 на αx_1 , отримаємо:

$$b_1 = \frac{a^2 (\sum_{i=1}^n (x_{1i} - \bar{x}_1)(y_i - \bar{y})) (\sum_{i=1}^n (x_{1i} - \bar{x}_1))^2 - a^2 (\sum_{i=1}^n (x_{1i} - \bar{x}_1)(y_i - \bar{y})) (\sum_{i=1}^n (x_{1i} - \bar{x}_1))^2}{a^2 (\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2)^2 - a^2 (\sum_{i=1}^n (x_{1i} - \bar{x}_1))^2} = \frac{0}{0}; \quad (5.8)$$

$$b_2 = \frac{a \sum_{i=1}^n (x_{1i} - \bar{x}_1)(y_i - \bar{y}) \sum_{i=1}^n (x_{1i} - \bar{x}_1)^2 - a (\sum_{i=1}^n (x_{1i} - \bar{x}_1)(y_i - \bar{y})) \sum_{i=1}^n (x_{1i} - \bar{x}_1)^2}{a^2 (\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2)^2 - a^2 (\sum_{i=1}^n (x_{1i} - \bar{x}_1))^2} = \frac{0}{0}; \quad (5.9)$$

Як бачимо з виразів (5.8) і (5.9), параметри b_1 і b_2 у разі мультиколінеарності стають невизначеними, тобто неможливо знайти окреме значення кожного з них.

Властивість 2. Якщо $r_{x_1 x_2} = 1$, то середні квадратичні відхилення параметрів прямують до нескінченності.

Розглянемо цю властивість на прикладі двофакторної регресійної моделі

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon,$$

де $x_2 = kx_1$ і $r_{x_1 x_2} = 1$.

Можна показати, що дисперсії оцінок параметрів дорівнюють (див. додаток 5.1):

$$\text{var}(b_1) = \frac{\sigma_\varepsilon^2}{(1 - r^2) \sum_{i=1}^n (x_{1i} - \bar{x}_1)^2}; \quad (5.10)$$

$$\text{var}(b_2) = \frac{\sigma_\varepsilon^2}{(1 - r^2) \sum_{i=1}^n (x_{2i} - \bar{x}_2)^2}. \quad (5.11)$$

Коли $r = 1$, тоді $\text{var}(b_1) = \infty$; $\text{var}(b_2) = \infty$.

Відповідно дорівнюють нескінченності і середні квадратичні відхилення.

Звичайно, досконала мультиколінеарність є дуже рідкісним явищем; частіше в економічних дослідженнях немає точної лінійної залежності між параметрами. Як же оцінювати параметри багатофакторної регресії за умови високої, але не досконалої мультиколінеарності? Розглянемо приклад. Нехай для двофакторної регресійної моделі маємо таку залежність між факторами:

$$x_2 = \alpha x_1 + e_1, \quad (5.12)$$

де $\alpha \neq 0$; e_1 — випадкова величина; $\sum_{i=1}^n (x_{1i} - \bar{x}_1) e_i = 0$.

У цьому випадку можна оцінити параметри b_1 і b_2 . Підставивши (5.12) у (5.6), отримаємо:

$$b_1 = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_{1i} - \bar{x}_1) (a^2 \sum_{i=1}^n (x_{1i} - \bar{x}_1)^2 + \sum_{i=1}^n e_i^2) - (a \sum_{i=1}^n (y_i - \bar{y})(x_{1i} - \bar{x}_1) + \sum_{i=1}^n ((y_i - \bar{y}) e_i) (a \sum_{i=1}^n (x_{1i} - \bar{x}_1)^2)) \sum_{i=1}^n (x_{1i} - \bar{x}_1)^2}{\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2 (a^2 (\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2 + \sum_{i=1}^n e_i^2) - (a \sum_{i=1}^n (x_{1i} - \bar{x}_1))^2)} \neq \frac{0}{0}, \quad (5.14)$$

виходячи з того, що $\sum_{i=1}^n (x_{1i} - \bar{x}_1) e_i = 0$.

Аналогічно можна отримати вираз для параметра b_2 . Тепер немає підстави вважати, що параметри неможливо оцінити. Звичайно, якщо e_i незначно відрізняється від нуля, тоді ми повертаємось до випадку “досконалої” колінеарності.

5.1.2. Теоретичні наслідки мультиколінеарності в загальному випадку

Згадаємо, що якщо умови класичної моделі задовольняються, то МНК-оцінки (оцінки, обчислені за методом найменших квадратів) параметрів регресії є BLUE-оцінками. Навіть при дуже високій, але недосконалій мультиколінеарності МНК-оцінки все ще зберігають властивість BLUE-оцінок.

Тоді чому виникає питання про мультиколінеарність? Звичайно, при побудові моделі навіть в умовах мультиколінеарності можна отримати незміщені та стійкі оцінки. Ефект мультиколінеарності виявляється у тому, що складно отримати значення параметрів з малою стандартною помилкою. Такий самий ефект спостерігається при невеликій кількості спостережень або при невеликій зміні значень. З точки зору теорії, мультиколінеарність та невелика кількість змінних — це одна і та сама проблема. Тому теоретично питання “Що робити з мультиколінеарністю?” збігається з питанням “Що робити, якщо в мене мало спостережень?” Але фактично проблема мультиколінеарності так часто виникає в емпіричних дослідженнях та створює такі серйозні проблеми оцінювання, що розглядати її тільки як порушення припущення моделі класичної лінійної регресії неможливо.

По-перше, справедливе твердження, що навіть у разі високої мультиколінеарності МНК-оцінки є незміщеними. Але незміщеність — це властивість часто повторюваної вибірки. Вона означає, що при постійних значеннях змінних величин x , якщо досліджувати повторювані вибірки та підраховувати МНК-оцінки для кожної з них, середнє значення змінних у цих вибірках прямує до правильних значень параметрів у будь-якій з наведених вибірок.

По-друге, справедливим є також те, що колінеарність не порушує властивостей мінімуму дисперсії: в класі лінійних незміщених оцінок МНК-оцінки мають мінімальну дисперсію, і тому вони ефективні. Однак це не значить, що дисперсія МНК-оцінки буде неминуче малою (відносно значення параметра) в будь-якій з наведених вибірок. Цю властивість ми проілюструємо дещо пізніше.

По-третє, мультиколінеарність — явище виключно регресійного аналізу вибірки в тому розумінні, що навіть якщо змінні x пов’язані в генеральній сукупності нелінійно, вони можуть мати лінійний зв’язок у кожному окремому випадку: коли ми постулюємо вибіркочну або узагальнену функцію регресії, то впевнені, що всі змінні величини x , які утворюють

модель, мають окремий або незалежний вплив на залежну змінну величину y . Однак може так трапитися, що в будь-якій з вибірок, яка використовується для перевірки узагальненої моделі, деякі або всі змінні величини x настільки висококолінеарні, що ми не можемо виявити іншого індивідуального впливу на y . Наша вибірка, так би мовити, нас підводить, хоча за теорією всі x важливі. Іншими словами, вже неможливо застосувати всі змінні величини x в аналізі.

Проілюструємо це прикладом “споживання — прибуток”. Економісти вважають, що важливим фактором у споживанні, окрім прибутку, є також заможність споживача. Таким чином, ми можемо записати:

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon_i,$$

де y_i — споживання, x_1 — доход, x_2 — багатство.

Може так трапитися, що коли ми отримуємо дані про прибуток та багатство, дві змінні величини x_1 та x_2 будуть високо або навіть повністю корельовані між собою. Більш заможні люди здебільшого прагнуть отримувати вищий прибуток. Таким чином, хоча теоретично прибуток і багатство є логічними кандидатами для з’ясування поведінки споживача, на практиці (тобто у вибірці) важко розібратися в окремих випадках впливу прибутку і багатства на споживання. В ідеальному випадку, щоб визначити індивідуальний вплив багатства і прибутку на споживання, нам потрібна достатня кількість вибірок зі спостереженнями щодо заможних індивідів з низьким прибутком та не дуже заможних людей з високим прибутком. Хоча це можливо теоретично, практично цього дуже важко добитися при проведенні загальних досліджень.

У зв’язку з цим той факт, що МНК-оцінки є BLUE-оцінками, незважаючи на мультиколінеарність, є маловтішним. Ми повинні знати, що відбувається або що може відбутися в кожній окремій вибірці.

5.1.3. Практичні наслідки мультиколінеарності

Першим практичним наслідком мультиколінеарності є велика дисперсія і коваріація оцінок параметрів, обчислених за методом найменших квадратів.

Знову для ілюстрації повернемося до двофакторної регресійної моделі

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

та її вибіркового аналога

$$y = b_0 + b_1 x_1 + b_2 x_2 + e.$$

Дисперсія оцінок параметрів b_1 і b_2 має вигляд (додаток 5.1.1):

$$\text{var}(b_1) = \frac{\sigma_\varepsilon^2}{(1-r^2) \sum_{i=1}^n (x_{1i} - \bar{x}_1)^2}; \quad (5.12)$$

$$\text{var}(b_2) = \frac{\sigma_\varepsilon^2}{(1-r^2) \sum_{i=1}^n (x_{2i} - \bar{x}_2)^2}; \quad (5.13)$$

$$\text{cov}(b_1, b_2) = \frac{-r\sigma_\varepsilon^2}{(1-r^2) \sqrt{\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2 \cdot \sum_{i=1}^n (x_{2i} - \bar{x}_2)^2}}. \quad (5.14)$$

де r — коефіцієнт кореляції між x_1 і x_2 .

З (5.12) і (5.13) ми бачимо, що якщо коефіцієнт між факторами x_1 і x_2 збільшується, то дисперсії оцінок параметрів теж збільшуються.

З (5.14) також зрозуміло, що при збільшенні коефіцієнта кореляції між факторами x_1 і x_2 коефіцієнт коваріації параметрів збільшується за абсолютною величиною. Причому при наближенні до граничного значення це збільшення зростає експоненційно. Проілюструємо це на прикладі параметра b_1 .

У формулі (5.12) позначимо:

$$\frac{\sigma_\varepsilon^2}{\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2} = p;$$

у формулі (5.14):

$$\frac{-\sigma_\varepsilon^2}{\sqrt{\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2 \sum_{i=1}^n (x_{2i} - \bar{x}_2)^2}} = D.$$

Послідовно збільшуючи значення коефіцієнта кореляції, отримуємо дані, наведені в табл. 5.1.

Як бачимо з табл. 5.1, при наближенні коефіцієнта кореляції до свого граничного значення дисперсії параметрів і коефіцієнт коваріації параметрів зростають із значною швидкістю. Так, при $r = 0,999$ $\text{var}(b_1)$ у 50 разів перевищують своє значення за умови, що немає мультиколінеарності ($r_{x_1x_2} = 0$).

Залежність дисперсії і коваріації оцінених параметрів від значень коефіцієнта кореляції графічно зображено на малюнках 5.1. і 5.2.

Таблиця 5.1

Вплив зростання $r_{x_1x_2}$ на $\text{var}(b_1)$ і $\text{cov}(b_1b_2)$

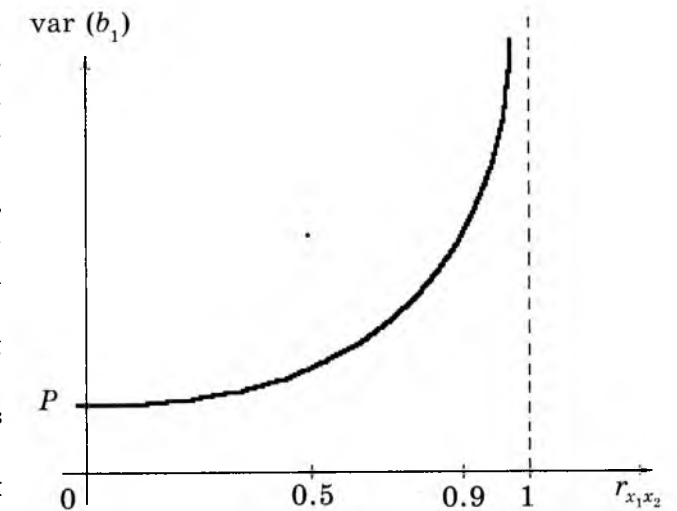
№ п/п	Значення $r_{x_1x_2}$	$\text{var}(b_1)$	$\text{cov}(b_1b_2)$
1	0	p	0
2	0.5	1.33·p	0.67·D
3	0.7	1.96·p	1.37·D
4	0.8	2.78·p	2.22·D
5	0.9	5.26·p	4.73·D
6	0.95	10.26·p	9.74·D
7	0.99	50.25·p	49.75·D
8	0.995	100·p	99.5·D
9	0.999	500·p	499.5·D

Другим практичним наслідком мультиколінеарності є збільшення інтервалу довіри. Оскільки збільшення коефіцієнта кореляції призводить до збільшення значень середньоквадратичних відхилень параметрів, то, звичайно, збільшується й інтервал довіри для них, що можна побачити з табл. 5.2.

Як бачимо, інтервал довіри за наявності високої мультиколінеарності ($r=0.999$) в $\sqrt{500}$ разів більший, ніж коли її немає ($r=0$).

Третім практичним наслідком мультиколінеарності є незначимість t -статистики.

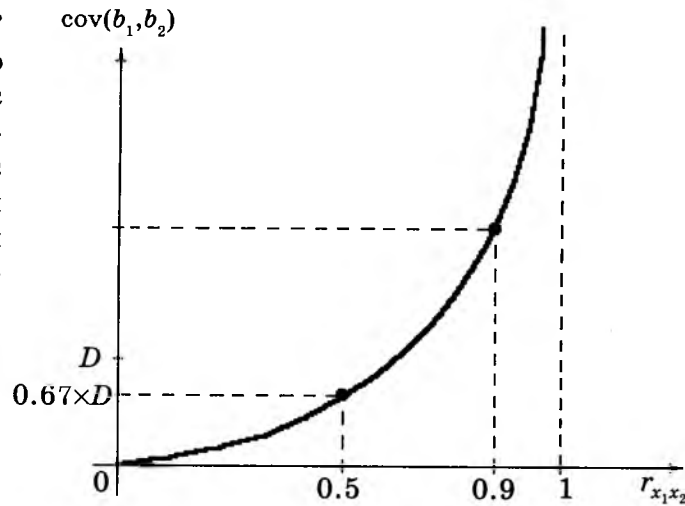
Нагадаємо: для того, щоб оцінити, чи значимо параметри багатofакторної регресії відрізняються від нуля, ми використовуємо t -статистику Ст'юдента. Для цього, наприклад, для параметра β_1 розраховуємо відношення: $t = \frac{b_1}{\hat{\sigma}_{b_1}}$ і порівнюємо його з табличним значенням t . У випадку мультиколінеарності $\hat{\sigma}_{b_1}$ нескінченно зростає, а t -значення прямує до нуля.

Малюнок 5.1. Залежність $\text{var}(b_1)$ від зростання значення коефіцієнта кореляції

Коли мультиколінеарність не є проблемою?

Якщо єдиною метою регресійного аналізу є прогноз, тоді мультиколінеарність не викликає проблем, оскільки чим вище значення R^2 , тим точніший прогноз. Це справедливо лише доти, "... доки значення залежних змінних, для яких і здійснюється прогноз, мають однакову майже лінійну залежність з початковою матрицею X ". Таким чином, якщо у побудованій регресії встановлено, що приблизно $x_1 = 2x_2$, то в наступних прикладах прогнозування x_1 повинно приблизно дорівнювати $2x_2$. Ця умова майже не здійснена на практиці. Більше того, якщо метою аналізу є не прогноз, а дійсні значення параметрів, мультиколінеарність перетворюється на проблему, оскільки вона призводить до великих стандартних помилок в оцінці параметрів.

Є випадок, коли мультиколінеарність не переростає в проблему: якщо R^2 велике і параметри регресії є значимими, оскільки t -статистика висока. Коли може виникнути така ситуація? Джонстон зазначає: "Вона може виникнути, якщо оцінки параметрів мають такі числові значення, які "кращі" за правильні значення, так що ефект залишається, незважаючи на підвищення стандартної помилки, і/або тому що правильні значення самі по собі настільки великі, що навіть занижена оцінка залишається значимою".



Малюнок 5.2. Залежність $\text{cov}(b_1, b_2)$ від зростання значення коефіцієнта кореляції.

Таблиця 5.2

Ефект впливу збільшення коефіцієнта кореляції на інтервал довіри для параметра β_1

Значення $r_{x_1x_2}$	95% інтервал довіри для β_1
0	$b_1 \pm 1,96\sqrt{p}$
0.5	$b_1 \pm 1,96\sqrt{1,33 \cdot p}$
0.95	$b_1 \pm 1,96\sqrt{10,26 \cdot p}$
0.99	$b_1 \pm 1,96\sqrt{500 \cdot p}$

$$\sqrt{p} = \sqrt{\frac{\sigma_\varepsilon^2}{\sum_{i=1}^n (x_{1i} - \bar{x})^2}}$$

5.1.4. Тестування наявності мультиколінеарності та засоби її вилучення

5.1.4.1. Тестування наявності мультиколінеарності

На жаль, немає єдиного методу для визначення мультиколінеарності. Тому наведемо кілька методів тестування наявності мультиколінеарності.

1. Високе значення R^2 і незначимість t -статистики

Одночасна наявність цих двох факторів є "класичною" ознакою мультиколінеарності.

Розглянемо p -факторну регресійну модель

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon_i.$$

У випадку мультиколінеарності можна визначити за t -статистикою Ст'юдента, що один або більше оцінених параметрів статистично незначимо відрізняються від нуля. При високому значенні R^2 ми приймаємо з великим ступенем імовірності F -критерій Фішера, бо він відкидає нульову гіпотезу, коли $\beta_1 = \beta_2 = \dots = \beta_p = 0$.

Тому високе значення R^2 і статистична незначущість деяких параметрів може свідчити про наявність мультиколінеарності.

2. Високе значення парних коефіцієнтів кореляції

Другим поширеним тестом на наявність мультиколінеарності є перевірка значень парних коефіцієнтів кореляції. Якщо значення хоча б одного коефіцієнта кореляції більше 0.8, то мультиколінеарність є серйозною проблемою.

Однак проблемним у цьому тесті є те, що високе значення парних коефіцієнтів кореляції — достатня, але не необхідна умова наявності мультиколінеарності.

Мультиколінеарність може бути навіть при відносно невеликих значеннях парних коефіцієнтів кореляції у більше, ніж двофакторній регресійній моделі.

3. F -тест для визначення мультиколінеарності

Цей тест було запропоновано Глаубером і Фарром. Наявність мультиколінеарності свідчить про те, що один або більше факторів пов'язані між собою лінійною або приблизно лінійною залежністю. Одним із способів визначення щільності регресійного зв'язку є побудова регресійної залежності кожного фактора x_i з усіма іншими факторами і обчислення відповідного коефіцієнта детермінації R^2 для цього допоміжного регресійного рівняння. Тому F -тест має й іншу назву: **побудова допоміжної регресії**.

Коефіцієнт детермінації $R^2_{x_i x_1 x_2 \dots x_p}$ є коефіцієнтом детермінації в регресії, яка пов'язує фактор x_i з іншими факторами. Наприклад, $R^2_{x_2 x_1 x_3 \dots x_p}$ є коефіцієнтом детермінації такої регресії: $x_2 = b_0 + b_1 x_1 + b_3 x_3 + \dots + b_p x_p + e$.

Для кожного коефіцієнта детермінації розраховуємо F_i -відношення:

$$F_i = \frac{(R_{x_i, x_1, \dots, x_p}^2)/(p-1)}{(1 - R_{x_i, x_1, \dots, x_p}^2)/(n-p)};$$

де n — кількість спостережень; p — кількість факторів.

F -тест перевіряє гіпотезу

$$H_0: R_{x_i, x_1, \dots, x_p}^2 = 0$$

проти гіпотези:

$$H_1: R_{x_i, x_1, \dots, x_p}^2 \neq 0.$$

Розраховані значення F_i порівнюємо з критичними значеннями $F_{\alpha, p}$, знайденими за таблицями F -розподілу Фішера з $(p-1)$ і $(n-p)$ ступенями вільності і заданим рівнем значимості. Якщо $F_i > F_{\alpha, p}$, тоді ми відкидаємо нуль-гіпотезу і вважаємо, що фактор x_i є мультиколінеарним; якщо ж $F_i < F_{\alpha, p}$, то приймаємо H_0 -гіпотезу і впевнюємось, що фактор x_i не є мультиколінеарним.

4. Характеристичні значення та умовний індекс

У деяких сучасних статистичних пакетах для перевірки наявності мультиколінеарності використовують характеристичні значення та умовний індекс. Ми не будемо детально розглядати, як обчислювати характеристичні значення, бо це потребує використання апарату теорії матриць. Відмітимо лише, що за цим тестом ми розраховуємо умовне число k :

$$k = \frac{\text{максимальне характеристичне значення}}{\text{мінімальне характеристичне значення}}$$

і умовний індекс (CI):

$$CI = \sqrt{k}.$$

Якщо $100 \leq k \leq 1000$, то це свідчить про помірну мультиколінеарність, при $k > 1000$ маємо високу мультиколінеарність. Аналогічно, якщо $10 \leq CI \leq 30$, це свідчить про помірну мультиколінеарність, а $CI > 0$ — про високу.

Звичайно, ми розглянули лише основні методи тестування мультиколінеарності. І, як бачимо, жоден з них не є універсальним. Кожен з розглянутих методів має свої переваги і недоліки, застосування того чи іншого методу залежить від вибору дослідника.

5.1.4.2. Визначення рівня мультиколінеарності

Визначення рівня мультиколінеарності за своїм змістом близьке до тестування її наявності. Ми вже знаємо, що коли йдеться про двофакторну модель, то для визначення мультиколінеарності можна обмежитись аналізом парних коефіцієнтів кореляції. Однак, якщо факторів більше, ніж два, то аналізу парних коефіцієнтів кореляції уже недостатньо.

Нехай ми маємо p -факторну регресійну модель

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + \varepsilon_i. \quad (5.15)$$

Тепер ми зможемо побудувати p багатфакторних регресій відповідно для факторів x_1, x_2, \dots, x_p .

$$\begin{aligned} x_{1j} &= a_0 + a_1 x_{2j} + \dots + a_{p-1} x_{pj}; \\ x_{2j} &= c_0 + c_1 x_{1j} + \dots + c_{p-1} x_{pj}; \\ &\dots \\ x_{pj} &= d_0 + d_1 x_{2j} + \dots + d_{p-1} x_{(p-1)j}. \end{aligned} \quad (5.16)$$

R_i^2 ($i = 1, p$) є коефіцієнтом детермінації для кожного i -го фактора. Він якраз і використовується для того, щоб визначити ступінь мультиколінеарності.

Для цього розраховуємо величину дисперсійно-інфляційного фактора VIF (variance inflationary factor) для кожної змінної:

$$VIF_i = \frac{1}{1 - R_i^2}. \quad (5.17)$$

Дослідники використовують значення $VIF_i = 10$ як критичне. Якщо $VIF_i \leq 10$, то можна стверджувати про недостатність зв'язку між i -м фактором і всіма іншими. Якщо $VIF_i \geq 10$, то це свідчить про наявність мультиколінеарності.

5.1.4.3. Засоби вилучення мультиколінеарності

Що робити, коли мультиколінеарність виявлено? Безпомилкових і абсолютно правильних порад нема, оскільки мультиколінеарність є прикладною проблемою. Звичайно, все залежить від ступеня мультиколінеарності, але у будь-якому випадку можна запропонувати декілька простих методів вилучення мультиколінеарності.

1. Використання первинної інформації

Аналіз і використання первинної інформації інколи дозволяє зняти проблему мультиколінеарності. Пояснимо це на прикладі, розглянувши залежність між споживанням, доходом та багатством. Цю залежність можна описати такою моделлю:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i, \quad (5.18)$$

де y — споживання, x_1 — доход, x_2 — багатство. Відомо, що доход та багатство є висококолінеарними факторами. Припустимо, попередньо визначено, що зв'язок між доходом та багатством є таким:

$$\beta_2 = 0.2\beta_1.$$

Тоді модель (5.18) можна переписати у вигляді:

$$y_i = \beta_0 + \beta_1 x_{1i} + 0.2\beta_1 x_{2i} + \varepsilon_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad (5.19)$$

де $x_i = x_{1i} + 0.2x_{2i}$. Звідси ми можемо знайти оцінку параметра $\beta_1 \Rightarrow b_1$, а потім і b_2 , виходячи із постульованої залежності між β_1 і β_2 .

Як ми отримуємо апріорну інформацію? Звичайно, багато в чому спираємося на економічну теорію. Наприклад, для виробничої функції Кобба — Дугласа $y = \beta_0 x_1^{\beta_1} x_2^{\beta_2} e^\varepsilon$ при постійному масштабі виробництва маємо: $\beta_1 + \beta_2 = 1$. Якщо при цьому спостерігається колінеарність між факторами x_1 (праця) і x_2 (капітал), тоді заміна змінних може зменшити або й усунути колінеарність. Проте до методу використання первинної інформації слід підходити з великою обережністю.

2. Об'єднання міжгалузевої та динамічної інформації

Окремим випадком методу використання первинної інформації, який ми розглядали вище, є об'єднання міжгалузевої інформації та інформації за часовими рядами. Цей метод відомий як метод зведення інформації. Припустимо, ми хочемо вивчити попит на автомобілі в Україні і маємо дані про продаж машин за певний час, середню ціну машини та доход споживача. Припустимо також, що

$$\ln y_t = \beta_0 + \beta_1 \ln P_t + \beta_2 \ln I_t + \varepsilon_t,$$

де y — кількість проданих машин, P — середня ціна, I — доход, t — час. Нашою метою є визначення еластичності ціни β_1 і еластичності доходу β_2 .

У часово-кількісних даних цінова змінна та змінна доходу схильні до високої колінеарності. Тому, якщо застосовувати попередню регресію, то перед нами постане звичайна проблема мультиколінеарності. Метод її

вирішення був запропонований Тобінім. Якщо ми працюємо з міжгалузевою інформацією (такою як бюджетні дослідження, що проводяться багатьма приватними та урядовими агентами), то можемо отримати більш надійні оцінки еластичності, наприклад, доходу β_2 , тому що в таких даних ціна змінюється неістотно. Використовуючи цю оцінку, попередню регресію можна записати у вигляді:

$$y_i^* = \beta_0 + \beta_1 \ln P_i + \varepsilon_i,$$

де $y^* = \ln y - \beta_2 \ln I$ і відповідає значенню y після усунення ефекту доходу. Тепер ми можемо отримати параметри еластичності ціни з попередньої регресії.

Незважаючи на те, що цей метод привабливий, при його використанні можуть виникнути проблеми з аналізом отриманих результатів, бо ми припустили, що отримана еластичність доходу дорівнює тій, яку б ми отримали за чистого аналізу часових рядів. Все ж таки, цей метод може застосовуватися в багатьох ситуаціях за умови, що міжгалузеві оцінки відрізняються неістотно.

3. Вилучення змінної (змінних) та помилки специфікації

Якщо ми маємо високу мультиколінеарність, тоді найкраще та найлегше просто відкинути одну із залежних змінних. Таким чином, у прикладі із споживанням, доходом та багатством, якщо опустити змінну, що відповідає багатству, ми отримаємо регресійну модель з однією незалежною змінною — доходом.

Але вилучення змінної з моделі може призвести до помилки специфікації. Помилка специфікації виникає через некоректне визначення моделі, що використовується в аналізі. Так, якщо за економічною теорією для пояснення розширення споживання модель повинна включати і доход, і багатство, тоді вилучення змінної багатства створюватиме помилку специфікації.

Детально проблеми специфікації ми розглядати не будемо, відмітимо тільки, що, якщо правильною моделлю є

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i$$

але ми помилково звели її до вигляду

$$y_i = \beta_0 + \beta_1 x_{2i} + \varepsilon_i, \quad (5.20)$$

тоді

$$E(b_{12}) = \beta_1 + \beta_2 b_{21} \neq \beta_1, \quad (5.21)$$

де b_{21} — нахил у регресії x_2 відносно x_1 . Як зрозуміло з (5.21), b_{12} є зміщеною оцінкою β_1 , коли b_{21} відмінне від нуля (тоді існує залежність між x_2 та x_1). Звичайно, якщо b_{21} дорівнює нулю, то з самого початку проблеми мультиколінеарності немає.

Отже, вилучення змінної з моделі з метою зниження мультиколінеарності, може призвести до зміщених оцінок. Наслідки від цього можуть бути гіршими, ніж сама проблема колінеарності. Адже при мультиколінеарності оцінки параметрів залишаються BLUE-оцінками, тоді як вилучення змінної може привести до зміщених оцінок, які не є BLUE-оцінками.

4. Перетворення змінних

Припустимо, дані за певний період часу щодо споживання, доходу та багатства. Одна з причин мультиколінеарності цих даних є їхня схильність змінюватись в одному напрямку, а один із шляхів зменшення такої залежності — використання перших різниць у моделі.

Наприклад, якщо залежність

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon_t \quad (5.22)$$

дійсна в час t , вона дійсна і для часу $t-1$. Тому отримуємо:

$$y_{t-1} = \beta_0 + \beta_1 x_{1,t-1} + \beta_2 x_{2,t-1} + \varepsilon_{t-1}. \quad (5.23)$$

Якщо підставимо (5.23) в (5.22), то матимемо:

$$y_t - y_{t-1} = \beta_1 (x_{1t} - x_{1,t-1}) + \beta_2 (x_{2t} - x_{2,t-1}) + v_t, \quad (5.24)$$

де $v_t = \varepsilon_t - \varepsilon_{t-1}$. Рівняння (5.24) відоме як *рівняння перших різниць*, бо ми отримали регресію не з початкових змінних, а з різниць послідовних значень змінних.

Цей прийом часто зменшує мультиколінеарність, бо, хоча значення x_1 і x_2 можуть мати високу кореляцію, їхні різниці не завжди висококорельовані.

Такі перетворення породжують певні додаткові проблеми. Випадкова величина v_t в (5.24) може не задовольняти припущенням моделі класичної лінійної регресії про незалежність. Як ми побачимо в параграфі про автокореляцію, коли початкове значення ε_t є послідовно незалежним або некорельованим, тоді випадкова величина v_t буде в багатьох випадках послідовно корельованою. Знову ж таки намагання поліпшити ситуацію призводять до гіршого стану, ніж був спочатку. Більше того, використання перших різниць призводить до зменшення ступенів вільності на одиницю. У роботі з малими вибірками цей фактор також треба брати до уваги.

5. Збільшення спостережень

Оскільки мультиколінеарність змінюється у кожній вибірці, то можливо, що в іншій моделі з такими ж змінними мультиколінеарність буде іншою. Іноді просте збільшення спостережень у моделі (якщо це можливо) пом'якшує проблему мультиколінеарності. Наприклад, у моделі з двома змінними ми бачимо, що

$$\text{var}(b_2) = \frac{\sigma^2}{\sum_{i=1}^n \tilde{x}_{2i}^2 (1 - r_{12}^2)},$$

де $\tilde{x}_{2i}^2 = (x_{2i} - \bar{x}_2)^2$.

Якщо збільшувати кількість спостережень, то $\sum \tilde{x}_{2i}^2$ завжди збільшуватиметься. Тому для кожного даного r_{12} дисперсія b_2 буде зменшуватися, зменшуючи таким чином стандартну помилку, що допомагає оцінити β_2 точніше.

Отримати додаткові дані не завжди легко, оскільки на практиці це часто вимагає значних витрат. Крім того, потрібно впевнитись, що економічна структура, пов'язана з новими спостереженнями, буде подібна до початкової структури.

6. Інші методи виправлення мультиколінеарності

Статистичні методи, такі як *факторний аналіз*, *метод головних компонент*, *гребенева регресія*, часто використовуються для виправлення мультиколінеарності. На жаль, ми їх не розглядаємо, бо вони потребують спеціальних математичних знань. Але всі ці методи є в сучасних комп'ютерних програмах і можуть застосовуватись на практиці.

5.1.5. Основні висновки щодо наявності мультиколінеарності в регресійній моделі

Однією з умов класичної лінійної регресії є припущення про відсутність мультиколінеарності між факторами x . Іншими словами, мультиколінеарність наявна тоді, коли між факторами x є лінійна або майже лінійна залежність.

Наслідки мультиколінеарності: за досконалої колінеарності між факторами x параметри регресії та їхні середньоквадратичні відхилення неможливо визначити. Якщо мультиколінеарність велика, але не досконала, визначити параметри регресії можна, але їхні середньоквадратичні відхилення будуть дуже великими. Як наслідок, значення параметрів для сукупності не можна визначити точно.

Хоча надійних методів тестування колінеарності не існує, є декілька її індикаторів.

1. Найкращою ознакою мультиколінеарності є високе значення коефіцієнта детермінації при незначимості параметрів за t -тестом.

2. У моделі з двома змінними найкращою ознакою мультиколінеарності є значення простого коефіцієнта кореляції.

3. У моделі з більш як двома змінними, простий коефіцієнт кореляції може бути низький за наявності мультиколінеарності. За таких умов потрібно брати до уваги часткові коефіцієнти кореляції.

4. Якщо коефіцієнт детермінації великий, а часткові коефіцієнти кореляції низькі, то мультиколінеарність можлива. Також можливо, що у моделі є надлишкові змінні. Але якщо коефіцієнт детермінації високий і часткові коефіцієнти кореляції високі, то мультиколінеарність не завжди можна виявити.

5. Ще один спосіб виявлення мультиколінеарності — побудова регресії для кожної змінної x_j від інших x та знаходження відповідного коефіцієнта детермінації R_j^2 . Більше значення R_j^2 відповідає вищій кореляції x_j з іншими змінними x .

Виявлення мультиколінеарності є лише частиною справи. Інша частина — як її позбутися. Немає якихось певних методів, можна дати лише окремі поради: (1) використання додаткової або первинної інформації; (2) об'єднання інформації; (3) відкидання змінної з високою кореляцією; (4) перетворення даних (використання перших різниць); (5) збільшення спостережень. Які поради спрацюють на практиці, залежить від істотності проблеми та її характеру.

Додаток 5.1

1. Двофакторна регресійна модель

Розглянемо двофакторну регресійну модель

$$y = b_0 + b_1x_1 + b_2x_2 + e.$$

Знайдемо оцінки невідомих параметрів методом найменших квадратів:

$$\sum_{i=1}^n l_i^2 = \sum_{i=1}^n (y_i - b_0 - b_1x_{1i} - b_2x_{2i})^2. \quad (5.0)$$

Для цього розрахуємо часткові похідні і прирівняємо їх до нуля; отримуємо систему нормальних рівнянь:

$$\begin{aligned} nb_0 + b_1 \sum_{i=1}^n x_{1i} + b_2 \sum_{i=1}^n x_{2i} &= \sum_{i=1}^n y_i; \\ b_0 \sum_{i=1}^n x_{1i} + b_1 \sum_{i=1}^n x_{1i}^2 + b_2 \sum_{i=1}^n x_{1i}x_{2i} &= \sum_{i=1}^n x_{1i}y_i; \\ b_0 \sum_{i=1}^n x_{2i} + b_1 \sum_{i=1}^n x_{1i}x_{2i} + b_2 \sum_{i=1}^n x_{2i}^2 &= \sum_{i=1}^n x_{2i}y_i. \end{aligned} \quad (5.1)$$

Якщо ввести заміну змінних:

$$(x_{1i} - \bar{x}_1) = \tilde{x}_{1i}; (x_{2i} - \bar{x}_2) = \tilde{x}_{2i}; (y_i - \bar{y}) = \tilde{y}_i,$$

то (5.1) можна переписати у вигляді:

$$b_1 \sum_{i=1}^n \tilde{x}_{1i}^2 + b_2 \sum_{i=1}^n \tilde{x}_{1i}\tilde{x}_{2i} = \sum_{i=1}^n \tilde{x}_{1i}\tilde{y}_i; \quad (5.2)$$

$$b_1 \sum_{i=1}^n \tilde{x}_{1i}\tilde{x}_{2i} + b_2 \sum_{i=1}^n \tilde{x}_{2i}^2 = \sum_{i=1}^n \tilde{x}_{2i}\tilde{y}_i.$$

З (5.2) отримаємо:

$$b_1 = \frac{\sum_{i=1}^n \tilde{x}_{1i}\tilde{y}_i \sum_{i=1}^n \tilde{x}_{2i}^2 - \sum_{i=1}^n \tilde{x}_{2i}\tilde{y}_i \sum_{i=1}^n \tilde{x}_{1i}\tilde{x}_{2i}}{\sum_{i=1}^n \tilde{x}_{1i}^2 \sum_{i=1}^n \tilde{x}_{2i}^2 - \left(\sum_{i=1}^n \tilde{x}_{1i}\tilde{x}_{2i}\right)^2}; \quad (5.3)$$

$$b_2 = \frac{\sum_{i=1}^n \tilde{x}_{1i}^2 \sum_{i=1}^n \tilde{x}_{2i}\tilde{y}_i - \sum_{i=1}^n \tilde{x}_{1i}\tilde{x}_{2i} \sum_{i=1}^n \tilde{x}_{1i}\tilde{y}_i}{\sum_{i=1}^n \tilde{x}_{1i}^2 \sum_{i=1}^n \tilde{x}_{2i}^2 - \left(\sum_{i=1}^n \tilde{x}_{1i}\tilde{x}_{2i}\right)^2}; \quad (5.4)$$

$$b_0 = \bar{y} - b_1\bar{x}_1 - b_2\bar{x}_2. \quad (5.5)$$

2. Розрахунок дисперсії параметрів двофакторної моделі

Коефіцієнт кореляції між факторами x_1 і x_2 :

$$r = \frac{(x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2)}{A_1 A_2}, \quad (5.6)$$

$$\text{де } A_1 = \sqrt{\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2}; \quad (5.7)$$

$$A_2 = \sqrt{\sum_{i=1}^n (x_{2i} - \bar{x}_2)^2}. \quad (5.8)$$

Вводячи (5.6) — (5.8) у рівняння (5.3), отримаємо:

$$\begin{aligned} b_1 &= \frac{A_2^2 \left[\sum_{i=1}^n (x_{1i} - \bar{x}_1)\tilde{y}_i \right] - r^2 A_1 A_2 \left[\sum_{i=1}^n (x_{2i} - \bar{x}_2)\tilde{y}_i \right]}{A_1^2 A_2^2 - r^2 A_1^2 A_2^2} = \\ &= \sum_{i=1}^n \left[\frac{(x_{1i} - \bar{x}_1)\tilde{y}_i - (rA_1 / A_2)(x_{2i} - \bar{x}_2)\tilde{y}_i}{(1 - r^2)A_1^2} \right] = \\ &= \sum_{i=1}^n \left[\frac{(x_{1i} - \bar{x}_1) - (rA_1 / A_2)(x_{2i} - \bar{x}_2)}{(1 - r^2)A_1^2} \right] \tilde{y}_i. \end{aligned} \quad (5.9)$$

Проведемо заміну у (5.9):

$$\sum_{i=1}^n (x_{1i} - \bar{x}_1) \tilde{y}_i = \sum_{i=1}^n (x_{1i} - \bar{x}_1) y_i ;$$

$$\sum_{i=1}^n (x_{2i} - \bar{x}_2) \tilde{y}_i = \sum_{i=1}^n (x_{2i} - \bar{x}_2) y_i .$$

Позначивши у (5.9) коефіцієнт при y_i через B_{1i} , отримаємо:

$$b_1 = \sum_{i=1}^n B_{1i} y_i . \quad (5.10)$$

Аналогічно:

$$b_2 = \sum_{i=1}^n \left(\frac{(x_{2i} - \bar{x}_2) - \frac{rA_2}{A_1} (x_{1i} - \bar{x}_1)}{(1-r^2)A_2^2} \right) y_i = \sum_{i=1}^n B_{2i} y_i . \quad (5.11)$$

Підставивши у (5.0) рівняння (5.10) і (5.11), отримаємо:

$$b_1 = b_0 \sum_{i=1}^n B_{1i} + b_1 \sum_{i=1}^n B_{1i} x_{1i} + b_2 \sum_{i=1}^n B_{1i} x_{2i} + \sum_{i=1}^n B_{1i} e_i ;$$

$$b_2 = b_0 \sum_{i=1}^n B_{2i} + b_1 \sum_{i=1}^n B_{2i} x_{1i} + b_2 \sum_{i=1}^n B_{2i} x_{2i} + \sum_{i=1}^n B_{2i} e_i . \quad (5.12)$$

Легко показати, що:

$$\sum_{i=1}^n B_{1i} = 0; \quad \sum_{i=1}^n B_{2i} = 0; \quad \sum_{i=1}^n B_{1i} x_{1i} = 1; \quad \sum_{i=1}^n B_{2i} x_{2i} = 1; \quad \sum_{i=1}^n B_{2i} x_{1i} = 0; \quad \sum_{i=1}^n B_{1i} x_{2i} = 0.$$

Тоді (5.12) і (5.13) можна переписати у вигляді:

$$b_1 - E(b_1) = b_1 - \beta_1 = \sum_{i=1}^n B_{1i} e_i ; \quad (5.14)$$

$$b_2 - E(b_2) = b_2 - \beta_2 = \sum_{i=1}^n B_{2i} e_i . \quad (5.14')$$

Виходячи з (5.14), отримаємо:

$$\text{var}(b_1) = E \left[\left(\sum_{i=1}^n B_{1i} e_i \right)^2 \right] - \left[E \left(\sum_{i=1}^n B_{1i} e_i \right) \right]^2 =$$

$$= \sum_{i=1}^n \sum_{j=1}^n B_{1i} B_{1j} E(e_i e_j) - \left(\sum_{i=1}^n B_{1i} \right)^2 [E(e_i)^2]^2 =$$

$$= \sigma_\varepsilon^2 \sum_{i=1}^n B_{1i}^2 . \quad (5.15)$$

Нагадаємо: $E(l_i) = 0$ і $E(l_i l_j) = 0$, коли $i \neq j$.

За означенням B_{1i} маємо:

$$B_{1i}^2 = \frac{(x_{1i} - \bar{x}_1)^2 + r^2 A_1^2 / A_2^2 (x_{2i} - \bar{x}_2)^2 - 2r(A_1/A_2)(x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2)}{(1-r^2)A_1^4} . \quad (5.16)$$

Виходячи з (5.16), (5.6), (5.7) і (5.8):

$$\sum_{i=1}^n B_{1i}^2 = \frac{A_1^2 + r^2(A_1^2/A_2^2)(A_2^2) - 2r(A_1/A_2)(rA_1A_2)}{(1-r^2)^2 A_1^4} =$$

$$= \frac{A_1^2(1-2r+r^2)}{(1-r^2)^2/A_1^4} = \frac{1}{(1-r^2)^2 A_1^2} .$$

Підставимо (5.17) у (5.15):

$$\text{var}(b_1) = \frac{\hat{\sigma}_\varepsilon^2}{(1-r^2) \times \sum_{i=1}^n (x_{1i} - \bar{x}_1)^2} . \quad (5.18)$$

Аналогічно, для параметра b_2 :

$$\text{var}(b_2) = \frac{\hat{\sigma}_\varepsilon^2}{(1-r^2) \times \sum_{i=1}^n (x_{2i} - \bar{x}_2)^2} . \quad (5.19)$$

3. Коефіцієнт коваріації для двофакторної регресії

Запишемо прогнозне значення для двофакторної регресійної моделі:

$$\hat{y}_{n+1} = b_0 + b_1 x_{1,n+1} + b_2 x_{2,n+1} = \bar{y} + b_1(x_{1,n+1} - \bar{x}_1) + b_2(x_{2,n+1} - \bar{x}_2) . \quad (5.20)$$

Тоді дисперсія \hat{y}_{n+1} матиме вигляд:

$$\text{var}(\hat{y}_{n+1}) = \text{var}[\bar{y} + b_1(x_{1,n+1} - \bar{x}_1) + b_2(x_{2,n+1} - \bar{x}_2)] =$$

$$= \text{var}(\bar{y}) + \text{var}[b_1(x_{1,n+1} - \bar{x}_1)] + \text{var}[b_2(x_{2,n+1} - \bar{x}_2)] +$$

$$+ 2 \text{cov}[b_1(x_{1,n+1} - \bar{x}_1), b_2(x_{2,n+1} - \bar{x}_2)] =$$

$$= \frac{\sigma_\varepsilon^2}{n} + (x_{1,n+1} - \bar{x}_1)^2 \text{var}(b_1) + (x_{2,n+1} - \bar{x}_2)^2 \text{var}(b_2) + 2(x_{1,n+1} - \bar{x}_1)(x_{2,n+1} - \bar{x}_2) \text{cov}(b_1, b_2). \quad (5.21)$$

З (5.14) і (5.14') можемо показати, що

$$\text{cov}(b_1, b_2) = E(b_1 - \beta_1)(b_2 - \beta_2) = \sum_{i=1}^n \sum_{j=1}^n B_{1i} B_{2j} E(e_i e_j). \quad (5.22)$$

Оскільки $E(e_i e_j) = 0$, коли $i \neq j$, то рівняння (5.22) можна скоротити:

$$\text{cov}(b_1, b_2) = \hat{\sigma}_\varepsilon^2 \sum_{i=1}^n \sum_{j=1}^n B_{1i} B_{2j} \delta_{ij}. \quad (5.23)$$

Легко показати, що

$$\sum_{i=1}^n B_{1i} B_{2i} = \frac{-r}{(1-r^2)A_1 A_2}. \quad (5.24)$$

Якщо (5.24) підставити у (5.23), то

$$\text{cov}(b_1, b_2) = \frac{-r \hat{\sigma}_\varepsilon^2}{(1-r^2) \sqrt{\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2 \sum_{i=1}^n (x_{2i} - \bar{x}_2)^2}}. \quad (5.25)$$

Підставляючи (5.17), (5.24) у (5.25), отримаємо:

$$\text{var}(y_{n+1}) = \left\{ \frac{1}{n} + \frac{(x_{1,n+1} - \bar{x}_1)^2}{(1-r^2) \sum_{i=1}^n (x_{1i} - \bar{x}_1)^2} + \frac{(x_{2,n+1} - \bar{x}_2)^2}{(1-r^2) \sum_{i=1}^n (x_{2i} - \bar{x}_2)^2} - \frac{2(x_{1,n+1} - \bar{x}_1)(x_{2,n+1} - \bar{x}_2)r}{(1-r^2) \sqrt{\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2 \sum_{i=1}^n (x_{2i} - \bar{x}_2)^2}} \right\} \hat{\sigma}_\varepsilon^2. \quad (5.26)$$

5.2. ГЕТЕРОСКЕДАСТИЧНІСТЬ ТА ЇЇ НАСЛІДКИ. ЗНАХОДЖЕННЯ ГЕТЕРОСКЕДАСТИЧНОСТІ ТА ЇЇ ВИЛУЧЕННЯ

5.2.1. Визначення гетероскедастичності та її природа

Одним з основних припущень моделі класичної лінійної регресії є припущення про сталість дисперсії кожної випадкової величини ε_i (гомоскедастичність).

Формалізовано це припущення записується у вигляді:

$$\text{var}(\varepsilon_i) = E\{(\varepsilon_i - E(\varepsilon_i))^2\} = \sigma_\varepsilon^2 = \text{const}.$$

Якщо це припущення не задовольняється у якомусь окремому випадку, то має місце гетероскедастичність:

$$\text{var}(\varepsilon_i) = \sigma_{\varepsilon_i}^2 \neq \text{const}.$$

Звичайно, нас буде цікавити питання про доцільність цього припущення і про те, що відбувається, коли припущення про сталість дисперсії випадкової величини ε не задовольняється. Отже, спробуємо розглянути:

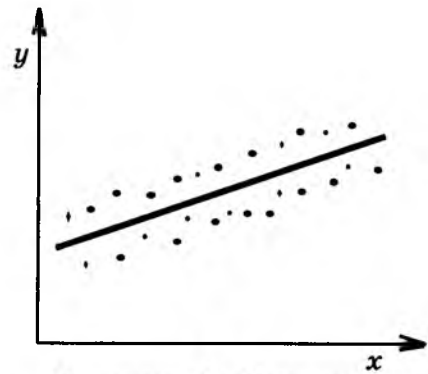
- 1) природу, або суть гетероскедастичності;
- 2) наслідки гетероскедастичності;
- 3) можливості тестування гетероскедастичності;
- 4) які корективні заходи потрібно вжити в разі порушення гетероскедастичності?

Суть припущення гомоскедастичності полягає в тому, що варіація кожної ε_i навколо її математичного сподівання не залежить від значення x . Дисперсія кожної ε_i зберігається сталою незалежно від малих чи великих значень факторів: σ_ε^2 не є функцією x_{ij} , тобто $\sigma_\varepsilon^2 \neq f(x_{1i}, x_{2i}, \dots, x_{pi})$.

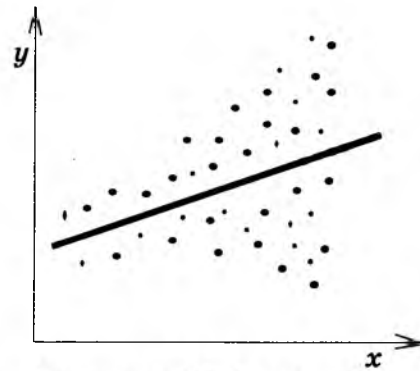
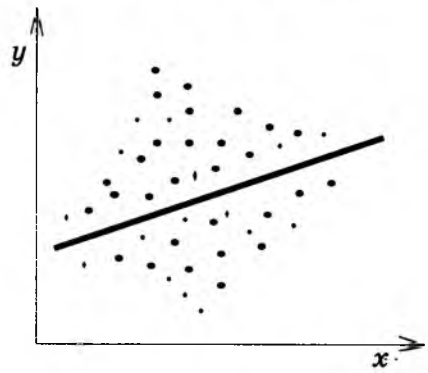
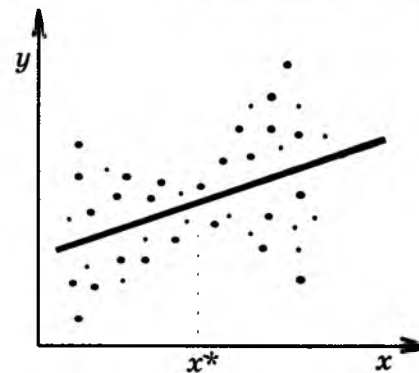
Графічно випадок гомоскедастичності для простої лінійної регресії показано випадковою дисперсією ε у межах сталої відстані навколо лінії регресії (див. мал. 5.3 а).

Якщо σ_ε^2 не є сталою, а її значення залежать від значень x , можемо записати $\sigma_\varepsilon^2 = f(x_{1i}, x_{2i}, \dots, x_{pi})$. У цьому разі маємо справу з гетероскедастичністю. Графічна форма розкиду спостережень залежить від форми гетероскедастичності, тобто форми зв'язку між σ_ε^2 та x_i . На мал. 5.3 (б, в, г) показано три різні форми гетероскедастичності. Зокрема на мал. 5.3б показано випадок (монотонно) зростаючої дисперсії ε_i (із зростанням x зростає і дисперсія ε).

Це загальноприйнята форма гетероскедастичності, що допускається в економетричних дослідженнях. На мал. 5.3в показано випадок спадної



Малюнок 5.3а. Гомоскедастичність

Малюнок 5.3б. Зростання дисперсії ε Малюнок 5.3в. Спадання дисперсії ε 

Малюнок 5.3г. Складний випадок

гетероскедастичності: коли x набуває більших значень, відхилення спостережень від лінії регресії зменшується, таким чином дисперсія випадкової змінної зменшується із зростанням x . На мал. 5.3г зображено більш складну форму гетероскедастичності: спочатку дисперсія ε зменшується із зростанням x , але після певного рівня x^* починає зростати із зростанням x . Зрозуміло, що форма гетероскедастичності залежить від знаків та значень коефіцієнтів у залежності $\sigma_{\varepsilon_i}^2 = f(x_{1i}, x_{2i}, \dots, x_{pi})$. Оскільки ε_i — неспостережувана випадкова величина, ми не знаємо справжньої форми гетероскедастичності.

У прикладних дослідженнях, як правило, використовується зручне припущення, що гетероскедастичність (у разі простої лінійної регресії) має форму $\sigma_{\varepsilon_i}^2 = k^2 x_i^2$, де k — константа, яку потрібно оцінити.

5.2.2. Правдоподібність припущення про гомоскедастичність

У багатьох економетричних дослідженнях може очікуватись, що припущення про сталу дисперсію випадкової змінної не зберігатиметься. Це можна легко зрозуміти, якщо врахувати фактори, вплив яких абсорбується значенням помилки. Згадаємо, що випадкова величина ε виражає вплив на залежну змінну помилок в її вимірюванні та неврахованих факторів. У обох випадках є підстави для зміни з часом дисперсії ε . Помилки вимірювання мають тенденцію до накопичення з плином часу, тому їхня величина збільшується. У такому разі дисперсія ε_i збільшується із зростанням значень x . З іншого боку, техніка вибірки та інші методи збору даних постійно вдосконалюються і тому похибки вимірювання можуть зменшуватися. У такому разі $\sigma_{\varepsilon_i}^2$ також зменшуватиметься. Але важливіше те, що багато із неврахованих змінних можуть змінюватись в однаковому з x напрямку, викликаючи, таким чином, збільшення відхилення спостережень від лінії регресії. Розглянемо, наприклад, просту лінійну регресійну модель

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,$$

де y_i — заощадження i -го домогосподарства;

x_i — дохід i -го домогосподарства (мал. 5.4 та 5.5).

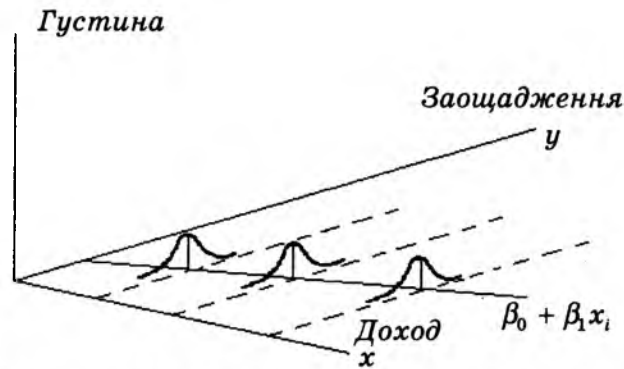
З мал. 5.4 та 5.5 видно, що із зростанням доходу заощадження також зростають. Але на мал. 5.4 дисперсія заощаджень залишається однаковою за усіх рівнів доходу, тоді як на мал. 5.5 вона зростає разом з доходом.

Яка ж ситуація типовіша? Звичайно, ситуація, зображена на мал. 5.5. Сім'ї з більшим доходом, як правило, показують більшу варіацію у своїй поведінці заощаджень, ніж сім'ї з низьким доходом. Сім'ї з високим доходом схильні дотримуватись певного стандарту життя, і коли їхній дохід падає, вони швидше скоротять свої заощадження, ніж споживання. З іншого боку, сім'ї з низьким доходом заощаджують із певною метою (наприклад, щоб сплатити розстрочку чи борг), і тому їхні заощадження більш регулярні. Це означає, що при високих доходах вони будуть високими, тоді як при низьких доходах — малими. Тому припущення про сталу дисперсію не витримується при оцінюванні функції заощаджень з сімейного бюджету¹.

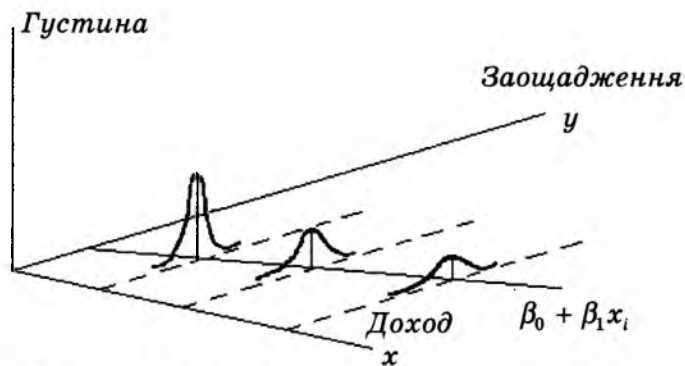
Якщо розглянути модель, котра пов'язує кількість помилок під час диктанту з кількістю годин, відведених на практику, то також можна помітити, що з плином часу дисперсія зменшується. Це очевидно: чим більше практикуєш, тим менше помиляєшся.

Аналогічно, компанії з більшими прибутками проводять ризикованішу дивідендну політику порівняно з компаніями, які мають менші прибутки. Банки, які мають у розпорядженні обладнання для обробки даних у місяч-

¹ Goldberger A.S. *Econometrics Theory*. — P. 231.



Малюнок 5.4. Випадок гомоскедастичності



Малюнок 5.5. Випадок гетероскедастичності

них та кварталних звітах своїх клієнтів, помиляються менше, ніж банки, які не мають подібних засобів обслуговування.

Розглянемо вибірку фірм певної галузі з метою оцінки виробничої функції Кобба — Дугласа:

$$y = \beta_0 L^{\beta_1} K^{\beta_2} e^{\varepsilon}$$

У цьому разі ε вміщує в собі такі фактори, як підприємництво, технологічні відмінності заводів різних фірм, відмінності у навиках чи організації та інші фактори. Ці фактори не мають значного коливання у малих фірмах і значно варіюють у великих фірмах. Отже, ε будуть гетероскедастичними.

Підсумовуючи, зазначимо, що апріорі є підстави стверджувати: на практиці припущення про гомоскедастичність порушуються. Тому важливо дослідити наслідки гетероскедастичності при оцінці невідомих параметрів та їхніх середньоквадратичних відхилень. Доведено, що звичайний випадок гетероскедастичності — це зростаюча дисперсія ε^1 .

¹ Katona G. Contributions of Survey Methods to Econometrics. — New York, 1954. — P. 203.

5.2.3. Наслідки порушення припущення про гомоскедастичність

Якщо припущення про гомоскедастичність випадкової величини не виконується, то ми не можемо застосувати формули дисперсії параметрів регресії для оцінки їхньої значимості та побудови інтервалів довіри. Згадаємо, що для простої лінійної регресії:

$$\text{var}(b_0) = \sigma_\varepsilon^2 \frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2}; \quad \text{var}(b_1) = \sigma_\varepsilon^2 \frac{1}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

Як бачимо, дисперсія σ_ε^2 не враховувалась у підсумовуванні, тому що припускалось, що вона постійна. З гетероскедастичністю дисперсія ε не буде скінченним сталим числом, а змінюватиметься із зростаючим розкидом значень.

Оцінка параметрів методом найменших квадратів за умов гетероскедастичності

Спробуємо відповісти на запитання, за інших однакових умов класичної моделі, що трапиться з оцінками параметрів за умов гетероскедастичності. ($E(\varepsilon_i^2) = \sigma_i^2$).

Щоб відповісти на це запитання, повернемося до простої лінійної регресії $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$. МНК-оцінка параметра β_1 має вигляд:

$$b_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}.$$

Якщо зберігаються припущення класичної моделі, включаючи гомоскедастичність, то b_1 є найкращою лінійною оцінкою без відхилень (BLUE-оцінкою). Цікаво, чи залишається b_1 BLUE-оцінкою в умовах гетероскедастичності? Можна показати (доведення наводити не будемо), що b_1 залишається лінійною оцінкою без відхилень, але вже не матиме найменшу дисперсію серед лінійних оцінок без відхилень, тобто не буде вже BLUE-оцінкою. На що ж *перетворюються* BLUE-оцінки параметрів за умов гетероскедастичності?

5.2.4. Оцінювання параметрів методом узагальнених найменших квадратів (УНК) у разі гетероскедастичності

На відміну від звичайного методу найменших квадратів (МНК), узагальнений метод (УНК) враховує інформацію про неоднаковість дисперсії і тому здатний створити BLUE-оцінки.

Щоб проілюструвати це, знову повернемося до простої лінійної регресії:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i. \quad (5.27)$$

Простою математичною маніпуляцією перепишемо (5.27) у вигляді:

$$y_i = \beta_0 x_{0i} + \beta_1 x_i + \varepsilon_i. \quad (5.28)$$

де $x_{0i} = 1$ для кожного i .

Припустимо, що наявна гетероскедастичність і всі дисперсії σ_i^2 відомі. Поділимо (5.28) на σ_i , отримаємо:

$$\frac{y_i}{\sigma_i} = \beta_0 \left(\frac{x_{0i}}{\sigma_i}\right) + \beta_1 \left(\frac{x_i}{\sigma_i}\right) + \left(\frac{\varepsilon_i}{\sigma_i}\right). \quad (5.29)$$

Для зручності перепишемо (5.29) у вигляді:

$$y_i^* = \beta_0^* x_{0i}^* + \beta_1^* x_i^* + \varepsilon_i^*, \quad (5.30)$$

де зірочками помічені початкові змінні, поділені на відомі σ_i . Позначення β_0^* та β_1^* використовуються для того, щоб відрізнити їх від звичайних параметрів β_0 та β_1 , отриманих методом найменших квадратів.

У чому полягає мета трансформації першої моделі? Тепер дисперсія трансформованої помилки ε^* є постійною величиною, тобто для моделі (5.30) зберігається припущення про гомоскедастичність, і ми переходимо до класичної регресійної моделі. Справді,

$$\text{var}(\varepsilon_i^*) = E(\varepsilon_i^*)^2 = E\left(\frac{\varepsilon_i}{\sigma_i}\right)^2 = \frac{1}{\sigma_i^2} E(\varepsilon_i^2). \quad (5.31)$$

Ми припускали, що σ_i^2 — відомі, звідси маємо:

$$\frac{1}{\sigma_i^2} E(\varepsilon_i^2) = \frac{1}{\sigma_i^2} \sigma_i^2 = 1. \quad (5.32)$$

Тепер оцінки параметрів будуть BLUE-оцінками, відомими як УНК-оцінки. Процедура знаходження невідомих параметрів трансформованої моделі називається методом узагальнених найменших квадратів (УНК).

Як же отримати оцінки для β_0^* та β_1^* ? Для моделі (5.30) запишемо вибірковий аналог:

$$y_i^* = b_0^* x_{0i}^* + b_1^* x_i^* + e_i^*. \quad (5.33)$$

Для того, щоб знайти невідомі параметри за методом узагальнених найменших квадратів, мінімізуємо:

$$\sum_{i=1}^n e_i^* = (y_i^* - b_0^* x_{0i}^* + b_1^* x_i^*)^2, \quad (5.34)$$

або

$$\sum_{i=1}^n \left(\frac{e_i}{\sigma_i}\right)^2 = \left[\left(\frac{y_i}{\sigma_i}\right) - b_0^* \left(\frac{x_{0i}}{\sigma_i}\right) - b_1^* \left(\frac{x_i}{\sigma_i}\right) \right]^2. \quad (5.35)$$

Розрахунки наводити не будемо і запишемо лише кінцевий результат. УНК-оцінка для β_1^* дорівнює:

$$b_1^* = \frac{\left(\sum_{i=1}^n \gamma_i\right) \left(\sum_{i=1}^n \gamma_i x_i y_i\right) - \left(\sum_{i=1}^n \gamma_i x_i\right) \left(\sum_{i=1}^n \gamma_i y_i\right)}{\left(\sum_{i=1}^n \gamma_i\right) \left(\sum_{i=1}^n \gamma_i x_i^2\right) - \left(\sum_{i=1}^n \gamma_i x_i\right)^2}, \quad (5.36)$$

дисперсія параметра дорівнює:

$$\text{var}(b_1^*) = \frac{\left(\sum_{i=1}^n \gamma_i\right)}{\left(\sum_{i=1}^n \gamma_i\right) \left(\sum_{i=1}^n \gamma_i x_i^2\right) - \left(\sum_{i=1}^n \gamma_i x_i\right)^2}, \quad (5.37)$$

де $\gamma_i = 1 / \sigma_i^2$.

Різниця між звичайним та узагальненим методом найменших квадратів

За методом звичайних найменших квадратів невідомі параметри знаходяться шляхом мінімізації суми квадратів відхилень фактичних значень від теоретичних. Для простої лінійної регресії маємо:

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2 \rightarrow \min.$$

В узагальненому методі найменших квадратів мінімізується вираз (5.35), який можна переписати у вигляді:

$$\sum_{i=1}^n \gamma_i e_i^2 = \sum_{i=1}^n \gamma_i (y_i - b_0^* - b_1^* x_i)^2, \quad (5.38)$$

де $\gamma_i = 1 / \sigma_i^2$ — вагові коефіцієнти.

Тобто в узагальненому методі найменших квадратів мінімізуємо зважену суму квадратів відхилень з вагами, обернено пропорційними до σ_i . Спостереження з більшою σ_i отримують у цьому методі відповідно менше значення ваги, і навпаки, спостереження з меншою σ_i отримують пропорційно більшу вагу при мінімізації суми квадратів відхилень. У літературі вираз (5.38) ще відомий як зважені найменші квадрати (ЗНК), а оцінки параметрів, отримані за цим методом, — як ЗНК-оцінки. У контексті гетероскедастичності можна користуватися обома термінами: і ЗНК, і УНК.

Наслідки гетероскедастичності при побудові інтервалів довіри параметрів

Ми вже показували, що оцінки параметрів, знайдені за методом узагальнених найменших квадратів, є лінійними оцінками без відхилень за умов гетероскедастичності. Крім того, оцінки, знайдені за узагальненим методом, мають ще й найменшу дисперсію, тобто вони є BLUE-оцінками.

Тепер спробуємо відповісти на запитання, що буде з інтервалами довіри та F -тестами, якщо ми будемо знаходити параметри методом звичайних найменших квадратів, незважаючи на гетероскедастичність?

Наприклад, дисперсія параметра b_1 , знайденого за методом звичайних найменших квадратів, в умовах гетероскедастичності матиме вигляд:

$$\text{var}(b_1) = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \sigma_i^2}{\left(\sum_{i=1}^n (x_i - \bar{x})^2 \right)^2}. \quad (5.39)$$

Вона, звичайно, відрізняється від дисперсії того самого параметра за умов гомоскедастичності:

$$\text{var}^*(b_1) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^n (x_i - \bar{x})^2}. \quad (5.40)$$

Звичайно, якщо $\sigma_i^2 = \sigma^2$ для кожного i , то формули (5.39) та (5.40) будуть ідентичними.

Припустимо, що дисперсії σ_i^2 відомі. Чи можемо ми побудувати інтервали довіри та перевірити гіпотези за допомогою звичайних t та F -тестів? Ні, тому що $\text{var}^*(b_1) < \text{var}(b_1)$, тобто інтервали довіри, побудовані з урахуванням дисперсії (5.39), будуть обов'язково більшими. Як наслідок, t та F -тести дають неточні результати за рахунок того, що $\text{var}(b_1)$ надмірно велика; параметр стає статистично незначимим (t -значення менше за критичне).

Стандартні пакети прикладних програм звичайно дають для дисперсії формулу (5.40), якщо не використовуються спеціальні методи. Дисперсія b_1 в (5.40) є безпомилковою оцінкою дисперсії b_1 , наведеної в (5.39), тобто вона в середньому переоцінює останню. Причиною помилки є те, що $\hat{\sigma}^2 = \sum_{i=1}^n e_i^2 / (n - 2)$ — оцінка дисперсії σ^2 не є вже безпомилковою оцінкою в умовах гетероскедастичності. Тому ми вже не можемо спиратися на звичайно обчислені інтервали довіри і на t та F -тести.

Іншими словами, якщо, незважаючи на гетероскедастичність, використовувати звичайні процедури перевірки гіпотез, то висновки можуть бути неправильними.

Стає зрозумілим, що потенційно гетероскедастичність є серйозною проблемою. Досліднику потрібно знати, чи є вона, чи її немає. Коли з'ясована її наявність, можна використовувати метод зважених (узагальнених) найменших квадратів чи застосовувати якийсь інший метод.

5.2.5. Тестування наявності гетероскедастичності

Отож перед нами постає важливе питання: як у конкретній ситуації можна довідатись про гетероскедастичність?

Як і у випадку мультиколінеарності, єдиних правил її виявлення немає, а є різноманітні тести. Розглянемо найпростіші з них за змістом та за розрахунками.

Аналіз змісту проблеми

Інколи при проведенні економетричних досліджень гетероскедастичність вгадується інтуїтивно або висувається як абсолютне припущен-

ня. Попередній аналіз проблеми, що вивчається, може навести на думку про наявність гетероскедастичності. Наприклад, при вивченні бюджету сім'ї можна помітити, що дисперсія залишків зростає відповідно до зростання доходу. При зведеному аналізі діяльності різних за розміром фірм також можна очікувати гетероскедастичність. І таких прикладів багато. Отже, перший крок до вияву гетероскедастичності — глибокий аналіз досліджуваної проблеми.

Графічний аналіз

Досить наочним та простим методом тестування припущення про наявність гетероскедастичності є графічний метод. Не завжди дослідник володіє необхідним для аналізу проблеми емпіричним матеріалом. Крім того, його висновки щодо наявності або відсутності гетероскедастичності носять суб'єктивний характер, і в цих умовах на допомогу приходять графіки.

Спочатку необхідно зробити аналіз регресійної моделі на базі припущення про відсутність гетероскедастичності, а потім за допомогою дослідження квадратів залишків e_i^2 з'ясувати, чи мають вони якусь систематичність. Звичайно, e_i^2 — це тільки оцінки невідомих ε_i^2 , але вони можуть успішно використовуватися, особливо при великих вибірках. Досліджуючи e_i^2 , можна отримати такі види графіків (мал. 5.6 (а—д)).

На мал.5.6 (а—д) квадрати залишків розсіяні навколо оціненої лінії регресії. Потрібно з'ясувати, чи оцінене середнє значення y систематично пов'язано з квадратом залишків. На мал. 5.6, а зображена ситуація, при якій систематичного зв'язку немає, отже, гетероскедастичності немає. На мал. 5.6, з показано лінійний зв'язок, а на 5.4, в — квадратичну залежність.

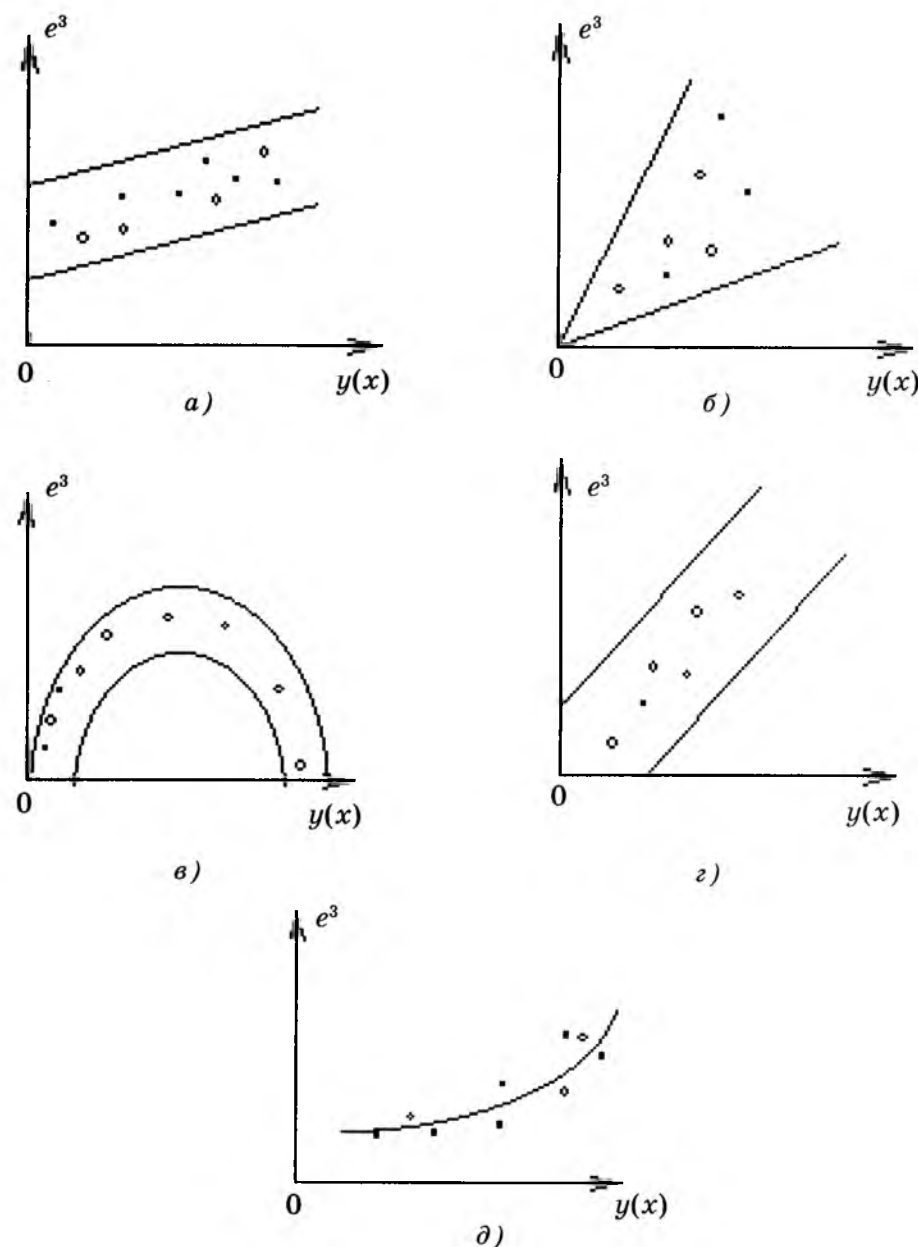
Завдяки графічному методу можна проаналізувати залежність не тільки між y та e^2 , а й між будь-яким фактором x_i та e^2 . Графіки залишаються такими ж при наявності гетероскедастичності. Графічний аналіз допомагає не тільки виявити гетероскедастичність, а й зробити висновок про саму форму зв'язку, що особливо корисно при трансформації наявних даних для побудови моделі з гомоскедастичністю помилок. Далі ми ще повернемося до розгляду цієї проблеми.

Тест рангової кореляції Спірмана

Це найпростіший тест, який можна використовувати як до малих, так і до великих вибірок.

Спочатку запишемо коефіцієнт рангової кореляції Спірмана:

$$r_s = 1 - 6 \left[\frac{\sum_{i=1}^n d_i}{n(n^2 - 1)} \right], \quad (5.41)$$



Малюнок 5.6. Різні типи графіків квадратів залишків

де d_i — різниця між рангами, що приписуються двом характеристикам i -го об'єкта;

n — кількість об'єктів, що ранжуються.

Наведений коефіцієнт рангової кореляції може використовуватись для визначення гетероскедастичності таким чином.

Припустимо, $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$.

Етап 1. Побудувати регресію для даних y та x і розрахувати відхилення e_i .

Етап 2. Нехтуючи знаком e_i , тобто беручи абсолютні значення $|e_i|$, ранжуємо $|e_i|$ та x_i у зростаючому чи спадному порядку і підрачуємо коефіцієнт рангової кореляції Спірмана.

Етап 3. Перевіряємо значимість отриманого коефіцієнта рангової кореляції за t -критерієм Ст'юдента. Для цього побудуємо t -статистику:

$$t = \frac{r_s \sqrt{n-2}}{\sqrt{1-r_s^2}}. \quad (5.42)$$

де n — кількість спостережень та $df = (n-2)$ — кількість ступенів вільності. При даних ступенях вільності за таблицями Ст'юдента знаходимо $t_{кр}$. Якщо розраховане значення перевищує $t_{кр}$ ($t > t_{кр}$), це підтверджує гіпотезу про гетероскедастичність. Якщо $t < t_{кр}$, тоді в регресійній моделі правильним є припущення про гомоскедастичність.

Тест Голдфелда та Квондта¹

Цей тест застосовується до великих вибірок. Спостережень має бути хоча б удвічі більше, ніж оцінюваних параметрів. Тест припускає нормальний розподіл та незалежність випадкової величини ε_i .

Для застосування тесту сформулюємо нульову та альтернативну гіпотези. **Нульова гіпотеза** $H_0: \varepsilon_i$ гомоскедастична проти **альтернативної гіпотези** $H_1: \varepsilon_i$ гетероскедастична (із зростаючою дисперсією).

Тест Голдфелда та Квондта складається з декількох етапів.

Етап 1. Ранжуємо спостереження незалежної змінної x в порядку зростання або спаду значення. У разі багатofакторної регресії, коли ми маємо більше, ніж одну незалежну змінну (фактор), вибираємо одну з них і за нею проводимо ранжування. Якщо важко априорі визначити змінну x для ранжування, то по черзі проводимо ранжування за кожною змінною окремо і в кожному випадку застосовуємо тест Голдфелда та Квондта.

Етап 2. Задаємо величину C — кількість центральних спостережень за незалежними змінними x , які ми будемо виключати з подальшого ана-

¹ Goldfeld S.M., Quandt R.E. Some Tests for Homoscedasticity//Journal of American Statistical Association. — 1965. — P. 539—547.

лізу. Експерименти, які проводили Голдфелд і Квондт, показали, що для вибірок з кількістю спостережень більшою, ніж 30 ($n > 30$) оптимальна кількість центральних спостережень, не врахованих у тесті, становить приблизно чверть усіх спостережень, наприклад 8 для $n = 30$, 16 для $n = 60$. Залишок $(n - c)$ спостережень ділиться на дві підвибірки однакового розміру $((n - c)/2)$, одна з яких включає малі значення X , інша — великі.

Етап 3. Будуємо окрему регресію для кожної підвибірки, розраховуємо суму квадратів залишків і отримуємо:

$\sum e_1^2$ — сума квадратів залишків для підвибірки із малими значеннями x , із $[(n - c)/2] - k$ ступенями вільності, де k — загальна кількість оцінюваних параметрів у моделі;

$\sum e_2^2$ — відхилення від підвибірки із великими значеннями x із такою самою кількістю ступенів вільності, $[(n - c)/2] - k$.

Якщо кожен з цих сум поділити на кількість ступенів вільності, отримаємо оцінки дисперсій випадкової величини ε в двох підвибірках.

Відношення двох дисперсій

$$F^* = \frac{\sum e_2^2 / [(n-c)/2] - k}{\sum e_1^2 / [(n-c)/2] - k} = \frac{\sum e_2^2}{\sum e_1^2}$$

має F -розподіл (із $v_1 = v_2 = [(n - c)/2] - k = [(n - c - 2k)/2]$ ступенями вільності, де n — загальна кількість спостережень, c — центральні невраховані спостереження, k — кількість оцінюваних параметрів із кожної регресії).

Якщо дві дисперсії рівні (отже, й гомоскедастичні), значення F^* буде дорівнювати одиниці. Якщо дисперсії відрізняються, то F^* матиме більше значення (із умови тесту $\sum e_2^2 > \sum e_1^2$). F^* порівнюється з теоретичним значенням F із $v_1 = v_2 = \{n - c - 2k\}/2$ ступенями вільності для заданого рівня значимості. Теоретичне (отримане з F -таблиць) значення F є значенням, яке мало б F , якби справджувалася нульова гіпотеза, тобто ε були б гомоскедастичними. Якщо $F^* > F$, ми допускаємо гетероскедастичність (відкидаємо нульову гіпотезу щодо відсутності відмінності між дисперсіями ε у двох підвибірках). Якщо $F^* < F$, ми допускаємо, що ε гомоскедастичні (іншими словами, приймаємо нульову гіпотезу). Чим більше відношення F^* , тим більша гетероскедастичність ε .

Тест Глейзера¹

Тест Глейзера є розширенням тесту Парка (якого ми не наводимо). Він також здійснюється в декілька етапів. Проілюструємо його на прикладі простої лінійної регресії.

¹ Glejser H. A New Test for Heteroscedasticity// Journal of American Statistical Association. — 1969. — Vol.64. — P. 316 — 323.

Етап 1. Знаходимо невідомі параметри простої лінійної регресії методом найменших квадратів та обчислюємо помилки e_i для кожного окремого спостереження.

Етап 2. Будуємо регресію, яка пов'язує абсолютні значення знайдених на першому етапі помилок ($|e_i|$) з незалежною змінною x . Ми беремо абсолютні значення помилок, а не їхні справжні значення, бо $\sum ex = 0$, і тому не можливо буде підібрати регресію $e = f(x)$. Фактична форма цієї регресії звичайно не відома, тому до неї можна підбирати різні форми кривих. Глейзер пропонував такі залежності:

$$\begin{aligned} |e_i| &= \beta_0 + \beta_1 x_i^2 + u_i, \\ |e_i| &= \beta_0 + \beta_1 x_i^{-1} + u_i = \beta_0 + \beta_1 / x_i + u_i, \\ |e_i| &= \beta_0 + \beta_1 x_i^{1/2} + u_i = \beta_0 + \beta_1 \sqrt{x_i} + u_i, \\ |e_i| &= \sqrt{\beta_0 + \beta_1 x_i} + u_i, \\ |e_i| &= \sqrt{\beta_0 + \beta_1 x_i^2} + u_i. \end{aligned}$$

Обираємо ту регресію, яка найкраще підходить з огляду на коефіцієнт кореляції та середні квадратичні відхилення параметрів b_0 та b_1 . (Зверніть увагу, що коли $b_0 = 0$ та $b_1 \neq 0$, така ситуація називається “чиста гетероскедастичність”; якщо $b_0 \neq 0$ та $b_1 \neq 0$, цей випадок називається “змішана гетероскедастичність”). Гетероскедастичність визначається в світлі статистичної значимості параметрів b_0 та b_1 , тобто ми виконуємо будь-який стандартний тест перевірки на значимість параметрів, і якщо вони значно відрізняються від нуля, то e_i є гетероскедастичними. Перевага тесту Глейзера в тому, що він дає також інформацію про форму гетероскедастичності, тобто про спосіб, яким пов'язані $\sigma_{e_i}^2$ та x_i . Ця інформація є важливою, як ми зараз побачимо, для “корекції” гетероскедастичності. Зазначимо, що у разі багатofакторної регресії на етапі 1 знаходимо помилки e_i для регресії, що вміщує всі фактори. На етапі 2 будуємо залежності між абсолютними величинами знайдених помилок та залежною змінною y .

Слід зазначити, що деякі статистики надають перевагу тестам рангової кореляції Спірмана і Голдфелда та Квондта перед тестом Глейзера для визначення гетероскедастичності. Якщо якимось із цих тестів виявлено гетероскедастичність, тоді можна експериментувати з функцією Глейзера ($|e| = f(x)$) з метою вирішення, які зміни початкових даних необхідні, щоб подолати гетероскедастичність.

5.2.6. Вилучення гетероскедастичності

Коли на базі будь-якого тесту встановлено гетероскедастичність, то для її вилучення змінюють початкову модель таким чином, щоб помилки мали постійну дисперсію. Далі невідомі параметри трансформованої моделі розраховуються за методом найменших квадратів. Трансформація моделі зводиться до зміни первісної форми моделі. Яким чином це проводиться, залежить від специфічної форми гетероскедастичності, тобто від форми залежності між дисперсією $\sigma_{e_i}^2$ та значеннями незалежних змінних:

$$\sigma_{e_i}^2 = f(x_i).$$

Розглянемо можливі випадки трансформації моделі на прикладі простої лінійної регресії. Припустимо, що ми маємо початкову модель $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, де випадкова величина ε_i гетероскедастична, але відповідає всім іншим класичним припущенням лінійної регресії.

Випадок 1. Припустимо, гетероскедастичність має форму $E(\varepsilon_i)^2 = \sigma_{\varepsilon_i}^2 = k^2 x^2$ (де k — скінченна константа), тобто дисперсія ε зростає пропорційно до x^2 . Виражаючи коефіцієнт пропорційності k^2 , маємо $k^2 = \sigma_{\varepsilon_i}^2 / x^2$. Це означає, що трансформація моделі полягає у діленні початкової моделі на $\sqrt{x^2} = x$. Трансформована таким чином модель має вигляд:

$$\frac{y_i}{x_i} = \frac{\beta_0}{x_i} + \beta_1 + \frac{\varepsilon_i}{x_i}.$$

Зверніть увагу на місце параметрів моделі: параметр при змінній $1/x_i$ у трансформованій моделі є перетином у початковій моделі, тоді як перетин трансформованої моделі є нахилом початкової.

Нове змінене значення випадкової величини $\frac{\varepsilon_i}{x_i}$ є гомоскедастичним, оскільки

$$E\left(\frac{\varepsilon_i}{x_i}\right)^2 = \frac{1}{x_i^2} E(\varepsilon_i^2) = \frac{1}{x_i^2} \sigma_{\varepsilon_i}^2$$

(за одним із припущень, X — множина фіксованих значень в усіх вибірках).

Однак у нашому прикладі припускалося, що $\sigma_{\varepsilon_i}^2 = k^2 x_i^2$,

тому $E\left(\frac{\varepsilon_i}{x_i}\right)^2 = \frac{1}{x_i^2} k^2 x_i^2 = k^2$. Отже, нова випадкова змінна має скінченну постійну дисперсію (що дорівнює k^2), і, таким чином, ми можемо застосова-

ти класичний метод найменших квадратів до розрахунку невідомих параметрів трансформованої моделі.

Випадок 2. Припустимо, що гетероскедастичність має форму

$$E(\varepsilon_i)^2 = \sigma_{\varepsilon_i}^2 = k^2 x_i.$$

Допустима трансформація полягає в діленні початкової моделі на \sqrt{x} , тобто трансформована модель має вигляд:

$$\begin{aligned} \frac{y}{\sqrt{x}} &= \frac{\beta_0}{\sqrt{x}} + \beta_1 \frac{x}{\sqrt{x}} + \frac{\varepsilon}{\sqrt{x}}; \\ (yx^{-1/2}) &= \beta_0 x^{-1/2} + \beta_1 x^{-1/2} + \frac{\varepsilon}{\sqrt{x}}. \end{aligned}$$

Для трансформованої моделі випадкова величина $\frac{\varepsilon}{\sqrt{x}}$ гомоскедастична із сталою дисперсією k^2 . Справді, маємо:

$$E\left(\frac{\varepsilon_i}{\sqrt{x_i}}\right)^2 = \frac{1}{x_i} E(\varepsilon_i^2) = \frac{1}{x_i} (\sigma_{\varepsilon_i}^2) = k^2,$$

оскільки в прикладі припускалося, що $\sigma_{\varepsilon_i}^2 = k^2 x$. Отже, виконавши вказані вище перетворення, ми виключили гетероскедастичність.

Випадок 3. Припустимо, що форма гетероскедастичності

$$E(\varepsilon_i)^2 = \sigma_{\varepsilon_i}^2 = k^2 (a_0 + a_1 x_i)^2.$$

Допустима трансформація полягає в діленні початкової моделі на $\sqrt{(a_0 + a_1 x)^2} = (a_0 + a_1 x)$, тобто

$$\frac{y_i}{a_0 + a_1 x_i} = \beta_0 \frac{1}{a_0 + a_1 x_i} + \beta_1 \frac{x_i}{a_0 + a_1 x_i} + \frac{\varepsilon_i}{a_0 + a_1 x_i}.$$

Нова випадкова величина є гомоскедастичною із сталою дисперсією, рівною k^2 . Справді, маємо:

$$E\left[\frac{\varepsilon_i}{a_0 + a_1 x_i}\right]^2 = \frac{1}{(a_0 + a_1 x_i)^2} E(\varepsilon_i^2) = \frac{k^2}{(a_0 + a_1 x_i)^2} (a_0 + a_1 x_i)^2 = k^2.$$

Загальний випадок

Коли гетероскедастичність має вираз

$$E(\varepsilon_i)^2 = \sigma_{\varepsilon_i}^2 = k^2 f(x_i),$$

де k — скінченна константа і $f(x_i)$ — функція від x , трансформація початкової початкової моделі здійснюється шляхом її ділення на $\sqrt{f(x_i)}$. Пояснимо, чому саме така трансформація необхідна. Взагалі, наведена вище трансформація еквівалентна застосуванню методу зважених найменших квадратів (ЗНК), який є особливим випадком методу узагальнених найменших квадратів (УНК). Опишемо метод зважених найменших квадратів (ЗНК), який забезпечує обґрунтування вищенаведеної трансформації.

У методі простих найменших квадратів ми мінімізуємо просту суму квадратів відхилень, $\sum e_i^2 = \sum (y_i - b_0 - b_1 x_i)^2$, у якій кожне відхилення має однакову вагу (сума ваг = 1). Тобто сума $\sum e_i^2$ є незваженою сумою квадратних відхилень, у якій припускається, що ε_i , оцінені за допомогою e_i . Хоча, якщо дисперсія ε_i не є сталою, а, скажімо, зростає із збільшенням x , зрозуміло, що більша дисперсія спостереження дає менш точну вказівку на те, де проходить правильна регресійна лінія. Тому здається правдоподібним надавати меншої уваги цим спостереженням у підборі лінії регресії порівняно з іншими спостереженнями. Цього можна досягнути наданням різної ваги кожній ε_i (чи її оцінці). Інколи доцільно використовувати як вагу частку $1/\sigma_{\varepsilon_i}^2$, тобто поділити кожне відхилення на дисперсію випадкової величини. У випадку, коли випадкова величина ε_i є великою, її дисперсія $\sigma_{\varepsilon_i}^2$ є теж великою і вага $1/\sigma_{\varepsilon_i}^2$ буде малою; таким чином, великим помилкам надаються малі ваги. Отже, замість мінімізації простої суми квадратів відхилень ми мінімізуємо зважену суму квадратів відхилень:

$$\sum \frac{e_i^2}{\sigma_{\varepsilon_i}^2} = \sum \frac{1}{\sigma_{\varepsilon_i}^2} (y_i - b_0 - b_1 x_i)^2.$$

Такий метод і називається методом зважених найменших квадратів (ЗНК). Прирівнявши часткові похідні зваженої суми квадратів до нуля і розв'язавши систему рівнянь, отримаємо формули для знаходження невідомих параметрів b_0 та b_1 , що можливо при відомій дисперсії $\sigma_{\varepsilon_i}^2$. Але на практиці ця дисперсія, як правило, невідома, і процедура її обчислення

може бути досить складною. Що ж робити, коли дисперсія невідома? Покажемо, що в такому випадку наведені вище трансформації початкової моделі аналогічні застосуванню методу зважених найменших квадратів для початкової моделі. Можна довести, що мінімізація зваженої суми квадратів відхилень виводить аналогічні формули для оцінок невідомих параметрів початкової моделі, як і застосування простого методу найменших квадратів до трансформованої моделі.

Доведення. Ми будемо використовувати простий випадок, у якому гетероскедастичність виражається як $\sigma_{\varepsilon_i}^2 = k^2 x_i^2$, тому трансформація

$\frac{y_i}{x_i} = b_0 \frac{1}{x_i} + b_1 + \frac{\varepsilon_i}{x_i}$, де трансформоване відхилення є гомоскедастичним із дисперсією k^2 .

Введемо позначення $v = \varepsilon / x$. Метод простих найменших квадратів (МНК) щодо трансформованої моделі полягає в знаходженні таких значень невідомих параметрів b_0 та b_1 , при яких мінімізується така сума квадратів відхилень:

$$\sum \hat{v}_i^2 = \sum \left(\frac{y_i}{x_i} - b_0 \frac{1}{x_i} - b_1 \right)^2 = \sum \frac{1}{x_i^2} (y_i - b_0 - b_1 x_i)^2.$$

Метод зважених найменших квадратів передбачає вибір b_0 та b_1 , які б мінімізували зважену суму квадратів відхилень:

$$\sum \frac{e_i^2}{\sigma_{\varepsilon_i}^2} = \sum \frac{1}{\sigma_{\varepsilon_i}^2} (y_i - b_0 - b_1 x_i)^2 = \sum \frac{1}{k^2 x_i^2} (y_i - b_0 - b_1 x_i)^2,$$

яку треба мінімізувати по відношенню до b_0 та b_1 . Обидві суми квадратів відхилень відрізняються тільки сталим значенням $1/k^2$. Отже, значення b_0 та b_1 , що мінімізують першу суму (метод МНК, застосований до трансформованої моделі), такі ж, як і ті, що мінімізують зважену суму квадратів відхилень (ЗНК). (Згадаємо, що мінімізація функції $f(z)$ буде й мінімізацією $kf(z)$, де k — будь-яка константа). Тому метод зважених найменших квадратів, що застосовується до початкової моделі, дає такі ж результати, як і метод найменших квадратів, який застосовується до трансформованої моделі.

5.2.6.1. Ефективність оцінок трансформованої моделі

Як було показано вище, тест, запропонований Глейзером, дає інформацію про форму гетероскедастичності. Експериментуючи з різними формами функції $|e| = f(x)$, ми можемо обрати найкращу, а потім перейти до трансформації початкової моделі, як описано раніше.

Можна довести, що оцінки трансформованої моделі мають меншу дисперсію (є ефективнішими), ніж оцінки, отримані із застосуванням методу найменших квадратів до початкової моделі.

Проілюструємо цей результат на прикладі простої лінійної регресії, припускаючи, що гетероскедастичність, наприклад, має вигляд: $\sigma_{\varepsilon_i}^2 = k^2 x_i^2$.

Нехай початкова (гетероскедастична) модель має вигляд:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \text{ із } E(\varepsilon_i^2) = k^2 x_i^2.$$

Застосуємо метод найменших квадратів для розрахунку невідомих параметрів гетероскедастичної моделі та розрахуємо дисперсію параметра b_1 (розрахунки наводити не будемо):

$$\begin{aligned} \text{var}(b_1) &= E(b_1 - \beta_1)^2 = E\left(\frac{\sum \tilde{x}_i \varepsilon_i}{\sum \tilde{x}_i^2}\right)^2 \\ &= E\left(\frac{\sum \tilde{x}_i^2 \tilde{\varepsilon}_i^2 + 2 \sum \tilde{x}_i \tilde{x}_j \varepsilon_i \varepsilon_j}{(\sum \tilde{x}_i^2)^2}\right) = \frac{\sum \tilde{x}_i^2 E(\varepsilon_i^2)}{(\sum \tilde{x}_i^2)^2} = k^2 \frac{\sum \tilde{x}_i^2 \tilde{x}_j^2}{(\sum \tilde{x}_i^2)^2} \end{aligned}$$

за умови припущення $E(\varepsilon_i^2) = k^2 x_i^2$ та $\tilde{x}_i^2 = (x_i - \bar{x})^2$.

Трансформована версія початкової моделі: $\frac{y_i}{x_i} = \beta_0 \frac{1}{x_i} + \beta_1 + v_i$,

де $v_i = \varepsilon_i / x_i$ із сталою дисперсією $\sigma_v^2 = k^2$.

Застосовуючи метод найменших квадратів, маємо:

$$\hat{b}_1 = (\overline{y/x}) - \hat{b}_0 (\overline{1/x})$$

та

$$\text{var}(\hat{b}_1) = \sigma_v^2 \frac{\sum (1/x_i)^2}{n \sum \left\{ (1/x_i) - (\overline{1/x}) \right\}^2} = k^2 \frac{\sum (1/x_i)^2}{n \sum \left\{ (1/x_i) - (\overline{1/x}) \right\}^2}.$$

Зрозуміло, що $\text{var}(\hat{b}) < \text{var}(b)$.

Слід звернути увагу, що гетероскедастичність може траплятися через невраховані змінні, тобто через погану специфікацію моделі. У цьому випадку можливим рішенням є включення неврахованих змінних у модель. Сліпе застосування трансформації, наведеної вище, зробить гомоскедастичною випадкову змінну, але оцінки параметрів можуть залишатися неправильними через неврахування важливих факторів. З економічної теорії відомо, що, наприклад, у функції заощаджень гетероскедастичність може виникати через зміни в економічній політиці (монетарна політика, податкові реформи, знецінення національної валюти). У цьому разі рішенням буде врахування у функції певних факторів, які б відбивали зміни в політиці уряду.

5.2.7. Ілюстративний приклад. Тест Голдфелда та Квондта

Для ілюстрації виявлення та усунення гетероскедастичності розглянемо приклад про визначення залежності між заощадженнями на душу населення та доходом¹. Дані наведені в табл. 5.3.

Припустимо, що залежність між заощадженнями та доходом на душу населення лінійна, і після знаходження параметрів за методом найменших квадратів отримаємо таку модель:

$$\hat{y} = -648.1 + 0.085x$$

(1181.1) (0.004) $R^2 = 0.912$.

Для перевірки наявності або відсутності гетероскедастичності в моделі скористаємося *тестом Голдфелда та Квондта*. Для цього проранжуємо значення незалежної змінної x у порядку зростання, виключимо дев'ять центральних значень. У результаті отримаємо дві підвибірки — одна з найбільшими значеннями x , інша — з найменшими (див. табл. 5.4).

Для кожної підвибірки даних побудуємо просту лінійну регресійну модель та розрахуємо параметри методом найменших квадратів. Після розрахунків отримаємо:

Підвибірка 1

$$\hat{y}_1 = -744.64 + 0.088x$$

(195.4) (0.016)

із $R^2 = 0.778$ та $\sum e_1^2 = 150867.9$.

Підвибірка 2

$$\hat{y}_2 = 1050.79 + 0.032x$$

(817.4) (0.025)

¹ Koutsoyiannis A. Theory of Econometrics. — 1973. — P. 187.

Таблиця 5.3

Заощадження та доход на душу населення

Період	Накопичення y	Доход x	Прогноз накопичень \hat{y}	Залишки $e_i = y_i - \hat{y}_i$	y/x	$1/x$ ($\times 10.0$)
1	264	8.777	94.98	169.02	30.079	1.139
2	105	9.210	131.64	-26.64	11.401	1.086
3	90	9.954	194.63	-104.63	9.042	1.005
4	131	10.508	241.53	-110.53	12.467	0.952
5	122	10.979	281.41	-159.41	11.112	0.911
6	107	11.912	360.40	-253.40	8.983	0.839
7	406	12.747	431.10	-25.10	31.851	0.785
8	503	13.499	494.77	8.23	37.262	0.741
9	431	14.269	559.96	-128.96	30.205	0.701
10	588	15.522	666.05	-78.05	37.882	0.644
11	898	16.730	768.32	129.68	53.676	0.598
12	950	17.663	847.31	102.69	53.785	0.566
13	779	18.575	924.53	-145.53	41.938	0.538
14	819	19.635	1014.27	-195.27	41.711	0.509
15	1222	21.163	1143.64	78.36	57.742	0.473
16	1702	22.880	1289.01	412.99	74.388	0.437
17	1578	24.127	1394.58	183.42	65.404	0.414
18	1654	25.604	1519.63	134.37	64.599	0.391
19	1400	26.500	1595.50	-195.50	52.830	0.377
20	1829	27.670	1694.55	134.45	66.100	0.361
21	2200	28.300	1747.89	452.11	77.739	0.353
22	2017	27.430	1674.23	342.77	73.533	0.365
23	2105	29.560	1854.57	250.43	71.211	0.338
24	1600	28.150	1735.19	-135.19	56.838	0.355
25	2250	32.100	2069.62	180.38	70.093	0.312
26	2420	32.500	2103.49	316.51	74.462	0.308
27	2570	35.250	2336.31	233.68	72.908	0.284
28	1720	33.500	2188.15	-468.15	51.343	0.299
29	1900	36.000	2399.81	-499.81	52.778	0.278
30	2100	36.200	2416.75	-316.75	58.011	0.276
31	2300	38.200	2586.08	-286.08	60.209	0.262

із $R^2 = 0.154$ та $\sum e_2^2 = 763760.5$.

За отриманими результатами обчислимо F -відношення Фішера:

$$F^* = \frac{\sum e_2^2}{\sum e_1^2} = \frac{763760.5}{150867.9} = 5.06.$$

Таблиця 5.4

Підвибірка 1			Підвибірка 2		
n_1	y	x_I	n_2	y	x_{II}
1	264	8.777	1	1829	27.670
2	105	9.210	2	1600	28.150
3	90	9.954	3	2020	28.300
4	131	10.508	4	2105	29.560
5	122	10.979	5	2250	32.100
6	107	11.912	6	2420	32.500
7	406	12.747	7	1720	33.500
8	503	13.499	8	2570	35.250
9	431	14.269	9	1900	36.000
10	588	15.522	10	2100	36.200
11	898	16.730	11	2300	38.200

Ступені вільності чисельника та знаменника для F -відношення будуть однаковими і дорівнюватимуть:

$$v_1 = v_2 = \frac{n - c - 2k}{2} = \frac{31 - 9 - (2)(2)}{2} = 9.$$

Задамо рівень значимості 5% та за таблицею F -розподілу Фішера для заданого рівня значимості і відповідних ступенів вільності знайдемо F -критичне. У нашому випадку воно дорівнює 3.18. Якщо F -розрахункове більше, ніж F -критичне ($F^* > F_{0.05}$), ми відкидаємо припущення про гомоскедастичність. Отже, наша модель буде гетероскедастичною.

Перевіримо цю модель ще за допомогою тесту рангової кореляції Спірмана.

Поставимо у відповідність до рангів x ранги значень помилок e та розрахуємо відповідні значення величини D (розрахунки наведено в табл. 5.5).

Коефіцієнт рангової кореляції, оцінений за даними табл. 5.5, дорівнює:

$$r' = 1 - \frac{6 \sum D^2}{n(n^2 - 1)} = 1 - \frac{6 \times 1.558}{31 \times (31^2 - 1)} \approx 0.686.$$

Отже, обидва тести виявляють гетероскедастичність у початковій моделі.

Для вилучення гетероскедастичності зробимо припущення щодо її форми. Нехай гетероскедастичність має форму

$$\sigma_{\varepsilon_i}^2 = k^2 x_i^2.$$

Таблиця 5.5

Ранг x	Ранг e	D	D^2
1	16	-15	225
2	3	-1	1
3	7	-4	16
4	8	-4	16
5	15	-10	100
6	23	-17	289
7	2	5	25
8	1	7	49
9	9	0	0
10	4	6	36
11	10	1	1
12	6	6	36
13	14	-1	1
14	19	-5	25
15	5	10	100
16	28	-12	144
17	18	-1	1
18	11	7	49
19	20	-1	1
20	12	9	81
21	29	-6	36
22	27	-7	49
23	22	2	4
24	13	9	81
25	17	8	64
26	25	1	1
27	21	7	49
28	30	-3	9
29	31	-2	4
30	26	4	16
31	24	7	49

У цьому разі допустима трансформація початкової моделі є такою:

$$\frac{y_i}{x_i} = b_0 \frac{1}{x_i} + b_1 \frac{\varepsilon_i}{x_i}.$$

Трансформовані змінні ($\frac{y_i}{x_i}$ та $\frac{1}{x_i}$) показані в останніх двох колонках табл. 5.4. Застосовуючи метод найменших квадратів до нових змінних, отримуємо таку трансформовану модель:

$$\frac{y_i}{x_i} = b_0 \frac{1}{x_i} + b_1 = 0.088 \frac{1}{x_i} - 0.072,$$

або

$$y^* = -0.072 + 0.088x^* \\ (0.007) (0.004) \\ R^2 = 0.0774.$$

Нагадаємо, що початкова модель мала вигляд:

$$y = -648.1 + 0.088x \quad R^2 = 0.912 \\ (118.1) (0.004).$$

Зверніть увагу, що R^2 трансформованої моделі менше, ніж R^2 початкової. Це через те, що індивідуальні спостереження у процесі трансформації набувають ваги.

5.2.8. Основні висновки щодо наявності гетероскедастичності в регресійній моделі

Одним із припущень класичної лінійної регресійної моделі є припущення про те, що випадкові величини ε_i мають однакову дисперсію. Якщо ця умова не задовольняється, то ми маємо випадок гетероскедастичності. Гетероскедастичність не впливає на значимість оцінок параметрів, розрахованих за методом найменших квадратів, але ці оцінки вже не будуть оцінками з мінімальною дисперсією, тобто вони більше вже не є BLUE-оцінками. BLUE-оцінки у разі гетероскедастичності забезпечуються методом зважених найменших квадратів.

Слід підкреслити важливу особливість: якщо за умов гетероскедастичності розрахувати невідомі параметри методом найменших квадратів, то дисперсії цих параметрів уже не можна розрахувати за відомими формулами. Так, наприклад, у разі простої лінійної регресії традиційна формула

дисперсії параметра b_1 має вигляд:

$$\hat{\sigma}_{b_1} = \frac{\sigma^2}{\sum \tilde{x}_i^2},$$

у той час як при гетероскедастичності вона дорівнює $\frac{\sum \tilde{x}_i^2 \sigma_i^2}{(\sum \tilde{x}_i^2)^2}$. Але за-

уважимо, що ця дисперсія вже не є мінімальною; отже, інтервали довіри, засновані на ній, занадто широкі і тест значимості менш потужний.

Ситуація може бути серйозною, якщо не враховувати гетероскедастичності і використовувати класичні формули оцінки дисперсій. Тоді, у випадку двох змінних, якщо брати $\sigma^2 / \sum \tilde{x}_i^2$ замість $(\sum \tilde{x}_i^2 \sigma_i^2) / (\sum \tilde{x}_i^2)^2$ (яка є неефективною), t - і F -тести значимості, засновані на оцінках дисперсій, будуть неправильними. Коротше, в цій ситуації звичайні процедури тестування мають сумнівну, якщо не найнижчу, цінність.

Якщо виявлена гетероскедастичність, а дисперсія випадкової величини σ_i^2 невідома, необхідно трансформувати початкову модель з метою вилучення гетероскедастичності. Якщо гетероскедастичні дисперсії σ_i^2 відомі, то невідомі параметри регресійної моделі розраховуються за методом зважених найменших квадратів. Знання σ_i^2 — взагалі рідкість. Тому на практиці найпоширенішим випадком розв'язку гетероскедастичних моделей залишається їх трансформація. У цьому розділі ми дослідили декілька широко використовуваних трансформацій і визначили їх специфічні властивості. При цьому ми постійно припускали, що випадкова змінна ε має нормальний розподіл. Припущення про нормальність необхідне для перевірки на значимість оцінок параметрів за статистичними тестами та для

побудови інтервалів довіри. Якщо припущення про нормальність порушується, то оцінки параметрів залишаються найкращими, але ми не можемо визначити їхню статистичну надійність за допомогою класичних тестів значимості (t , F і т.д.), тому що ці тести базуються на нормальному законі розподілу. Але якщо розподіл не є нормальним, ми можемо використати центральну граничну теорему, за якою навіть якщо розподіл не є нормальним, вибірковий розподіл математичного сподівання наближається до нормального, коли розмір вибірки прямує до нескінченності. На перший погляд може здатися, що й центральна гранична теорема не дуже допоможе, бо на практиці кількість спостережень не буває великою. Але численні прикладні дослідження показали, що загалом хороша апроксимація досягається і для досить малої кількості спостережень (10, 20 і т.д.). У кожному окремому економетричному дослідженні бажано встановити ймовірність припущення про нормальність. Але на практиці найчастіше ігнорують це припущення, маючи на увазі, що центральна гранична теорема правильна у будь-якому випадку, чи намагаються довести припущення про нормальність на апіорних підставах, вважаючи, наприклад, що випадкова величина головним чином абсорбує вплив численних неважливих факторів та нестійких елементів людської поведінки.

5.3. АВТОКОРЕЛЯЦІЯ

5.3.1. Природа автокореляції. Основні поняття та означення

Одним із припущень класичного регресійного аналізу є припущення про незалежність випадкових величин. Якщо це припущення порушується, то ми маємо справу з автокореляцією. Звичайно, важливо зрозуміти, що викликає автокореляцію, які її практичні та теоретичні наслідки, чи змінюються методи знаходження невідомих параметрів моделі в умовах автокореляції і, нарешті, чи є ефективні методи її тестування.

Спробуємо послідовно відповісти на всі ці запитання. Як уже зазначалося вище, в регресійній моделі автокореляція наявна у разі, коли випадкові величини залежні між собою, тобто: $E(\varepsilon_i \varepsilon_j) \neq 0, i \neq j$.

Потрібно розрізняти поняття автокореляції і серійної кореляції. Автокореляцією називається залежність між значеннями однієї вибірки з записанням в один лаг. Наприклад, якщо між значеннями однієї вибірки $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ та $\varepsilon_2, \varepsilon_3, \dots, \varepsilon_{p+1}$ є залежність, то маємо справу з автокореляцією, якщо така залежність є між значеннями двох різних вибірок $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ та w_2, w_3, \dots, w_{p+1} , то це свідчить про наявність серійної кореляції. Автокореляція може бути як позитивною, так і негативною. Графічно ці випадки відображено в розділі 2, мал.2.6. Автокореляція може ви-

никнути у зв'язку з інерційністю та циклічністю багатьох економічних процесів. Провокувати автокореляцію може і неправильно специфікована функціональна залежність у регресійних моделях та лагові запізнення в економічних процесах.

5.3.2. Тестування автокореляції

Найбільш відомим і поширеним тестом перевірки моделі на наявність кореляції між залишками є тест Дарбіна — Уотсона. На відміну від багатьох інших тестів, перевірка за тестом Дарбіна — Уотсона складається з декількох етапів і включає зони невизначеності.

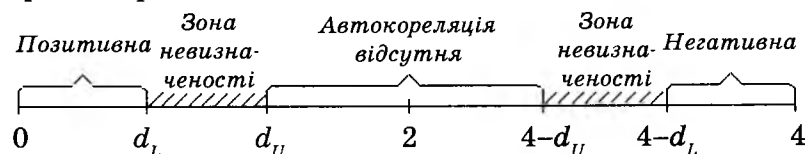
Розглянемо порядок тестування за критерієм Дарбіна — Уотсона.

1. На першому етапі розраховується значення d -статистики за формулою:

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}.$$

У теорії доведено, що значення d -статистики Дарбіна — Уотсона знаходяться в межах від 0 до 4.

2. Задаємо рівень значимості α та підраховуємо кількість факторів (k) у досліджуваній моделі. Припустимо $k = p$. За таблицею Дарбіна — Уотсона при заданому рівні значимості α , кількості факторів $k = p$ та кількості спостережень n , знаходимо два значення d_L та d_U . Якщо розраховане значення d -статистики знаходиться в проміжку від 0 до d_L ($0 < d < d_L$), то це свідчить про наявність позитивної автокореляції. Якщо значення d потрапляє в зону невизначеності, тобто набуває значення $d_L \leq d \leq d_U$, або $4 - d_U \leq d \leq 4 - d_L$, то ми не можемо зробити висновки ні про наявність, ні про відсутність автокореляції. Якщо $4 - d_L < d < 4$, то маємо негативну автокореляцію. Нарешті, якщо $d_U < d < 4 - d_U$, то автокореляції немає. Всі ці випадки проілюстровано на мал. 5.7.



Малюнок 5.7. Зони автокореляційного зв'язку за критерієм Дарбіна — Уотсона

Розглянемо приклад. Припустимо, для певної простої регресійної моделі, яка має один фактор ($k = 1$), кількість спостережень дорівнює $n = 20$ та розраховане значення d -статистики дорівнює 0.34. Прийmemo, що рівень значимості, тобто ризик відкинути правильну гіпотезу, дорівнює 5%. За таблицею Дарбіна — Уотсона при $k = 1$ та $n = 20$ знаходимо $d_L = 1.20$;

$d_U = 1.41$. Відповідно відкидаємо гіпотезу про відсутність автокореляції та приймаємо гіпотезу про наявність позитивної автокореляції.

5.3.3. Оцінка параметрів регресійної моделі при наявності автокореляції

Чи впливає на оцінку параметрів наявність автокореляції? І якщо впливає, то яким чином цього впливу уникнути? Спробуємо відповісти на ці запитання. Для спрощення викладок розглянемо просту лінійну регресійну модель:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t, \quad t = \overline{1, n}. \quad (5.3.1)$$

Припустимо, що всі класичні припущення виконуються, крім припущення про незалежність випадкових величин, тобто:

$$E(\varepsilon_t \varepsilon_{t+j}) \neq 0, \quad (j \neq 0).$$

Припустимо також, що між випадковими величинами є лінійна залежність:

$$\varepsilon_t = \rho \varepsilon_{t-1} + u_t, \quad -1 < \rho < 1, \quad (5.3.2)$$

де ρ — коефіцієнт автокореляції; u_t — випадкова величина, для якої виконуються всі класичні припущення методу найменших квадратів:

$$E(u_t) = 0; \quad \text{var}(u_t^2) = \sigma^2; \quad \text{cov}(u_t, u_{t+s}) = 0; \quad s \neq 0. \quad (5.3.3)$$

Модель (5.3.3) відома під назвою авторегресивна модель Маркова першого порядку (AR(1)), або авторегресивна лагова модель (авторегресивні лагові моделі детальніше буде розглянуто в наступному параграфі). У такій інтерпретації коефіцієнт автоковаріації ρ ще називається коефіцієнтом автокореляції першого порядку, або коефіцієнтом автокореляції з лагом 1.

Ми не розглядаємо використання для опису залежності між залишками авторегресивних моделей вищих порядків, бо це суттєво не впливає на подальші викладки.

Отже, для того, щоб дослідити вплив автокореляції на оцінку невідомих параметрів, повернемося до моделі (5.3.1). Розглянемо для спрощення тільки оцінку параметра β_1 , яка за методом найменших квадратів знаходиться за формулою:

$$b_1 = \frac{\sum_{t=1}^n (x_t - \bar{x})(y_t - \bar{y})}{\sum_{t=1}^n (x_t - \bar{x})^2} = \frac{\sum_{t=1}^n \tilde{x}_t \tilde{y}_t}{\sum_{t=1}^n \tilde{x}_t^2}.$$

Нагадаємо, що дисперсія параметра b_1 при відсутності автокореляції дорівнює:

$$\text{var}(b_1) = \frac{\sigma^2}{\sum_{t=1}^n \tilde{x}_t^2}. \quad (5.3.4)$$

За наявності автокореляції, наприклад типу AR(1), дисперсія параметра b_1 змінює своє значення (доведення цього факту ми не наводимо):

$$\text{var}(b_1)_{AR(1)} = \frac{\sigma^2}{\sum_{t=1}^n \tilde{x}_t^2} + \frac{2\sigma^2}{\sum_{t=1}^n \tilde{x}_t^2} \left[\rho \frac{\sum_{t=1}^n \tilde{x}_t \tilde{x}_{t+1}}{\sum_{t=1}^n \tilde{x}_t^2} + \rho^2 \frac{\sum_{t=1}^{n-2} \tilde{x}_t \tilde{x}_{t+2}}{\sum_{t=1}^n \tilde{x}_t^2} + \dots + \rho^{m-1} \frac{\tilde{x}_t \tilde{x}_n}{\sum_{t=1}^n \tilde{x}_t^2} \right]. \quad (5.3.5)$$

Звичайно, якщо $\rho = 0$, то обидві формули будуть однаковими, але при наявності автокореляції дисперсія параметра b_1 відрізнятиметься від значення дисперсії за відсутності автокореляції. Розглянемо, як цей факт буде впливати на оцінки параметрів. Чи залишатимуться у такому разі оцінки параметрів BLUE-оцінками?

На жаль, це не так. Теоретично доведено (доказу ми не наводимо), що при наявності автокореляції оцінки параметрів, залишаючись лінійними та незміщеними, не будуть мати найменшу дисперсію, тобто не будуть ефективними, а значить, і BLUE-оцінками. Якщо це так, то постає інше запитання: чи є ще якийсь метод оцінювання, який в умовах автокореляції дає BLUE-оцінки? Можна показати, що при наявності автокореляції в моделях простої лінійної регресії¹ BLUE-оцінкою параметра β_1 буде така:

$$\hat{b}_1^* = \frac{\sum_{t=2}^n (x_t - \rho x_{t-1})(y_t - \rho y_{t-1})}{\sum_{t=2}^n (x_t - \rho x_{t-1})^2} + A, \quad (5.3.6)$$

де A — коригуючий параметр.

Дисперсія параметра знаходиться за формулою

$$\text{var}(b_1^*) = \frac{\sigma^2}{\sum_{t=2}^n (x_t - \rho x_{t-1})^2} + D,$$

де D — коригуючий параметр.

Саме такі оцінки, як для простої лінійної регресії, так і для багатофакторної, дає метод узагальнених найменших квадратів (УНК), який уже було розглянуто в параграфі про гетероскедастичність. Таким чином, за наявності автокореляції перевагу при оцінці невідомих параметрів слід віддати методу узагальнених найменших квадратів, а не методу найменших квадратів.

Якщо все ж таки використовувати метод найменших квадратів в умовах автокореляції, то це призведе до таких наслідків.

1. Оцінка дисперсії випадкової величини часто переоцінює дійсну дисперсію і, як наслідок, маємо переоцінений коефіцієнт детермінації R^2 .

2. Дисперсія параметрів, наприклад $\text{var}(b_1)_{AR(1)}$, породжує помилки при використанні t - та F -тестів.

Висновки

Одним із припущень класичної лінійної регресії є припущення про незалежність випадкових величин. Якщо це припущення порушується, то наявна серійна кореляція або автокореляція.

Автокореляція може виникати з багатьох причин: по-перше, її викликає інерційність економічних процесів і, як наслідок, залежність між даними в часових рядах; по-друге, некоректно специфіковані моделі, маніпуляції з даними, введення лагових змінних.

При автокореляції небажана оцінка параметрів методом найменших квадратів, бо вона призводить до неефективних оцінок і неможливості застосування t - та F -тестів. Поширеним методом оцінки невідомих параметрів при наявності автокореляції є метод узагальнених найменших квадратів. Тестування автокореляції, як правило, проводиться за d -тестом Дарбіна — Уотсона, хоча є й інші не менш відомі тести.

¹ Загальний випадок розглянуто в кн.: Johnston J. Econometric Methods. — 3rd. ed. — New York: McGraw-Hill Book Company, 1984.

5.4. АВТОРЕГРЕСИВНІ І ДИСТРИБУТИВНО-ЛАГОВІ МОДЕЛІ

5.4.1. Природа авторегресивних моделей. Приклади практичного застосування авторегресивних моделей

У регресійному аналізі, якщо регресійна модель включає не лише поточні, а й попередні (лагові, або затримані) значення незалежних змінних (x), вона має назву *дистрибутивно-лагова модель*. У той же час, якщо до моделі включене одне або більше попередніх значень залежної змінної (y), вона має назву *авторегресивна модель*. Таким чином,

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \varepsilon_t$$

є *дистрибутивно-лаговою моделлю*, а

$$y_t = \alpha + \beta x_t + \gamma y_{t-1} + \varepsilon_t$$

є прикладом *авторегресивної моделі*. Такі авторегресивні моделі також відомі під назвою *динамічних моделей*, оскільки вони відображають часові зміни залежної змінної щодо її попереднього (попередніх) значень. Авторегресивні і дистрибутивно-лагові моделі широко використовуються в економетричному аналізі, і в цьому параграфі ми докладно розглянемо їх, щоб з'ясувати такі питання.

1. Яку роль відіграють лаги в економіці?
2. Яка причина цих лагів?
3. Чи є якимось теоретичне обґрунтування широкого використання лагових моделей в економетриці?
4. Як пов'язані між собою авторегресивні і дистрибутивно-лагові моделі?
5. Які статистичні проблеми можуть виникнути при оцінці параметрів таких моделей?

5.4.1.1. Приклади використання лагових моделей в економіці. Роль “часу” або “часового лагу” в економіці

В економіці рідко трапляється миттєва залежність змінної y (залежної змінної) від іншої незалежної змінної (змінних) x . Дуже часто значення y змінюється через невеликий проміжок часу після зміни значення x . Такий проміжок часу називається *часовим лагом*. Щоб проілюструвати природу часового лагу, розглянемо кілька прикладів.

Приклад 5.4.1. Функція споживання. Припустимо, людина отримала щорічне підвищення заробітної плати на 1200 гривень, припустимо також,

що це підвищення — постійне, зберігатиметься й протягом наступних років. Як буде впливати таке підвищення доходу на схему щорічних витрат на споживання цієї людини?

Загальноприйнята практика показує, що люди, отримавши таке підвищення доходу, не витрачають його одразу. Припустимо, що людина може вирішити підвищити свої витрати на споживання на 480 гривень у перший рік після підвищення доходу, на 360 гривень на другий рік і на 240 гривень на третій рік, заощаджуючи решту. На кінець третього року щорічні витрати на споживання зростуть на 1080 гривень. Ми можемо записати функцію споживання для цієї людини таким чином:

$$y_t = \text{const} + 0.4x_t + 0.3x_{t-1} + 0.2x_{t-2} + \varepsilon_t, \quad (5.4.1)$$

де y — витрати на споживання; x — дохід.

Функція, або модель (5.4.1) показує, що ефект підвищення доходу на 1200 гривень був розподілений на 3 роки згідно з наведеними вище припущеннями. Моделі, подібні до (5.4.1), називають *дистрибутивно-лаговими моделями*, оскільки ефект від певної причини (наприклад, підвищення доходу) був розподілений серед декількох часових періодів (лагів).

У загальному дистрибутивно-лагова модель має такий вигляд:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \beta_k x_{t-k} + \varepsilon_t. \quad (5.4.2)$$

Вона називається дистрибутивно-лаговою моделлю з кінцевим лагом в k періодів. Коефіцієнт β_0 відомий як *короткостроковий, або впливовий мультиплікатор*, тому що він показує вплив змінної x на зміну значення y в поточний період часу. Якщо зміна значення x триває, то $(\beta_0 + \beta_1)$ показує зміну y у другий період, $(\beta_0 + \beta_1 + \beta_2)$ — в третій і так далі. Ці часткові суми мають назву *проміжних інтервалів*. Нарешті, через k періодів отримаємо:

$$\sum_{i=1}^k \beta_i = \beta_0 + \beta_1 + \beta_2 + \dots + \beta_k = \beta, \quad (5.4.3)$$

що має назву *довгострокового, або загального дистрибутивно-лагового мультиплікатора*.

Якщо позначимо

$$\beta_i^* = \frac{\beta_i}{\sum \beta_i} = \frac{\beta_i}{\beta}, \quad (5.4.4)$$

то отримаємо “стандартизоване” β_i . Часткові суми стандартизованих β_i показують пропорцію довгострокового або загального впливу на певний часовий період.

Повертаючись до регресії споживання (5.4.1), ми бачимо, що короткостроковий мультиплікатор, який є нічим іншим як короткостроковою граничною схильністю до споживання (MPC), дорівнює 0.4, а довгостроковий мультиплікатор, який є довгостроковою схильністю до споживання, дорівнює $0.4 + 0.3 + 0.2 = 0.9$. Це значить, що після підвищення доходу на 1 гривню споживач підвищить свій рівень споживання приблизно на 0.4 гривні в рік підвищення, ще на 0.3 гривні на наступний рік і ще на 0.2 гривні на третій рік. Таким чином, довгостроковий вплив підвищення доходу на 1 гривню становить 0.9 гривні. Якщо ми поділимо кожне β_i на 0.9, то отримаємо, відповідно, 0.44, 0.33 і 0.23, які показують, що 44 відсотки загального впливу на y після одиначної зміни в x відчуватимуться негайно, 77 відсотків — через рік, і 100 відсотків — на кінець другого року.

Приклад 5.4.2. Зв'язок між грошима та цінами. Як вважають монетаристи, інфляція — це в основному монетарне явище в тому сенсі, що постійне збільшення рівня цін прямо залежить від збільшення грошової пропозиції, яка значно перевищує ту кількість грошей, яка реально потрібна *економічними одиницями*. Звичайно, цей зв'язок між рівнем інфляції і грошовою пропозицією не є миттєвим. Дослідження показали, що часовий лаг між ними сягає від 3 до 20 кварталів. Результати одного з таких досліджень наведено в роботі К. Карлсона¹. Вони базуються на періоді спостережень з січня 1955 по квітень 1969 рр. (див. табл. 5.6).

$$\dot{P} = -0.146 + \sum_{t=0}^{20} m_t \dot{M}_{-1}, \quad (0.395)$$

де \dot{P} = річна норма приросту дефлятора ВВП;

\dot{M} = річна норма приросту грошової пропозиції.

Дослідження показують, що ефект 1%-ної зміни грошової пропозиції відчувається протягом 20 кварталів. Довгостроковий вплив 1%-ної зміни в грошовій пропозиції на інфляцію дорівнює $\sum m_{t-1}$. До речі, оскільки P і M подані у відсотковій формі, m_t (β_i у загальній формі) показує еластичність P щодо M , яка є відсотковою зміною ціни y відповідь на 1%-не збільшення грошової пропозиції. Таким чином, $m_0=0.041$ означає, що при 1%-ному збільшенні грошової пропозиції короткострокова цінова еластичність становить 0.04 відсотка. Довгострокова еластичність становить 1.03 відсотка. Це означає, що в довгостроковому періоді збільшення грошової пропозиції на 1% спричинить майже такий самий відсоток збільшення цін. Тобто при 1%-ному збільшенні грошової пропозиції в довгостроковому періоді рівень інфляції збільшиться на 1%.

Приклад 5.4.3. Лаг між витратами на дослідження та продуктивністю праці. Рішення інвестувати певні дослідження та кінцевий результат інвестування, який приводить, наприклад, до підвищення продуктивності праці, відділені значним лагом. Крім того, цей лаг фактично складається з декількох лагів, таких як лаг між вкладанням коштів і часом, коли фактично почи-

¹ Carlson Keith M. The Lag from Money to Prices//Review. — Federal Reserve Bank of St. Louis. — 1980. — Oct. — Tabl.1.

Таблиця 5.6

	Коефіцієнт	t		Коефіцієнт	t		Коефіцієнт	t
m_0	0.041	1.276	m_8	0.048	3.249	m_{16}	0.069	3.943
m_1	0.034	1.538	m_9	0.054	3.783	m_{17}	0.062	3.712
m_2	0.030	1.903	m_{10}	0.059	4.305	m_{18}	0.053	3.511
m_3	0.029	2.171	m_{11}	0.065	4.673	m_{19}	0.039	3.338
m_4	0.030	2.235	m_{12}	0.069	4.795	m_{20}	0.022	3.191
m_5	0.033	2.294	m_{13}	0.072	4.694			
m_6	0.037	2.475	m_{14}	0.073	4.468	$\sum m_i$	1.031	7.870
m_7	0.042	2.798	m_{15}	0.072	4.202	Середній лаг	10.959	5.634
R^2	0.525							
S.E	1.066							
D.W.	2.00							

нають з'являтися нові ідеї і винаходи; лаг між винаходженням нового приладу або ідеї і його розвитком до стадії комерційного застосування (*виробництва*) і, нарешті, лаг, представлений процесом заміни старих приладів на кращі нові зразки.

Наведені вище приклади — це лише невеликі зразки використання лагів в економіці. Безсумнівно, кожен може сам навести кілька подібних прикладів із свого власного досвіду.

5.4.1.2. Причини лагів

Хоча приклади, наведені в параграфі 5.4.1.1, відображають природу такого явища як часовий лаг, вони не повністю пояснюють, чому лаги виникають. Загалом є три основні причини виникнення лагів.

1. *Психологічні причини.* Внаслідок інерції люди не змінюють своїх споживацьких звичок одразу після зниження цін або підвищення доходів. Можливо, це відбувається тому, що сам процес зміни може принести певну миттєву незадоволеність. Таким чином, ті, хто раптом стали мільйонерами внаслідок великого виграшу в лотереї, можуть і не змінити свого сти-

лю життя, оскільки вони одразу можуть і не знати, як реагувати на такий неочікуваний подарунок. Звичайно, з часом вони можуть навчитися жити відповідно своїм новим можливостям. Реакція на збільшення доходу може залежати також від того, чи є ця зміна постійною, чи тимчасовою. Якщо це тільки одноразове збільшення, а згодом прибуток повернеться до попереднього рівня, можна або заощадити весь “прибуток”, або все витратити.

2. Технологічні причини. Припустимо, що ціна капіталу залежить від зменшення праці, роблячи таким чином заміну капіталу працею економічно можливою. Звичайно, збільшення капіталу займає певний час. Більше того, якщо очікується тимчасовий спад цін, фірми можуть і не поспішати замінювати капітал працею, особливо якщо вони сподіваються, що після тимчасового спаду ціна капіталу може піднятися набагато вище за попередній рівень. Іноді неповне знання теж пояснює лаги.

Зараз ринок переповнений усіма видами кишенькових калькуляторів з різноманітними функціями і за різноманітними цінами. Більше того, з часу їхнього запровадження в кінці 60-х років ціни на більшість калькуляторів дуже знизились. Внаслідок цього майбутні споживачі калькуляторів можуть вагатися, чи купувати певну модель, аж до тих пір, доки вони не матимуть змоги вивчити можливості і ознайомитись із цінами всіх конкуруючих моделей. Більше того, вони можуть утриматися від купівлі в надії на подальше зниження цін або удосконалення калькуляторів.

3. Інституціональні причини також сприяють виникненню лагів. Наприклад, контрактові зобов'язання для деяких фірм можуть запобігати заміні одних джерел праці або сировини на інші. Можна також навести ще один приклад: ті, хто зберігає свої заощадження на довгострокових рахунках (1 або 2 роки), фактично зв'язані, хоча на грошовому ринку можуть бути такі умови, що в іншому місці можна отримати вищі прибутки. Подібним чином роботодавці часто пропонують своїм працівникам вибір між різними типами медичного страхування, але якщо людина вже обрали певний тип, вона не може його змінити протягом, як мінімум, одного року. Хоча це зроблено для адміністративних зручностей, працівник фактично зв'язаний протягом року.

З огляду на вказані причини, лаги посідають значне місце в економіці. Це чітко відображено в коротко- і довгостроковій методології економіки. Саме тому можна сказати, що короткострокові ціни, або еластичність, з доходом звичайно менші (за абсолютним значенням), ніж відповідна довгострокова еластичність, або що короткострокова гранична схильність до споживання в основному менша за довгострокову граничну схильність до споживання.

5.4.2. Оцінка параметрів дистрибутивно-лагових моделей

Якщо припустити, що дистрибутивно-лагові моделі відіграють важливу роль в економіці, як можна оцінити параметри такої моделі? Нехай ми маємо таку дистрибутивно-лагову модель з однією пояснювальною змінною:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \varepsilon_t, \quad (5.4.5)$$

де ми не визначаємо довжину лагу. Така модель має назву *нескінченна (лагова) модель*, тоді як модель типу (5.4.1) називається *скінченною дистрибутивно-лаговою моделлю*, оскільки в ній визначена довжина лагу k . Надалі будемо використовувати модель (5.4.5) як загальний випадок. Зразу ж постає питання, як оцінити невідомі параметри α і β_i в моделі (5.4.5)? Це можна зробити за двома способами: послідовного оцінювання та апріорного оцінювання, припускаючи, що β_i мають певну систематичну закономірність.

5.4.2.1. Послідовна оцінка дистрибутивно-лагових моделей

Оскільки припускається, що x_t нестохастичні (або принаймні не корелюються з помилковою e_t), то x_{t-1} , x_{t-2} , ... теж нестохастичні. Таким чином, до (5.4.5) можна застосувати метод найменших квадратів. Цей підхід застосували Ф.Альт і Дж.Тінберген. Вони припустили, що для того, щоб оцінити (5.4.5), потрібно діяти *послідовно*: тобто спочатку побудувати регресію y_t за x_t та оцінити невідомі параметри, потім — регресію y_t за x_t і за x_{t-1} , потім — за x_t , x_{t-1} і x_{t-2} і так далі. Ця послідовна процедура припиняється, коли параметри при лагових змінних x_t починають бути статистично незначними та (або) коефіцієнт хоча б однієї змінної змінює свій знак. Згідно з цим принципом Альт досліджував залежність споживання пального і мастил y від надходження нових замовлень x . Базуючи свої спостереження на щоквартальних даних за 1930—1939 рр., він отримав такі результати:

$$\hat{y}_t = 8.37 + 0.171x_t;$$

$$\hat{y}_t = 8.27 + 0.111x_t + 0.064x_{t-1};$$

$$\hat{y}_t = 8.27 + 0.109x_t + 0.071x_{t-1} - 0.055x_{t-2};$$

$$\hat{y}_t = 8.32 + 0.108x_t + 0.063x_{t-1} + 0.022x_{t-2} - 0.020x_{t-3}.$$

Альт обрав друге рівняння як “найкраще”, тому що в останніх двох знак x_{t-2} не був стабільним, а в четвертому знак x_{t-3} став від'ємним, що могло спричинити труднощі при економічній інтерпретації.

Хоча метод послідовних оцінок здається повним, він має багато недоліків.

1. Спочатку невідомо, яка максимальна тривалість лагу.
 2. При оцінці послідовних лагів залишається менше ступенів вільності, що робить економічні висновки дещо непевними.
 3. В економічних даних послідовні значення змінних звичайно мають високу кореляцію; таким чином, з'являється проблема мультиколінеарності.
- З огляду на наведені вади метод послідовних оцінок має дуже мало переваг для того, щоб його можна було рекомендувати для використання на практиці.

5.4.2.2. Підхід Койка до дистрибутивно-лагових моделей

Койк запропонував досить цікавий метод оцінки дистрибутивно-лагових моделей. Припустимо, ми починаємо з дистрибутивно-лагової моделі з невизначеним лагом (5.3.1). Припускаючи, що β_k мають той самий знак, Койк припустив також, що вони змінюються в геометричній прогресії:

$$\beta_k = \beta_0 \lambda^k \quad k = 0, 1, \dots, \quad (5.4.6)$$

де λ такі, що $0 < \lambda < 1$ — темп зменшення дистрибутивного лагу, а $(1 - \lambda)$ — швидкість пристосування. Співвідношення (5.4.6) показує, що кожний наступний коефіцієнт β менший, ніж попередній (оскільки $\lambda < 1$), тобто з кожним наступним кроком у минуле вплив лагу на y_t поступово зменшується, що є досить імовірним припущенням. Значення лагового коефіцієнта β_k залежить, крім загального β_0 , також і від λ . Чим ближче значення λ до 1, тим повільніший темп зменшення β_k , а чим ближче він до 0, тим швидше спадає β_k . У попередньому випадку віддалені в минулому значення x досить сильно впливали на y_t , тоді як у нашому випадку їхній вплив на y_t швидко зменшується. Це добре видно в табл. 5.7.

Таблиця 5.7

λ	β_0	β_1	β_2	β_3	β_4	β_5	...	β_{10}
0.75	β_0	$0.75\beta_0$	$0.56\beta_0$	$0.42\beta_0$	$0.32\beta_0$	$0.24\beta_0$...	$0.06\beta_0$
0.25	β_0	$0.25\beta_0$	$0.06\beta_0$	$0.02\beta_0$	$0.004\beta_0$	$0.001\beta_0$...	0

Слід зазначити, що метод Койка має такі переваги:

- припускаючи, що λ можуть бути від'ємними, Койк абстрагувався від зміни знака коефіцієнта при β_i ;
- завдяки тому, що $\lambda < 1$ віддалені за часом, значення β_i стали менш впливовими, ніж поточні;
- сума β_i , яка складає довгостроковий мультиплікатор, є скінченною, тобто

$$\sum_{k=0}^{\infty} \beta_k = \beta_0 \left(\frac{1}{1 - \lambda} \right). \quad (5.4.7)$$

Як результат (5.4.6), модель з кінцевим лагом (5.4.7) можна записати таким чином:

$$y_t = \alpha + \beta_0 x_t + \beta_0 \lambda x_{t-1} + \beta_0 \lambda^2 x_{t-2} + \dots + \varepsilon_t. \quad (5.4.8)$$

Як бачимо, модель (5.4.8) також незручна для оцінки, оскільки залишається дуже велика (фактично нескінченна) кількість оцінюваних параметрів, крім того, параметр λ входить до моделі в нелінійній формі: тобто метод лінійної (за параметрами) регресії не можна застосувати до цієї моделі. Але Койк пропонує модифікований метод, який полягає в тому, що в модель (5.4.8) вводиться затримка на один період. Виходячи з цього, модель записується таким чином:

$$y_{t-1} = \alpha + \beta_0 x_{t-1} + \beta_0 \lambda x_{t-2} + \beta_0 \lambda^2 x_{t-3} + \dots + \varepsilon_{t-1}. \quad (5.4.9)$$

Далі помножимо (5.4.9) на λ і отримаємо:

$$\lambda y_{t-1} = \lambda \alpha + \lambda \beta_0 x_{t-1} + \beta_0 \lambda^2 x_{t-2} + \beta_0 \lambda^3 x_{t-3} + \dots + \lambda \varepsilon_{t-1}. \quad (5.4.10)$$

Віднявши (5.4.10) від (5.4.8), маємо:

$$y_t - \lambda y_{t-1} = \alpha(1 - \lambda) + \beta_0 x_t + (\varepsilon_t - \lambda \varepsilon_{t-1}), \quad (5.4.11)$$

або

$$y_t = \alpha(1 - \lambda) + \beta_0 x_t + \lambda y_{t-1} + v_t, \quad (5.4.12)$$

де $v_t = (\varepsilon_t - \lambda \varepsilon_{t-1})$. Ця процедура відома як **перетворення Койка**. Порівнюючи (5.4.12) з (5.4.5), бачимо надзвичайне спрощення моделі. Якщо раніше нам треба було оцінювати параметр $\alpha\lambda$ та нескінченну кількість параметрів β_i , тепер достатньо оцінити лише три змінних: α , β_0 і λ , тобто немає причин очікувати мультиколінеарність. Фактично ми позбулись мультиколінеарності заміною x_{t-1} , x_{t-2} , ... на одну змінну, тобто y_{t-1} .

Зазначимо деякі особливості трансформації Койка.

1. Трансформація Койка переводить дистрибутивно-лагову модель в авторегресивну, оскільки серед незалежних змінних залишається y_{t-1} .
2. Поява y_{t-1} може спричинити ряд статистичних проблем: y_{t-1} , як і y_t , — стохастична; це означає, що в модель ми вводимо стохастичну змінну.

Згадаємо, що класичний метод найменших квадратів базується на загальному припущенні про те, що незалежні змінні або нестохастичні, або, якщо стохастичні, то розподілені незалежно від випадкової величини ε . Таким чином, ми повинні перевірити, чи задовольняє y_{t-1} цьому припущенню.

3. У початковій моделі (5.4.5) помилка дорівнювала ε_t , а в перетвореній $v_t = (\varepsilon_t - \lambda\varepsilon_{t-1})$. Тепер статистичні властивості v_t залежать від статистичних властивостей ε_t . Як буде показано пізніше, якщо ε_t серійно не корельовані, то v_t серійно корельовані. Таким чином, ми можемо зіткнутися з проблемою серійної кореляції щодо стохастичної пояснювальної змінної y_{t-1} .

4. Наявність лагового значення y порушує одне з припущень d -тесту Дарбіна — Уотсона. Отже, нам потрібно розробити альтернативу для тестування серійної кореляції при лаговому y . Цією альтернативою є h -тест Дарбіна, який ми розглянемо пізніше.

Середні і медіанні лаги

Як ми вже бачили в (5.4.4), часткові суми стандартизованих β_k показують пропорцію довгострокового або загального впливу на певному проміжку часу. Але на практиці для характеристики природи лагової структури в дистрибутивно-лаговій моделі використовують середні або медіанні лаги.

Медіанний лаг. Це час, потрібний для половинної (50% -ної) загальної зміни y після одиничної зміни x . Для моделі Койка медіанний лаг дорівнює

$$-\left(\frac{\log 2}{\log \lambda}\right) \quad (5.4.13)$$

Отже, якщо $\lambda=0.2$, то медіанний лаг дорівнюватиме 0.4306, але якщо $\lambda=0.8$, то медіанний лаг 3.1067. У першому випадку 50% зміни загального значення y відбудеться менш ніж через половину часового періоду, а в другому випадку для половинної зміни y потрібно більше ніж 3 часових періоди. Але це не повинно викликати здивування, оскільки відомо, що чим вище значення λ , тим менша швидкість впливу, і навпаки (див. приклад 5.4.9).

Середній лаг. Середній лаг обчислюється за такою формулою:

$$\frac{\sum_0^{\infty} k\beta_k}{\sum_0^{\infty} \beta_k} \quad (5.4.14)$$

Він фактично є зваженим середнім усіх лагів, включених до моделі, з відповідними коефіцієнтами β . Для моделі Койка середній лаг обчислюється за формулою

$$\left(\frac{\lambda}{1-\lambda}\right) \quad (5.4.15)$$

Таким чином, якщо $\lambda = 1/2$, то середній лаг дорівнює 1, тобто половинний вплив зміни залежної змінної y стане відчутним протягом першого проміжку часу (див. приклад 5.4.10).

Отже, медіанний і середній лаги використовуються для вимірювання швидкості, з якою y відповідає на зміни в x . У прикладі, наведеному в підрозділі 5.4.1, середній лаг дорівнює приблизно 11 кварталам, тобто потрібен певний час, щоб відчути зміну цін після зміни грошової пропозиції.

5.4.3. Перша модифікація моделі Койка: модель адаптивних очікувань

Модель Койка фактично є послідовною моделлю, оскільки її можна одержати чисто алгебраїчним шляхом; вона позбавлена будь-якого теоретичного обґрунтування. Але цей розрив можна подолати, якщо підійти до неї з іншої точки зору. Припустимо, що ми маємо таку модель:

$$y_t = \beta_0 + \beta_1 x_t^* + \varepsilon_t, \quad (5.4.16),$$

де y — попит на гроші;

x^* — точка рівноваги, оптимум, очікувана довгострокова або звичайна відсоткова ставка;

ε — випадкова величина.

Рівняння (5.4.16) показує, що попит на гроші є функцією від очікуваної відсоткової ставки.

Оскільки очікувану змінну x^* не можна спостерігати безпосередньо, запропонуємо таку гіпотезу формування очікування:

$$x_t^* - x_{t-1}^* = \gamma(x_t - x_{t-1}^*), \quad (5.4.17)$$

де $\gamma (0 < \gamma < 1)$ — коефіцієнт очікування.

Гіпотеза (5.4.17) відома як *гіпотеза адаптивного очікування*, або *помилкового навчання*.

Гіпотеза (5.4.17) передбачає, що “сили, задіяні в економіці, пристосовують свої очікування до попереднього досвіду, і зокрема вчать на своїх попередніх помилках”. Якщо точніше, (5.4.17) стверджує, що очікування кожного періоду коригуються на частку γ від розриву між поточним значенням змінної та її попереднім очікуваним значенням. Для нашої мо-

делі це означає, що очікування щодо відсоткової ставки кожного проміжку часу коригуються на частку γ від розходження між відсотковою ставкою у поточному періоді і її очікуваним значенням у попередньому періоді. Це твердження можна сформулювати інакше, якщо переписати (5.4.17) у вигляді

$$x_t^* = \gamma x_t + (1 - \gamma)x_{t-1}^*. \quad (5.4.18)$$

Це показує, що очікуване значення відсоткової ставки в момент часу t є зваженим середнім фактичного значення відсоткової ставки в момент t і його очікуваного значення в попередній період з коефіцієнтами γ і $1-\gamma$ відповідно. Якщо $\gamma = 1$ і $x_t^* = x_t$, це означає, що очікування справджуються негайно і повністю, тобто в поточний період часу. Якщо, з іншого боку, $\gamma = 0$ і $x_t^* = x_{t-1}^*$, значить очікування статичні, тобто “умови, що є сьогодні, зберігатимуться й у наступні періоди. Майбутні очікувані значення будуть ототожнюватись із поточними значеннями.”

Підставляючи (5.4.18) у (5.4.16), отримаємо:

$$\begin{aligned} y_t &= \beta_0 + \beta_1[\gamma x_t + (1 - \gamma)x_{t-1}^*] + \varepsilon_t = \\ &= \beta_0 + \beta_1\gamma x_t + \beta_1(1 - \gamma)x_{t-1}^* + \varepsilon_t. \end{aligned} \quad (5.4.19)$$

Тепер затримаємо (5.4.16) на один період, помноживши його на $1-\gamma$, і віднімемо добуток від (5.4.19). Після простих алгебраїчних перетворень ми отримаємо:

$$\begin{aligned} y_t &= \gamma \beta_0 + \gamma \beta_1 x_t + (1 - \gamma)y_{t-1} + \varepsilon_t - (1 - \gamma)\varepsilon_{t-1} \\ &= \gamma \beta_0 + \gamma \beta_1 x_t + (1 - \gamma)y_{t-1} + v_t, \end{aligned} \quad (5.4.20)$$

де $v_t = \varepsilon_t - (1 - \gamma)\varepsilon_{t-1}$.

Перед тим, як продовжити далі, покажемо різницю між (5.4.16) і (5.4.20). У першій моделі β_1 оцінює середню зміну y у відповідь на одиничну зміну x^* , рівноважне або довгострокове значення x . У (5.4.20), з іншого боку, β_1 оцінює середню зміну y у відповідь на одиничну зміну фактичного, або спостережуваного значення x . На практиці спочатку оцінюють (5.4.20). Як тільки отримано оцінку γ за допомогою коефіцієнта лагового значення y , можна легко обчислити β_1 , просто поділивши коефіцієнт $\gamma\beta_1$ при x_t на γ .

Модель адаптивних очікувань (5.4.20) і модель Койка (5.4.12) безумовно схожі між собою, хоча в них і різні інтерпретації коефіцієнтів. Зауважимо, що, як і модель Койка, модель адаптивних очікувань теж авторегресивна. Ми ще повернемося до оцінки моделі адаптивних очікувань і пізніше розглянемо деякі приклади. Тепер, коли ми побіжно розглянули

модель адаптивних очікувань, задамо питання: наскільки вона є реалістичною? Справді, вона більш приваблива, ніж чисто алгебраїчний підхід Койка, але чи має сенс гіпотеза адаптивних очікувань? На користь цієї гіпотези є такі аргументи.

Вона забезпечує досить прості засоби моделювання очікувань в економічній теорії. Гіпотеза, що люди навчаються з попереднього досвіду, є набагато розумнішою, ніж припущення про те, що вони всі позбавлені пам'яті, невід'ємної характеристики тези про статичні очікування. Більше того, твердження, що давніший досвід має менший вплив, ніж нещодавній, відповідає здоровому глузду, і його можна повністю підтвердити простими спостереженнями.

5.4.4. Друга модифікація моделі Койка: модель часткових пристосувань

Модель адаптивних очікувань — це один спосіб модифікації моделі Койка. Іншу модифікацію запропонував М. Нерлоу у своїй так званій *моделі часткових пристосувань*. Щоб проілюструвати цю модель, розглянемо модель гнучкого акселератора з економічної теорії, яка припускає, що є рівноважна, оптимально бажана або довгострокова кількість капіталу, необхідна для того, щоб виробляти певну кількість продукції при даному рівні технології, відсотковій ставці і т. ін. Для спрощення припустимо, що бажаний рівень капіталу y_t^* є лінійною функцією від випуску x :

$$y_t^* = \beta_0 + \beta_1 x_t + \varepsilon_t. \quad (5.4.21)$$

Оскільки бажаний рівень капіталу не можна спостерігати явно, Нерлоу запропонував гіпотезу, відому тепер під назвою *гіпотеза часткових пристосувань*:

$$y_t - y_{t-1} = \delta(y_t^* - y_{t-1}), \quad (5.4.22)$$

де δ ($0 < \delta < 1$) відоме під назвою *коефіцієнт пристосування*; $y_t - y_{t-1}$ — фактична зміна; а $y_t^* - y_{t-1}$ — бажана зміна.

Оскільки $y_t - y_{t-1}$, зміна запасу капіталу між двома періодами є нічим іншим, як інвестуванням, (5.4.22) можна переписати в іншому вигляді:

$$I_t = \delta(y_t^* - y_{t-1}), \quad (5.4.23)$$

де I_t — інвестування за період t .

Рівняння (5.4.22) показує, що фактична зміна в запасі капіталу (інвестиціях) у будь-який період часу t є певною часткою δ від бажаної зміни за цей період. Якщо $\delta=1$, це означає, що фактичний запас капіталу дорівнює бажаному, тобто фактичний запас пристосовується до бажаного в той самий період часу. Проте, якщо $\delta=0$, ніщо не змінюється, оскільки фактич-

ний запас у момент часу t такий самий, як і в попередній період часу. Реально δ лежить між цими екстремальними значеннями, оскільки пристосування до бажаного рівня капіталу найчастіше буває неповним через інертність, контрактівні зобов'язання і т. ін. Звідси впливає назва — **модель часткових пристосувань**. Зауважимо, що механізм пристосування можна переписати в іншому вигляді:

$$y_t = \delta y_t^* - (1 - \delta)y_{t-1}. \quad (5.4.24)$$

Це показує, що запас капіталу, який ми спостерігаємо в момент t , є зваженим середнім бажаного запасу капіталу в цей момент і запасу в попередній період, δ і $(1 - \delta)$ — вагові коефіцієнти. Тепер підставимо (5.4.21) до (5.4.24):

$$\begin{aligned} y_t &= \delta(\beta_0 + \beta_1 x_t + \varepsilon_t) + (1 - \delta)y_{t-1} = \\ &= \delta\beta_0 + \delta\beta_1 x_t + (1 - \delta)y_{t-1} + \delta\varepsilon_t. \end{aligned} \quad (5.4.25)$$

Ця модель має назву модель часткових пристосувань.

Оскільки (5.4.21) відображає довгостроковий або рівноважний попит на запас капіталу, (5.4.25) можна назвати *короткостроковою* функцією попиту на наявний запас капіталу, тому що в короткостроковому періоді наявний запас капіталу не обов'язково може дорівнювати своєму довгостроковому рівню. Як тільки ми оцінимо короткострокову функцію (5.4.25) і одержимо оцінку коефіцієнта пристосування δ (з коефіцієнта y_{t-1}), зможемо легко обчислити довгострокову функцію, поділивши $\delta\beta_0$ і $\delta\beta_1$ на δ і опустивши лагове значення y , яке потім дає (5.4.21).

Щоб проілюструвати наші викладки, припустимо, що $\delta = 0.5$. Це передбачає, що фірма планує закрити половину розриву між бажаним і фактичним запасом капіталу в кожному періоді. Таким чином, у перший період вона пересувається до y_2 з рівнем інвестицій $(y_2 - y_1)$, що у свою чергу дорівнює $0.5(y_1^* - y_1)$. У кожному наступному періоді вона закриває половину розриву між капітальним запасом на початку періоду і бажаним капітальним запасом y^* .

Модель часткових пристосувань нагадує як модель Койка, так і модель адаптивних очікувань у тому плані, що вона теж є авторегресивною. Але вона має значно простішу помилку: початкову помилку ε_t , помножену на константу δ . Слід зазначити, що, незважаючи на зовнішню схожість, модель адаптивних очікувань і модель часткових пристосувань концептуально дуже різні. Перша базується на невизначеності (майбутні ціни, ставки відсотків і т.ін.), тоді як остання залежить від технічних і інституціональних обмежень, інерції, вартості обміну та ін. Однак обидві ці моделі набагато більш обґрунтовані теоретично, ніж модель Койка.

5.4.5. Комбінація моделей адаптивних очікувань і часткових пристосувань

Розглянемо таку модель:

$$y_t^* = \beta_0 + \beta_1 x_t^* + \varepsilon_t, \quad (5.4.26)$$

де y_t^* — бажаний запас капіталу, а x_t^* — очікуваний рівень випуску.

Оскільки як y_t^* , так і x_t^* не можна спостерігати безпосередньо, можна використовувати механізм часткового пристосування для y_t^* і механізм адаптивних очікувань для x_t^* , щоб отримати таке рівняння (приклад 5.4.6):

$$\begin{aligned} y_t &= \beta_0 \delta + \beta_1 \delta x_t + [(1 - \gamma) + (1 - \delta)]y_{t-1} - \\ &- (1 - \delta)(1 - \gamma)y_{t-2} + [\delta\varepsilon_t - (1 - \gamma)\varepsilon_{t-1}] = \\ &= \alpha_0 + \alpha_1 x_t + \alpha_2 y_{t-1} + \alpha_3 y_{t-2} + v_t, \end{aligned} \quad (5.4.27)$$

де $v_t = [\delta\varepsilon_t - (1 - \gamma)\varepsilon_{t-1}]$.

Ця модель також є авторегресивною, а єдиною відмінністю від чистої моделі адаптивних очікувань є те, що серед незалежних змінних поряд з y_{t-1} є також y_{t-2} . Як і в моделях Койка та адаптивних очікувань, помилка в (5.4.27) є середнім ковзним процесом. Іншою відмінністю цієї моделі є те, що хоча модель є лінійною за параметрами α_i , вона нелінійна за початковими параметрами.

Відоме застосування (5.4.27) запропонував Фрідман у своїй *гіпотезі постійного доходу*, яка стверджує, що "*постійне*" або *довгострокове споживання є функцією від "постійного" або довгострокового доходу*.

При оцінці (5.4.27) виникають ті ж проблеми, що й при оцінці моделей Койка або адаптивних очікувань, — усі ці моделі є авторегресивними зі подібною структурою помилок. Крім того, (5.4.27) включає в себе також деякі проблеми нелінійних оцінок, які ми коротко розглянемо в прикладі 5.4.13, але не будемо детально розглядати в цьому підручнику.

5.4.6. Оцінювання параметрів авторегресивних моделей

Ми розглянули три моделі.

Модель Койка:

$$y_t = \alpha(1 - \lambda) + \beta_0 x_t + \lambda y_{t-1} + (\varepsilon_t - \lambda\varepsilon_{t-1}). \quad (5.4.12)$$

Модель адаптивних очікувань:

$$y_t = \gamma \beta_0 + \gamma \beta_1 x_t + (1 - \gamma)y_{t-1} + [\varepsilon_t - (1 - \gamma)\varepsilon_{t-1}]. \quad (5.4.19)$$

Модель часткових пристосувань:

$$y_t = \delta\beta_0 + \delta\beta_1 x_t + (1 - \delta)y_{t-1} + \delta\varepsilon_t. \quad (5.4.25)$$

Усі ці моделі можна подати в такій загальній формі:

$$y_t = \alpha_0 + \alpha_1 x_t + \alpha_2 y_{t-1} + v_t, \quad (5.4.28)$$

тобто всі вони за природою авторегресивні.

Постає проблема оцінювання невідомих параметрів цих моделей, тому що до них не можна прямо застосувати метод найменших квадратів з двох причин: серед пояснювальних змінних є стохастичні, а також існує можливість серійної кореляції.

Як уже зазначалося, для того, щоб застосувати класичний метод найменших квадратів, треба показати, що стохастична пояснювальна змінна y_{t-1} розподілена незалежно від випадкової величини v_t . Щоб визначити, чи це справджується, необхідно знати властивості v_t . Якщо припустити, що початкова випадкова величина ε_t задовольняє всім класичним припущенням, таким як $E(\varepsilon_t) = 0$, $\text{var}(\varepsilon_t) = \text{const}$ (припущення гомоскедастичності) і $\text{cov}(\varepsilon_t, \varepsilon_{t+s}) = 0$ для $s \neq 0$ (припущення відсутності автокореляції), v_t може й не успадковувати всі ці властивості. Розглянемо, наприклад, помилку в моделі Койка, $v_t = (\varepsilon_t - \lambda\varepsilon_{t-1})$. Якщо всі припущення щодо ε_t правильні, то легко показати, що v_t серійно корельоване, тому що

$$E(v_t v_{t-1}) = -\lambda \sigma^2 \quad (5.4.29)$$

і це не дорівнює нулеві (інакше λ буде нульовим). А оскільки y_{t-1} у моделі Койка належить до пояснювальних змінних, воно, очевидно, повинно корелюватися з v_t (через наявність у ньому ε_{t-1}). Фактично, це можна показати таким чином:

$$\text{cov}[y_{t-1}, (\varepsilon_t - \lambda \varepsilon_{t-1})] = -\lambda \sigma^2, \quad (5.4.30)$$

що відповідає (5.8.2). Читач може переконатись, що аналогічні міркування справджуються і для моделі адаптивних очікувань.

Для чого потрібно знати, що в моделі Койка, як і в моделі адаптивних очікувань, стохастична незалежна змінна y_{t-1} корелюється з випадковою величиною v_t ? Як зазначалося вище, **коли пояснювальна змінна корелюється з випадковою величиною, оцінки методу найменших квадратів не лише зміщені, але й не є напевно консистентними**; таким чином, навіть якщо розмір вибірки нескінченно зростатиме, оцінки не наблизяться до реальних значень генеральної сукупності. Отже, оцінка моделі Койка і моделі адаптивних очікувань за допомогою звичайного методу найменших квадратів може призвести до істотно помилкових результатів.

Модель часткових пристосувань відрізняється від попередніх. У цій моделі $v_t = \delta\varepsilon_t$, де $0 < \delta < 1$. Таким чином, якщо ε_t задовольняє припущен-

ням класичної лінійної регресійної моделі, то $\delta\varepsilon_t$ теж буде задовольняти цим припущенням. Отже, оцінювання моделі часткових пристосувань методом найменших квадратів дає консистентні оцінки, хоча вони будуть зміщеними (в кінцевих або маленьких вибірках). Інтуїтивно причина консистентності оцінок полягає ось у чому: хоча y_{t-1} залежить від ε_{t-1} і всіх попередніх значень, вона не пов'язана з поточним значенням ε_t . Таким чином, доки ε_t буде серійно незалежною, y_{t-1} теж буде незалежною або, принаймні, некорельованою з ε_t , задовольняючи, таким чином, важливе припущення методу найменших квадратів — некорельованість між пояснювальною змінною (змінними) і випадковою величиною.

Хоча оцінювання моделі часткових пристосувань методом найменших квадратів забезпечує консистентні оцінки через те, що помилка в такій моделі має просту структуру, не можна припускати, що ця модель більш застосовувана, ніж модель Койка або модель адаптивних очікувань. Модель потрібно вибирати на основі серйозних теоретичних міркувань, а не просто тому, що її легше статистично оцінити. Кожну модель треба застосовувати з огляду на її власні переваги, зважаючи на елементи стохастичності, які можуть у ній трапитись. Якщо до таких моделей, як модель Койка або модель адаптивних очікувань, метод найменших квадратів не можна застосовувати прямо, потрібно знайти інші методи, щоб оцінити їх. Є кілька альтернативних методів оцінювання, хоча деякі з них можуть бути складними для обчислень. Розглянемо один з таких методів.

5.4.7. Метод допоміжних змінних

Причина неможливості застосування методу найменших квадратів до моделі Койка або моделі адаптивних очікувань полягає в тому, що пояснювальна змінна y_{t-1} корелюється з помилкою v_t . Якщо якимось чином усунути цю кореляцію, то можна застосувати метод найменших квадратів, щоб отримати консистентні оцінки, як було зазначено вище (*зауваження*: оцінки можуть бути зміщеними в малих вибірках). Як цього досягти?

Припустимо, що ми знайшли “замінник” для y_{t-1} , який сильно корелює з y_{t-1} , але не корелює з v_t , де v_t — це помилка, яка з'являється в моделі Койка або в моделі адаптивних очікувань. Такий заміник має назву **допоміжної змінної**. Розглянемо x_{t-1} як допоміжну змінну для y_{t-1} , а потім припустимо, що параметри регресії (5.4.28) можна отримати, розв'язавши таку систему нормальних рівнянь:

$$\begin{aligned} \sum y_t &= N a_0 + a_1 \sum x_t + a_2 \sum y_{t-1}; \\ \sum y_t x_t &= a_0 \sum x_t + a_1 \sum x_t^2 + a_2 \sum y_{t-1} x_t; \\ \sum y_t x_{t-1} &= a_0 \sum x_{t-1} + a_1 \sum x_t x_{t-1} + a_2 \sum y_{t-1} x_{t-1}. \end{aligned} \quad (5.4.31)$$

При цьому a_i є оцінками дійсних параметрів α_i . Зауважимо, що якби ми застосували метод найменших квадратів прямо до (5.4.28), система рівнянь звичайного методу найменших квадратів мала б такий вигляд:

$$\begin{aligned} \sum y_t &= N a_0 + a_1 \sum x_t + a_2 \sum y_{t-1}; \\ \sum y_t x_t &= a_0 \sum x_t + a_1 \sum x_t^2 + a_2 \sum y_{t-1} x_t; \\ \sum y_t y_{t-1} &= a_0 \sum y_{t-1} + a_1 \sum x_t y_{t-1} + a_2 \sum y_{t-1}^2. \end{aligned} \quad (5.4.32)$$

Різниця між цими двома системами рівнянь повинна бути очевидною. Теоретично було показано, що a_i , обчислені з (5.4.31), консистентні, в той час як обчислення за (5.4.32) може й не дати консистентних оцінок. Це трапляється тому, що y_{t-1} і $v_t [= \varepsilon_t - \lambda \varepsilon_{t-1} \text{ або } \varepsilon_t - (1 - \gamma)\varepsilon_{t-1}]$ можуть корелюватися, а x_t і x_{t-1} не корелюються з v_t . (Спробуйте це показати самостійно.)

Хоча метод допоміжних змінних легко застосовувати, як тільки знайдено придатний замітник, але може виникнути проблема мультиколінеарності, тому що x_t і x_{t-1} , які входять до системи нормальних рівнянь (5.4.31), можуть бути висококорельованими (часові ряди звичайно мають високу кореляцію між послідовними значеннями). Таким чином, хоча цей метод дозволяє отримати консистентні оцінки, вони можуть виявитися неефективними.

Перед тим як просуватися далі, маємо відповісти на запитання: яким чином можна знайти “добрий” замітник для y_{t-1} , який би високо корелювався з y_{t-1} , але не корелювався з v_t ? В літературі можна знайти декілька відповідей, їх ми розглянемо пізніше. Треба зазначити, що не завжди легко знайти добрі замітники, тому метод допоміжних змінних має малу область застосування на практиці і потрібно звертатися до методів оцінки з максимальною вірогідністю, що виходить за рамки цієї книги.

5.4.8. Виявлення автокореляції в авторегресивних моделях:

h -тест Дарбіна

Як ми вже бачили, в авторегресивних моделях є імовірність того, що випадкові величини v_t можуть бути корельовані. Це спричинює певні проблеми при оцінюванні: у моделі часткових пристосувань випадкова величина v_t не має серійної кореляції (першого порядку), якщо ε_t в початковій моделі були некорельовані, тоді як у моделях Койка і адаптивних очікувань v_t були корельовані, навіть якщо ε_t були незалежні. Таким чином, постає запитання: як виявити кореляцію в помилці авторегресивної моделі?

Вище зазначалось: d -тест Дарбіна — Уотсона не можна використовувати для визначення серійної кореляції (першого порядку) в авторегресивних моделях, тому що обчислене значення d в таких моделях звичайно

наближається до 2, що очікується лише у справді випадкових послідовностях. Іншими словами, якщо ми ретельно обчислимо значення d для цих моделей, то з’явиться закладена помилка, яка не дозволить виявити серійну кореляцію (першого порядку). Нещодавно Дарбін запропонував власний тест серійної кореляції (першого порядку) в авторегресивній моделі для великих вибірок. Цей тест називається h -тестом і має такий вигляд:

$$h = \hat{\rho} \sqrt{\frac{N}{1 - N[\text{var}(a_2)]}}, \quad (5.4.33)$$

де N — розмір вибірки; $\text{var}(a_2)$ — дисперсія оцінки параметра a_2 при Y_{t-1} ; $\hat{\rho}$ — оцінка коефіцієнта серійної кореляції (першого порядку) ρ . Для великих вибірок Дарбін показав, що коли $\rho = 0$, h розподілено за стандартним нормальним законом (тобто нормальний розподіл, де математичне сподівання дорівнює нулеві, а дисперсія — одиниці). Звідси випливає, що статистичну значимість спостережуваного h можна легко визначити з таблиці нормального стандартного розподілу.

На практиці, звичайно, не потрібно обчислювати $\hat{\rho}$, тому що його можна приблизно обчислювати через оцінене значення d таким чином:

$$\hat{\rho} = 1 - \frac{1}{2} d, \quad (5.4.34)$$

де d — значення звичайного тесту Дарбіна — Уотсона.

Таким чином, (5.4.33) можна переписати у вигляді:

$$h = \left(1 - \frac{1}{2} d\right) \sqrt{\frac{N}{1 - N(\text{var}(a_2))}}. \quad (5.4.35)$$

Застосування h -тесту можна розбити на такі кроки.

1. Оцінити (5.4.28) за методом найменших квадратів.
2. Обчислити $\text{var}(a_2)$.
3. Обчислити $\hat{\rho}$, як зазначено в (5.4.34).
4. Обчислити h з (5.4.33) або (5.4.35).
5. Припустивши, що N — велике, можна впевнитись, що

$$h \sim N(0,1), \quad (5.4.36)$$

тобто h розподілене за нормальним законом з математичним сподіванням, що дорівнює нулеві, та дисперсією одиниця. Із закону нормального розподілу відомо, що

$$\text{Pr} [-1.96 \leq h \leq 1.96] = 0.95, \quad (5.4.37)$$

тобто h лежить між -1.96 і $+1.96$ з імовірністю 95% .

Отже, розв'язок виглядає таким чином:

а) якщо $h > 1.96$, то відкинути нуль-гіпотезу про те, що немає позитивної автокореляції першого порядку;

б) якщо $h < -1.96$, то також відкинути нуль-гіпотезу про те, що немає негативної автокореляції першого порядку;

в) якщо h знаходиться між -1.96 та 1.96 , то не відкидати нуль-гіпотезу про те, що немає автокореляції першого порядку (позитивної чи негативної).

Щоб проілюструвати це, припустимо, що у вибірці, яка нараховує 100 спостережень, встановлено, що $d=1.9$ і $\text{var}(a_2) = 0.005$. Тоді

$$h = \left[1 - \frac{1}{2}(1.9)\right] \sqrt{\frac{100}{1 - 100(0.005)}} = 0.7071.$$

Оскільки обчислене значення h лежить у межах (5.4.37), ми не можемо з 5% -ним рівнем довіри відкинути гіпотезу про те, що немає позитивної автокореляції.

Особливості h -тесту

1. Немає значення, скільки змінних x або скільки попередніх значень y включено до регресійної моделі. Щоб обчислити значення h , потрібно розглянути лише дисперсію коефіцієнта попереднього значення y_{t-1} .

2. Тест не можна застосовувати, якщо $[N \times \text{var}(a_2)]$ перевищує 1 . (Спробуйте показати це самостійно.) Однак на практиці цього, звичайно, не трапляється.

3. Оскільки тест розрахований на великі вибірки, застосування його до малих вибірок не є строго обґрунтованим. Властивості тесту для малих вибірок ще не повністю встановлені.

5.4.9. Ілюстративний приклад: попит на гроші в Україні

Розглянемо залежність грошового запасу України від національного доходу, цін і довгострокової відсоткової ставки. Найбільш поширений вигляд моделей цього типу такий:

$$M_t^* = \beta_0 R_t^{\beta_1} Y_t^{\beta_2} e^{\varepsilon_t}, \quad (5.4.38)$$

де M_t^* — бажаний або довгостроковий попит на гроші (реальний грошовий баланс);

R_t — довгострокова відсоткова ставка, %;

y_t — сукупний дійсний національний доход.

Для статистичної оцінки (5.4.38) можна зручно подати в логарифмічній формі:

$$\ln M_t^* = \ln \beta_0 + \beta_1 \ln R_t + \beta_2 \ln y_t + \varepsilon_t. \quad (5.4.39)$$

Оскільки бажану зміну попиту не можна спостерігати безпосередньо, припустимо гіпотезу часткових пристосувань, тобто

$$\frac{M_t}{M_{t-1}} = \left(\frac{M_t^*}{M_{t-1}^*} \right)^\delta, \quad 0 \leq \delta \leq 1. \quad (5.4.40)$$

Рівняння (5.4.40) показує, що постійний відсоток різниці між фактичним і бажаним реальним грошовим балансом ліквідується протягом одного періоду (року). (Спробуйте це показати.) В логарифмічній формі рівняння (5.4.40) можна подати таким чином:

$$\ln M_t - \ln M_{t-1} = \delta(\ln M_t^* - \ln M_{t-1}^*). \quad (5.4.41)$$

Підставляючи $\ln M_t^*$ з (5.4.39) у рівняння (5.4.41), після перетворень отримаємо:

$$\ln M_t = \delta \ln \beta_0 + \beta_1 \delta \ln R_t + \beta_2 \delta \ln y_t + (1 - \delta) \ln M_{t-1} + \delta \varepsilon_t, \quad (5.4.42)$$

яке можна назвати *короткостроковою функцією попиту* на гроші. Припустимо, що ε_t (а отже, і $\delta \varepsilon_t$) задовольняє всім класичним припущенням. Нехай ми отримали такі результати (дані умовні):

$$\begin{aligned} \hat{\ln} M_t &= 1.6027 - 0.1024 \ln R_t + 0.6869 \ln y_t + 0.5284 \ln M_{t-1} \\ &\quad (1.2404) \quad (0.3678) \quad (0.3427) \quad (0.2007) \\ t &= (1.3066) \quad (-0.2784) \quad (2.0108) \quad (2.6328) \\ R^2 &= 0.9227 \quad d = 1.8624. \end{aligned} \quad (5.4.43)$$

Оцінена функція короткострокового попиту показує, що короткострокова еластичність відсоткової ставки статистично незначима, але короткострокова еластичність доходу статистично значима з 5% -ним рівнем довіри (односторонній тест). Коефіцієнт пристосування $\delta = 1 - 0.5284 = 0.4716$, припускаючи, що 47% -на розбіжність між бажаним і фактичним реальним грошовим балансом покривається протягом року. Щоб повернутися до довгострокової функції попиту (5.4.39), треба скрізь поділити короткострокову функцію попиту на d і опустити $\ln M_{t-1}$. Отже, отримано такий результат:

$$\ln M_t^* = 2.2520 - 0.2169 \ln R_t + 1.4565 \ln y_t. \quad (5.4.44)$$

Як бачимо, довгострокова еластичність попиту за доходом на гроші 1.4565 значно більша, ніж короткострокова еластичність 0.6869.

Зверніть увагу, що оцінений d -критерій Дарбіна — Уотсона дорівнює 1.8624, тобто близький до 2. Це підтверджує наше попереднє зауваження про те, що в авторегресивних моделях обчислене d звичайно наближається до 2. Таким чином, ми не можемо довіряти обчисленому значенню d , щоб з'ясувати, чи є серійна кореляція. Хоча розмір нашої вибірки досить невеликий, що, строго кажучи, не дозволяє застосовувати h -тест, тим не менше, ми його наводимо, щоб проілюструвати механізм обчислювання. Використовуючи оцінене значення d і формулу (5.4.35), отримуємо:

$$h = \left[1 - \frac{1}{2}(1.8624)\right] \sqrt{\frac{16}{1 - 16(0.0403)}} = 0.4617,$$

де дисперсія затриманої залежної змінної отримана із стандартної помилки цієї змінної, тобто $(0.2007)^2$.

Хоча оцінене h досить мале, що обумовлює прийняття гіпотези про відсутність серійної кореляції (першого порядку), але до цього висновку треба ставитись із застереженням, зважаючи на невеликий розмір вибірки.

5.4.10. Ілюстративний приклад: короткострокова і довгострокова функції сукупного споживання у США

Розглянемо широко відомий приклад М. Ловеля — моделювання короткострокової та довгострокової функції сукупного споживання у США. Припустимо, споживання C лінійно залежить від постійного доходу X^* :

$$C_t = \beta_1 + \beta_2 x_t^* + \varepsilon_t. \quad (5.4.45)$$

Оскільки x^* не можна спостерігати безпосередньо, потрібно визначити механізм, який би генерував постійний доход. Припустимо, що ми приймаємо гіпотезу адаптивних очікувань, яку ми визначили в (5.4.17). Використовуючи (5.4.17), після спрощень отримуємо таке рівняння (порівняйте з 5.4.20):

$$C_t = \alpha_1 + \alpha_2 x_t + \alpha_3 C_{t-1} + v_t, \quad (5.4.46)$$

де $\alpha_1 = \gamma\beta_1$; $\alpha_2 = \gamma\beta_2$; $\alpha_3 = (1 - \gamma)$ і $v_t = \varepsilon_t - (1 - \gamma)\varepsilon_{t-1}$.

Як відомо, β_2 відображає середню зміну споживання на підвищення постійного доходу, наприклад на 1 дол., в той час як α_2 відображає середню зміну споживання на підвищення поточного доходу на 1 дол.

Базуючись на щоквартальних даних для США за 1946 — 1972 рр., М. Ловел отримав такі результати (сукупна пропозиція і сукупний доступний доход були взяті з урахуванням індексу цін, щоб перетворити їх на реальні показники):

$$\begin{aligned} \hat{C}_t &= 2.361 + 0.2959x_t + 0.6755C_{t-1} & (5.4.47) \\ &(1.229) (0.0582) \quad (0.0666) \\ \bar{R}^2 &= 0.999 \\ d &= 1.77. \end{aligned}$$

Ця регресійна модель показує, що гранична схильність до споживання (МРС) дорівнює 0.2959, або близько 0.3. Це означає, що підвищення поточного або спостережуваного доходу на 1 дол. підвищить споживання у середньому на 30 центів. Але якщо це підвищення доходу збережеться, то фактично МРС від рівня поточного доходу буде $\beta_2 = \gamma\beta_2/\gamma = 0.2959/0.3245$, або близько 91 цента. Іншими словами, коли споживачі пристосуються до зміни доходу на 1 дол., вони підвищать споживання на 91 цент.

Тепер припустимо, що наша функція споживання має такий вигляд:

$$C_t^* = \beta_1 + \beta_2 x_t + \varepsilon_t. \quad (5.4.48)$$

При цьому постійне або довгострокове споживання C_t є лінійною функцією від поточного або спостережуваного доходу. Оскільки C_t^* не можна спостерігати безпосередньо, застосуємо модель часткових пристосувань (5.4.22). Використовуючи цю модель, після алгебраїчних перетворень отримуємо:

$$\begin{aligned} C_t &= \delta\beta_1 + \delta\beta_2 x_t + (1 - \delta)C_{t-1} + \delta\varepsilon_t \\ &= \alpha_1 + \alpha_2 x_t + \alpha_3 C_{t-1} + v_t. \end{aligned} \quad (5.4.49)$$

За своїм зовнішнім виглядом ця модель нічим не відрізняється від моделі адаптивних очікувань (5.4.46). Таким чином, результати регресії, отримані в (5.4.47), можна еквівалентно застосувати і тут. Але в інтерпретації цих двох моделей є головна відмінність, якщо залишити осторонь проблеми оцінювання в авторегресивній і, можливо, серійно корельованій моделі (5.4.46). Модель (5.4.48) — це довгострокова, або рівноважна функція споживання, в той час як (5.4.49) — це короткострокова функція споживання, β_2 вимірює довгострокове МРС, а $\alpha_2 (= \delta\beta_2)$ — короткострокове МРС. Першу модель можна отримати з останньої, поділивши її на коефіцієнт пристосування δ .

Повертаючись до (5.4.47), можемо інтерпретувати 0.2959 як короткострокове МРС. Оскільки $\delta = 0.3245$, довгострокове МРС дорівнює 0.91 дол., або 91 цент. Зауважимо, що коефіцієнт пристосування дорівнює 0.33, тобто

в кожний період часу споживачі будуть пристосовувати своє споживання лише на одну третину від бажаного довгострокового рівня.

Цей приклад ілюструє один дуже важливий момент: за зовнішнім виглядом модель адаптивних очікувань і модель часткових пристосувань настільки схожі, що, наприклад, з (5.4.47) не можна одразу визначити, яка саме модель тут подана. Ось чому так важливо точно визначити теоретичне підґрунтя моделі, обраної для емпіричного аналізу, а потім діяти належним чином. Якщо це звичка або інерція, яка характеризує поведінку споживача, то придатною буде модель часткових пристосувань. З іншого боку, якщо поведінка споживача далекоглядна, тобто базується на очікуваному майбутньому доході, то краще обрати модель адаптивних очікувань. В останньому випадку потрібно приділити багато уваги саме оціночній проблемі, щоб отримати спроможні оцінки. В першому випадку спроможні оцінки може забезпечити метод найменших квадратів, якщо будуть виконані всі його припущення.

5.4.11. Підхід Ш. Альмона до дистрибутивно-лагових моделей: поліноміальний лаг Альмона

Хоча дистрибутивно-лагова модель Койка широко використовується на практиці, вона базується на припущенні, що коефіцієнти β спадають у геометричній прогресії в міру зростання довжини лагу. Це припущення може бути занадто строгим у деяких ситуаціях, і схема дистрибутивно-лагових моделей Койка не спрацює. У складніших випадках параметри β_i можна виразити як функцію від i , тривалості лагу (часу) і підібрати відповідні криві, які відобразатимуть цю функціональну залежність. Саме цей підхід і запропонований Ш. Альмоном. Щоб проілюструвати його метод, повернемося до скінченної дистрибутивно-лагової моделі, яку ми розглядали раніше:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \beta_k x_{t-k} + \varepsilon_t. \quad (5.4.50)$$

Її можна записати в більш компактному вигляді:

$$y_t = \alpha + \sum_{i=0}^k \beta_i x_{t-i} + \varepsilon_t. \quad (5.4.51)$$

Відповідно до *теорему Веєрштрасса* Альмон припустив, що β_i можна апроксимувати поліномом відповідного ступеня від i , тривалості лагу. Наприклад:

$$\beta_i = a_0 + a_1 i + a_2 i^2, \quad (5.4.52)$$

яка є квадратичним поліномом від i другого ступеня або поліномом третього ступеня:

$$\beta_i = a_0 + a_1 i + a_2 i^2 + a_3 i^3. \quad (5.4.53)$$

У загальному випадку можна записати:

$$\beta_i = a_0 + a_1 i + a_2 i^2 + \dots + a_m i^m, \quad (5.4.54)$$

що є поліномом m -го ступеня від i . Припускається, що m (ступінь полінома) менше, ніж k (максимальна довжина лагу).

Щоб пояснити, *як працює схема Альмона*, припустимо, що β_i змінюються таким чином, що можна обрати поліноміальну апроксимацію другого ступеня (вигляд залежності краще за все обирати за зовнішнім виглядом графіка залежності величини параметра від лагу). Підставляючи (5.4.52) до (5.4.51), отримаємо:

$$\begin{aligned} y_t &= \alpha + \sum_{i=0}^k (a_0 + a_1 i + a_2 i^2) x_{t-i} + \varepsilon_t = \\ &= \alpha + a_0 \sum_{i=0}^k x_{t-i} + a_1 \sum_{i=0}^k i x_{t-i} + a_2 \sum_{i=0}^k i^2 x_{t-i} + \varepsilon_t. \end{aligned} \quad (5.4.55)$$

Визначаючи

$$\begin{aligned} Z_{0t} &= \sum_{i=0}^k x_{t-i}; \\ Z_{1t} &= \sum_{i=0}^k i x_{t-i}; \\ Z_{2t} &= \sum_{i=0}^k i^2 x_{t-i}, \end{aligned} \quad (5.4.56)$$

можна переписати (5.4.55) як

$$y_t = \alpha + a_0 Z_{0t} + a_1 Z_{1t} + a_2 Z_{2t} + \varepsilon_t. \quad (5.4.57)$$

У моделі Альмона y залежить від штучно створених змінних Z , а не від початкових змінних x . Зауважимо, що (5.4.57) можна оцінити за звичайним методом найменших квадратів. Оцінки α і a_i , отримані таким чином, матимуть усі бажані статистичні властивості, якщо випадкова величина ε_t задовольнятиме припущенням класичної моделі лінійної регресії. З цього боку *модель Альмона має чітку перевагу перед моделлю Койка*, оскільки остання, як ми бачили, має деякі серйозні проблеми через присутність у моделі пояснювальної змінної y_{t-1} , яка, імовірно, корелюється з випадковою величиною ε_t .

Оцінивши коефіцієнти a_i з (5.4.57), початкові β_i можна отримати з (5.4.52) [у загальному випадку з (5.4.54)] таким чином:

$$\begin{aligned}
 \hat{\beta}_0 &= \hat{\alpha}_0 \\
 \hat{\beta}_1 &= \hat{\alpha}_0 + \hat{\alpha}_1 + \hat{\alpha}_2 \\
 \hat{\beta}_2 &= \hat{\alpha}_0 + 2\hat{\alpha}_1 + 4\hat{\alpha}_2 \\
 \hat{\beta}_3 &= \hat{\alpha}_0 + 3\hat{\alpha}_1 + 9\hat{\alpha}_2 \\
 &\dots\dots\dots \\
 \hat{\beta}_k &= \hat{\alpha}_0 + k\hat{\alpha}_1 + k^2\hat{\alpha}_2.
 \end{aligned}
 \tag{5.4.58}$$

Перед застосуванням методу Альмона потрібно вирішити такі практичні проблеми.

1. Максимальна тривалість лагу k має бути визначена заздалегідь. Це найголовніший недолік методу Альмона. Дослідник повинен визначити найпридатнішу тривалість лагу. На практиці, звичайно, припускають, що k достатньо мала. Таким чином, у регресії, яка включає в себе щоквартальні дані за 10 років, ми, найшвидше, оберемо максимальний лаг 8 — 10 кварталів. Але якщо будуть лише щорічні дані за ті ж 10 років, нам може і не знадобитися лаг більший, ніж за 2 — 3 роки. У будь-якому разі дослідник сам має визначити максимальне значення k .

2. Визначивши k , треба також визначити ступінь полінома m . В загальному випадку ступінь полінома має бути принаймні на одиницю більший за кількість точок екстремума кривої, що показує залежність β_i від i . Тобто заздалегідь потрібно знати кількість точок екстремуму, таким чином, вибір m є великою мірою суб'єктивним. Але в деяких випадках теорія може допомогти знайти потрібний вигляд кривої. На практиці припускають, що за допомогою полінома низького ступеня (скажімо, m дорівнює 2 або 3) можна отримати добрі результати. Якщо ми обрали певне значення m і хочемо з'ясувати, чи не буде кращим поліном вищого ступеня, потрібно діяти таким чином.

Припустимо, нам потрібно зробити вибір між поліномом другого та третього ступеня. Рівняння для полінома другого ступеня наведено в (5.4.57). Для полінома третього ступеня відповідне рівняння має такий вигляд:

$$y_t = \alpha + a_0 Z_{0t} + a_1 Z_{1t} + a_2 Z_{2t} + a_3 Z_{3t} + \varepsilon_t, \tag{5.4.59}$$

$$\text{де } Z_{3t} = \sum_{i=0}^k i^3 x_{t-1}.$$

Після знаходження параметрів регресії (5.4.59), якщо ми знайдемо, що a_2 статистично значиме, а a_3 — ні, можемо припустити, що достатньою буде апроксимація поліномом другого ступеня.

Але слід бути обережними щодо проблеми мультиколінеарності, яка

може виникнути через те, що значення Z_i було отримано через значення x_t , як показано в (5.4.56) [див. також (5.4.60)]. У випадку мультиколінеарності \hat{a}_3 може стати статистично незначимою не тому, що дійсне значення a_3 дорівнює нулеві, а просто тому, що вибірка не дозволяє оцінити окремий вплив Z_3 на y . Отже, у нашому прикладі перед тим, як дійти висновку, що не можна обирати поліном третього ступеня, треба впевнитись, що немає мультиколінеарності.

Таким чином, з чисто емпіричної точки зору вибір ступеня полінома базується на статистичній значимості коефіцієнтів a_i в моделях, подібних до (5.4.59), але необхідно враховувати проблему мультиколінеарності.

3. Визначивши m і k , можна легко знайти значення Z_i . Наприклад, $m = 2$ і $k = 5$. Тоді

$$\begin{aligned}
 Z_{0t} &= \sum_{i=0}^k x_{t-i} = (x_t + x_{t-1} + x_{t-2} + x_{t-3} + x_{t-4} + x_{t-5}); \\
 Z_{1t} &= \sum_{i=0}^k i x_{t-i} = (x_{t-1} + 2x_{t-2} + 3x_{t-3} + 4x_{t-4} + 5x_{t-5}); \\
 Z_{2t} &= \sum_{i=0}^k i^2 x_{t-i} = (x_{t-1} + 4x_{t-2} + 9x_{t-3} + 16x_{t-4} + 25x_{t-5}).
 \end{aligned}
 \tag{5.4.60}$$

Зауважимо, що Z_i — це лінійні комбінації початкових x_t , а також те, що Z_i можуть бути мультиколінеарними.

Перед тим, як перейти до числового прикладу, зазначимо переваги методу Альмона. По-перше, він забезпечує гнучкий спосіб залучення до моделі цілого ряду лагових структур, у той час як модель Койка досить суворо вимагає від коефіцієнтів β_i , щоб вони спадали в геометричній прогресії. По-друге, на відміну від методу Койка, в моделі Альмона не потрібно турбуватися про те, що серед пояснювальних змінних є залежні, а отже, ми позбавляємось проблем, які можуть виникнути у зв'язку з цим. Нарешті, якщо обрано поліном досить низького ступеня, кількість оцінюваних коефіцієнтів (a_i) буде набагато менша, ніж початкова кількість їх (β_i).

Тепер повернемося до проблем, пов'язаних із застосуванням методу Альмона. По-перше, ступінь полінома, як і максимальне значення лагу, обирається дуже суб'єктивно. По-друге, з причин, зазначених вище, змінні Z можуть бути мультиколінеарними. Отже, в моделях, подібних до (5.4.60), оцінені коефіцієнти a_i можуть мати великі стандартні помилки (які стосуються значення цих коефіцієнтів), і, таким чином, один або більше коефіцієнтів можуть виявитися статистично незначимими згідно з t -тестом Ст'юдента. Але це не обов'язково означає, що один або більше початкових коефіцієнтів β також виявляться статистично незначимими. (Доведення цього твердження не розглядається). В результаті може виявитись, що проблема мультиколінеарності не є такою серйозною, як здається.

Для ілюстрації методу Альмона розглянемо ілюстративний приклад

залежності кількості товарів y від обсягу продажу x у виробничому секторі України за 1977 — 1996 рр. Умовні дані наведено в табл. 5.8. При цьому припустимо, що кількість товарів залежить від продажу в поточному і трьох попередніх роках, тобто можна записати таку модель:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \varepsilon_t. \quad (5.4.61)$$

Далі припустимо, що β_i можна апроксимувати поліномом другого ступеня, як показано в (5.4.62). Потім, згідно з (5.4.55), запишемо

$$y_t = \alpha + \alpha_0 Z_{0t} + \alpha_1 Z_{1t} + \alpha_2 Z_{2t} + \varepsilon_t, \quad (5.4.62)$$

де

$$\begin{aligned} Z_{0t} &= \sum_{i=0}^3 x_{t-i} = (x_t + x_{t-1} + x_{t-2} + x_{t-3}); \\ Z_{1t} &= \sum_{i=0}^3 i x_{t-i} = (x_{t-1} + 2x_{t-2} + 3x_{t-3}); \\ Z_{2t} &= \sum_{i=0}^3 i^2 x_{t-i} = (x_{t-1} + 4x_{t-2} + 9x_{t-3}). \end{aligned} \quad (5.4.63)$$

Змінні Z утворюються, як показано в табл. 5.8. Використовуючи дані значень y і Z_i , отримуємо такі результати регресії:

$$\begin{aligned} \hat{y}_t &= -7140.7564 + 0.6612Z_{0t} + 0.9020Z_{1t} - 0.4322Z_{2t} \quad (5.4.64) \\ &\quad (1992.9809) (0.1655) \quad (0.4831) \quad (0.1665) \\ t &= (-4.0847) (3.9960) \quad (1.8671) \quad (-2.5961) \\ \bar{R}^2 &= 0.9961 \quad df=13. \end{aligned}$$

(*Зауваження:* оскільки ми припустили, що лаг дорівнює 3, загальна кількість спостережень зменшена від 20 до 17).

За допомогою оцінених коефіцієнтів a в рівнянні (5.4.54) оцінимо коефіцієнти β із співвідношень (5.4.58) таким чином:

$$\begin{aligned} \hat{\beta}_0 &= \hat{\alpha}_0 = 0.6612; \\ \hat{\beta}_1 &= (\hat{\alpha}_0 + \hat{\alpha}_1 + \hat{\alpha}_2) = (0.6612 + 0.9020 - 0.4322) = 1.1310; \\ \hat{\beta}_2 &= (\hat{\alpha}_0 + 2\hat{\alpha}_1 + 4\hat{\alpha}_2) = [0.6612 + 2(0.9020) - 4(0.4322)] = 0.7364; \\ \hat{\beta}_3 &= (\hat{\alpha}_0 + 3\hat{\alpha}_1 + 9\hat{\alpha}_2) = [0.6612 + 3(0.9020) - 9(0.4322)] = -0.5226. \end{aligned}$$

Таблиця 5.8

Рік	y	x	Z_0	Z_1	Z_2
1977	45.069	26.480
1978	50.642	27.740
1979	51.871	28.736
1980	50.070	27.280	110.236	163.656	378.016
1981	52.707	30.219	113.975	167.972	391.884
1982	53.814	30.796	117.031	170.987	397.963
1983	54.939	30.896	119.191	173.074	397.192
1984	58.213	33.113	125.024	183.145	426.051
1985	60.043	35.032	129.837	187.293	433.861
1986	63.383	37.335	136.376	193.946	445.548
1987	68.221	41.003	143.483	206.738	475.480
1988	77.965	44.869	158.239	220.769	505.631
1989	84.655	46.449	169.656	238.880	544.896
1990	90.875	50.282	182.603	259.196	594.952
1991	97.074	53.555	195.155	227.787	639.899
1992	101.645	52.859	203.145	293.466	672.724
1993	102.445	55.917	212.613	310.815	719.617
1994	107.719	62.017	224.348	322.300	749.348
1995	120.870	71.398	242.191	332.428	761.416
1996	147.135	82.078	271.410	363.183	822.719

Отже, оцінена дистрибутивно-лагова модель має такий вигляд:

$$\begin{aligned} y_t &= -7140.7564 + 0.6612x_t + 1.1310x_{t-1} + 0.7364x_{t-2} - 0.5226x_{t-3} \quad (5.4.65) \\ &\quad (1992.9803) (0.1655) \quad (0.1798)50 \quad (0.1625)50 \quad (0.2307)50 \\ t &= (-4.0847) (3.9960) \quad (6.2903) \quad (4.5317) \quad (-2.2653). \end{aligned}$$

Використаємо наш приклад, щоб зазначити деякі додаткові властивості методу Альмона.

1. Стандартні помилки коефіцієнтів a отримані безпосередньо з методу найменших квадратів, але стандартні помилки деяких оцінених коефіцієнтів β , що є нашою головною метою, не можна отримати таким чином. Ці стандартні помилки можна легко обчислити з оцінених коефіцієнтів a , використовуючи відому формулу із статистики (див. вправу 5.4.21).

2. Оцінки коефіцієнтів β , отримані в (5.4.65), називаються *необмеженими оцінками* в тому сенсі, що на них не накладається жодних попередніх обмежень. Однак у деяких ситуаціях на β_i можуть бути накладені так звані кінцеві точкові обмеження, якщо припустити, що β_0 і β_k (поточний і k -ий лаговий коефіцієнт) дорівнюють нулеві. Через психологічні, інституціональні і технологічні причини значення пояснювальної змінної в поточному періоді може й не мати жодного впливу на поточне значення залежної змінної, що, таким чином, виправдовує нульове значення β_0 . З тих самих причин після певного часу k пояснювальна змінна може й не

впливати на залежну змінну, тобто і β_k теж дорівнюватиме нулеві. Також інколи при оцінці коефіцієнтів β_i на суму їх накладається таке обмеження: вона повинна дорівнювати одиниці.

5.4.12. Причинність в економіці: тест Гренжера (Granger)

Хоча регресійний аналіз розглядає залежність однієї змінної від інших, це зовсім не обов'язково означає, що між змінними і явищами є причинний зв'язок. Розглянемо таку ситуацію. Припустимо, що дві змінні, скажімо валовий національний продукт (ВНП) і грошова пропозиція M , впливають одна на одну з деяким (дистрибутивним) лагом. Чи можливо у такому випадку сказати, що саме зміна грошової пропозиції приводить до зміни ВНП, чи саме зміна ВНП змінює грошову пропозицію, або чи є зворотний зв'язок між цими двома впливами? Якщо коротко, питання полягає в тому, чи можна *статистично* визначити напрям причинності (взаємовідношення “причина-наслідок”).

Не будемо дуже заглиблюватись у це питання, оскільки воно заведе нас далеко в бік від теми нашої розмови, а лише розглянемо відносно простий тест причинності, запропонований Гренжером. Пояснимо цей тест на прикладі взаємовідношення між ВНП і M .

Тест Гренжера

Тест причинності Гренжера припускає, що інформація, потрібна для прогнозування відповідних змінних, у нашому випадку ВНП і M , повністю міститься в часових рядах. Тест включає оцінювання двох регресій:

$$GNP_t = \sum_{i=1}^n \alpha_i M_{t-i} + \sum_{j=1}^n \beta_j GNP_{t-j} + \varepsilon_{1t}; \quad (5.4.66)$$

$$M_t = \sum_{i=1}^m \lambda_i M_{t-i} + \sum_{j=1}^n \delta_j GNP_{t-j} + \varepsilon_{2t}, \quad (5.4.67)$$

де припускається, що випадкові величини ε_{1t} і ε_{2t} не корельовані між собою.

Рівняння (5.4.66) показує, що поточне значення ВНП залежить як від попередніх значень самого ВНП, так і від M , а (5.4.67) визначає подібну поведінку і для M . Розглянемо чотири випадки.

1. *Однобічна причинність від M до ВНП наявна*, якщо оцінені коефіцієнти при M в (5.4.66) статистично відрізняються від нуля як група (тобто $\sum \alpha_i \neq 0$), а множина оцінених коефіцієнтів при ВНП в (5.4.67) статистично не відрізняються від нуля ($\sum \delta_i = 0$).

2. *Навпаки, однобічна причинність від ВНП до M наявна*, коли множина коефіцієнтів при M в (5.4.66) статистично не відрізняються від нуля (тобто $\sum \alpha_i = 0$), а множина коефіцієнтів при ВНП в (5.4.67) статистично

відрізняються від нуля (тобто $\sum \delta_i \neq 0$).

3. *Зворотний зв'язок, або двобічна причинність, буває*, коли множини коефіцієнтів при M і ВНП статистично відрізняються від нуля в обох регресіях.

4. *Нарешті, між M і ВНП причинного зв'язку немає*, тобто вони *незалежні*, коли множини коефіцієнтів при M і ВНП статистично не відрізняються від нуля в обох регресіях.

Емпіричні результати

Р.Гафер застосував тест Гренжера, щоб з'ясувати характер причинності між ВНП і M для США за період від січня 1960 р. до квітня 1980 р. Він використав чотири затриманих значення двох змінних у кожній з двох регресій і отримав такі результати (див. табл. 5.9).

Таблиця 5.9

Напрямок причинності	Значення F	Рішення
$M \rightarrow y$	2.68	Не відкидати гіпотезу
$y \rightarrow M$	0.56	Відкинути гіпотезу

Ці результати показують, що напрямок причинності — від M до ВНП, оскільки оцінене значення F значиме при 5% -му рівні довіри; критичне значення F дорівнює 2.50 (для ступенів вільності відповідно 4 та 71). З іншого боку, немає “зворотного причинного зв'язку” між ВНП і M , оскільки оцінене значення F статистично незначиме.

5.4.13. Основні висновки про авторегресивні моделі

Через психологічні, технологічні та інституціональні причини для прийняття і виконання якогось економічного рішення потрібний певний час. У результаті цього економічна залежна змінна y може реагувати на зміну в економічній визначальній змінній x через деякий проміжок часу. Цей проміжок часу називається *лагом*, а *регресійні моделі, в які включений такий лаг, називаються лаговими регресійними моделями*.

Є два типи лагових змінних: лагові пояснювальні змінні (які або нестохастичні, або, якщо стохастичні, то розподіляються незалежно від випадкової величини) і лагові залежні змінні. *Регресійні моделі, які включають у себе як поточні, так і попередні значення нестохастичних змінних x , мають назву дистрибутивно-лагові моделі*, тому що вплив пояснювальної змінної (змінних) на залежну змінну поширюється або розподіляється на декілька часових періодів. Крім того, регресійні моделі, в яких попередні значення залежної змінної включені разом з пояснювальними змінними, мають назву *авторегресивні моделі*.

Якщо дистрибутивно-лагова модель містить кілька лагів, оцінка її ме-

тодом найменших квадратів хоча й можлива в принципі, але ускладнена на практиці, тому що це може призвести до проблеми мультиколінеарності. Як відомо, за наявності мультиколінеарності, метод найменших квадратів дає хоча й незміщені, але вкрай неточні оцінки. Для різних лагових коефіцієнтів потрібно зробити деякі попередні припущення. Однією з таких процедур є широко вживана дистрибутивно-лагова модель Койка, яка припускає, що лагові коефіцієнти спадають у геометричній прогресії при віддаленні у минуле. За цієї умови модель, яка містить нескінченну кількість лагів, можна скоротити до моделі, яка містить лише поточне значення нестохастичної змінної (змінних) x і одне попереднє значення залежної змінної серед її пояснювальних змінних.

Незважаючи на відмінні результати, це спрощення має свою ціну. Модель Койка створює деякі серйозні статистичні проблеми, оскільки вона містить стохастичну пояснювальну змінну (лагову y_{t-1}), яка може бути сильно корельованою із випадковою величиною. В цій ситуації економічна теорія показує, що оцінки, обчислені за допомогою методу найменших квадратів, не лише зміщені, а й неконсистентні: тобто якщо навіть розмір вибірки нескінченно збільшити, оцінки не наблизяться до своїх справжніх значень генеральної сукупності. Якщо коротко, вони залишаться асимптотично зміщеними. Таким чином, необхідно застосовувати альтернативні методи оцінювання.

У цьому розділі ми розглядали один із таких методів, а саме *метод допоміжних змінних*. Ключова ідея цього методу полягає в тому, щоб замінити лагову стохастичну пояснювальну змінну y_{t-1} на іншу змінну, яка б була сильно корельованою з y_{t-1} , але некорельованою з випадковою величиною. Цей метод забезпечує консистентні оцінки.

Хоча в емпіричній економетриці модель Койка досить популярна, вона не має теоретичного підґрунтя. Це ускладнення подолане за допомогою моделі адаптивних очікувань і моделі часткових пристосувань. У цих моделях враховується, яким чином економічні агенти формують свої очікування щодо невизначених економічних подій і як вони пристосовуються, якщо їхні очікування не збігаються із дійсністю. Унікальна відмінність обох цих моделей полягає в тому, що вони у своїй кінцевій формі нагадують модель Койка: вони також авторегресивні і використовують ті самі змінні. В моделі адаптивних очікувань з'являється така ж проблема, як і в моделі Койка. Модель часткових пристосувань можна оцінити за допомогою звичайного методу найменших квадратів.

Незважаючи на популярність, модель адаптивних очікувань критикується прихильниками моделі раціональних очікувань, які стверджують, що очікування спрямовані наперед, оскільки вони використовують усю доступну інформацію, а не лише минулу, як це стверджують прихильники гіпотези адаптивних очікувань. Але гіпотеза раціональних очікувань теж має своїх критиків. Приймавши початкові припущення гіпотези раціо-

нальних очікувань, вони наполягають на тому, що люди зовсім не обов'язково формують свої сподівання саме таким чином, як стверджує ця гіпотеза. Ця тема дуже суперечлива, і математичне обґрунтування гіпотези раціональних очікувань не входить до мети цієї книги.

Як було зазначено раніше, d -тест Дарбіна — Уотсона не прийнятний для перевірки автокореляції (першого порядку) в авторегресивних моделях, тому що в них значення d перевищує 2, а це значення очікується лише в справді випадкових послідовностях. Нещодавно Дарбін сам запропонував так звану h -статистику для тестування серійної кореляції в авторегресивних моделях. Однак цей тест можна застосовувати лише для великих вибірок.

Альтернативою підходу Койка до дистрибутивно-лагових моделей є поліноміальна дистрибутивно-лагова модель Ш.Альмона. Базуючись на теоремі Веєрштрассе, Альмон припустив, що лагові коефіцієнти β_i можна апроксимувати поліномом відповідного ступеня від i , тривалості лагу. Хоча метод Альмона уникає певних проблем, пов'язаних з моделлю Койка, його практична слабкість полягає в тому, що як ступінь поліному, так і максимальну довжину лагу дослідник повинен визначити перед початком самого дослідження.

Незважаючи на проблеми, що трапляються при оцінюванні, дистрибутивно-лагові і авторегресивні моделі виявились дуже корисними в емпіричній економіці, тому що вони перетворюють моделі, які б у будь-якому іншому випадку залишилися статичними, на динамічні, за допомогою фактору часу. Такі моделі допомагають розрізняти короткостроковий і довгостроковий вплив на залежну змінну при одиничній зміні значення незалежної змінної (змінних). Таким чином, для оцінювання коротко- і довгострокової еластичності за ціною, доходом, нормою заміни та іншими схожими показникам такі моделі виявились дуже корисними.

Тема причин і наслідків, або причинності, дуже важлива в економіці, як і в науці в цілому. Але вона досить складна, ставить велику кількість різноманітних філософських питань. Ми лише ледь торкнулися цієї теми, коли розглядали статистичний тест причинності, запропонований Гренжером. Але читач має пам'ятати, що це не єдиний тест на перевірку наявності причинного зв'язку для часових рядів. Взагалі питання причинного зв'язку розглядаються в теорії коінтеграції, що виходить за межі цього підручника.

5.5. DUMMY-ЗМІННІ

5.5.1. Природа dummy-змінних

Досі ми розглядали рівняння багатфакторної регресії, де фактори набували кількісних значень. Звичайно, економічні явища набагато різноманітніші. На залежну змінну поряд з кількісними факторами впливають і якісні: якість продукції, стать, релігія, страйки, війни, зміни в економічній політиці і т. ін.

Потрібно вміти вводити якісні дані в багатфакторні регресійні моделі, оцінювати параметри і аналізувати отримані результати. **Часто якісні змінні є бінарними: вони отримують “значення 1” при наявності певної якості і “значення 0” при її відсутності. Такі змінні називаються dummy-змінними.** Звичайно dummy-змінні не обов’язково приймають значення (0,1). Пара (0,1) може легко трансформуватись у будь-яку іншу пару лінійним перетворення $y = a + bz$ ($b \neq 0$), де a та b константи, а $z = 1$ або нулю. Наприклад, коли $z = 1$, $y = a + b$, а коли $z = 0$, $y = a$.

Dummy-змінні можуть використовуватись у регресійних моделях поряд з кількісними змінними, а можуть утворювати регресійні моделі, в яких усі фактори є dummy-змінними. Такі моделі називають АOV-моделями (analysis-of-variance).

Розглянемо як приклад найпростішу лінійну регресійну модель залежності рейтингу студентів НаУКМА від успішного закінчення навчання в школі:

$$y_i = \alpha_0 + \alpha_1 d_i + \varepsilon_i, \quad (5.5.1)$$

де y_i — триместровий середній рейтинг успішності студента 1 курсу;
 d_i — (dummy-змінна) = 1, якщо студент відмінник у школі, або його середній бал >4.8 ; =0, у протилежному випадку.

Модель (5.5.1) встановлює залежність успішного навчання в університеті від успішного навчання в школі. Як і у разі класичної лінійної регресійної моделі, ми отримуємо:

$$\text{середній рейтинг відмінників } E(y_i | d_i = 1) = \alpha_0 + \alpha_1; \quad (5.5.2)$$

$$\text{середній рейтинг невідмінників } E(y_i | d_i = 0) = \alpha_0.$$

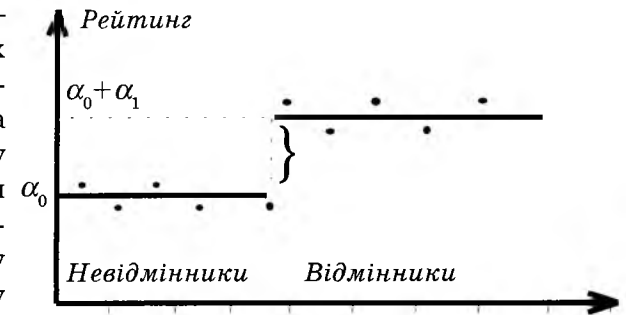
У цьому разі перетин α_0 показує середній рейтинг невідмінників; $(\alpha_0 + \alpha_1)$ — середній рейтинг відмінників; а нахил α_1 — наскільки середній рейтинг відмінників відрізняється від середнього рейтингу невідмінників.

Базуючись на реальних даних рейтингу студентів першого курсу НаУКМА, проведемо розрахунки за моделлю (5.5.1). Графічно отримані результати подано на мал.5.8.

Як бачимо з графіка, вхідні дані розподіляються на дві групи: відмінники та невідмінники. Сама функція є ступінчастою функцією — середній рейтинг студентів невідмінників та середній рейтинг відмінників.

5.5.2. Регресія однієї кількісної та однієї якісної змінної двох класів, або категорій

Звичайно, більш поширеним випадком в економічних дослідженнях є випадок змішаних факторів — якісних та кількісних. Моделі такого типу називаються АCOV-моделями (analysis of covariance). Розглянемо як приклад найпростішу АCOV-модель, яка вміщує одну якісну та одну кількісну змінну.



Малюнок 5.8. Функція залежності рейтингу студентів НаУКМА від успішності в школі

Ускладнимо попередню розглянуту модель і додамо ще один фактор — бал вступного іспиту; отримаємо таку АCOV-модель:

$$y_i = \alpha_0 + \alpha_1 d_i + \beta_1 x_i + \varepsilon_i, \quad (5.5.3)$$

де y_i — середній триместровий рейтинг студента;
 $d_i = 1$, якщо студент у минулому відмінник;
 $= 0$, у протилежному випадку;
 x_i — бал на вступному іспиті;
 ε — випадкова величина.

Припускаючи, як завжди, $E(\varepsilon_i) = 0$, отримуємо, що середній рейтинг невідмінника

$$E(y_i | d_i = 0) = \alpha_0 + \beta_1 x_i; \quad (5.5.4)$$

середній рейтинг відмінника

$$E(y_i | d_i = 1) = (\alpha_0 + \alpha_1) + \beta_1 x_i. \quad (5.5.5)$$

Геометрично АCOV-модель зображено на мал. 5.9 ($\alpha_0 > 0$).

Як бачимо, модель (5.5.3) залежно від відмінників і невідмінників розпадається на дві функції з однаковим нахилом, але різним перетином.

Іншими словами, рівень середнього триместрового рейтингу студентів відмінників відрізняється від рейтингу невідмінників на α_1 , але величина зміни середнього триместрового рейтингу студентів щодо балу, отриманого на вступному тестуванні, однакова для відмінників та невідмінників.

При умові, що нахил статистично значимий, зробимо перевірку гіпотези, за якою дві регресії (5.5.4) і (5.5.5) мають однаковий перетин (тобто немає впливу відмінної успішності в школі). Це можна легко здійснити при розгляді регресії (5.5.3) і визначити статистичну значимість оцінено-

го коефіцієнта α_1 на основі традиційного t -тесту. Якщо t -тест показує, що α_1 статистично значимий, нуль-гіпотезу про те, що дві регресії мають однаковий перетин — відхиляємо.

Задамо запитання, що додають до регресії dummy -змінні, які особливості виникають при введенні dummy -змінних у модель?

По-перше, dummy -змінні відокремлюють різні класи або різні категорії. У нашому прикладі завдяки введенню dummy -змінних ми розрізняли дві категорії студентів: відмінників та невідмінників.

Отже, одна dummy -змінна може розрізняти дві категорії. Припустимо, що ми хочемо ускладнити наш приклад, поділивши ще детальніше категорії. Що при цьому виникає? Розглянемо таку модель:

$$y_i = \alpha_0 + \alpha_1 d_{1i} + \alpha_2 d_{2i} + \beta_1 x_{1i} + \varepsilon_i, \quad (5.5.6)$$

де y_i і x_{1i} визначені попередньо;

$d_1 = 1$, якщо студент відмінник;

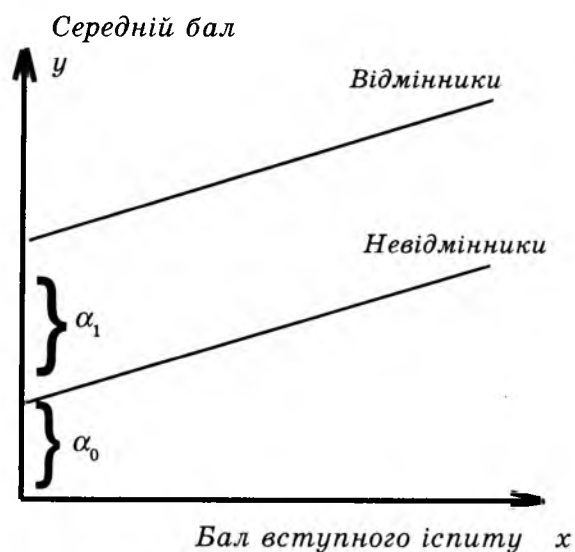
$= 0$ в усіх інших випадках;

$d_2 = 0$, якщо студент невідмінник;

$= 1$ в усіх інших випадках.

Модель (5.5.6) не може бути оцінена через повну колінеарність між d_1 і d_2 . Для пересвідчення, припустимо, що ми маємо модель, яка складається з трьох студентів відмінників та двох невідмінників. Інформаційна матриця при цьому виглядатиме таким чином:

		d_1	d_2	x	
Відмінник	y_1	1	1	0	x_1
Відмінник	y_2	1	1	0	x_2
Невідмінник	y_3	1	0	1	x_3
Відмінник	y_4	1	1	0	x_4
Невідмінник	y_5	1	0	1	x_5



Малюнок 5.9. Функція залежності рейтингу студентів НАУКМА від успішності в школі та балу на вступному іспиті

Перша колонка праворуч відображає перетин α_0 . Тепер можна побачити, що $d_1 = 1 - d_2$, або $d_2 = 1 - d_1$, тобто d_1 і d_2 — колінеарні. У разі досконалої мультиколінеарності звичайне оцінювання методом найменших квадратів неможливе. Найпростіший спосіб вирішення цієї проблеми — визначити dummy -змінні способом, описаним у моделі (5.5.3), скажімо, використати тільки одну dummy -змінну, якщо є тільки дві категорії або два класи якісної змінної. В даному випадку, попередня матриця інформації не матиме колонки d_2 , і проблема мультиколінеарності зникає. Тому в загальному випадку для уникнення проблеми мультиколінеарності потрібно дотримуватися такого правила: якщо якісна змінна має m категорій, у модель вводимо лише $m-1$ dummy -змінну. В нашому прикладі кінцевий результат навчання в школі ми підрозділяли на дві категорії — відмінник та невідмінник, тому ми ввели у модель лише одну dummy -змінну. Якщо це правило не виконано, ми попадаємо в так звану dummy -пастку, тобто ситуацію досконалої мультиколінеарності.

По-друге, припустимо, що ми ввели інші значення 1 і 0 для того, щоб відрізнити дві категорії. Нехай тепер $D=1$ для невідмінників і $D=0$ для відмінників. Тоді дві регресії, отримані з (5.5.3), будуть:

$$\text{студенти-невідмінники: } E(y_i / d_{1i} = 1) = (\alpha_0 + \alpha_1) + \beta_1 x_{1i}; \quad (5.5.7)$$

$$\text{студенти-відмінники: } E(y_i / d_{1i} = 0) = \alpha_0 + \beta_1 x_{1i}. \quad (5.5.8)$$

Якщо порівняти ці рівняння з рівняннями (5.5.4) та (5.5.5) у попередніх моделях, то α_1 показує, наскільки середній рейтинг студента-відмінника в школі відрізняється від середнього рейтингу невідмінника в школі. При наявності залежності α_1 буде негативним, тоді як попередньо очікувалось, що воно матиме позитивне значення. Тому під час інтерпретації результатів моделей, які використовують dummy -змінні, важливо знати, які саме категорії позначались 1, а які 0.

По-третьє, клас, або категорія, позначена 0 (нулем), часто розглядається як базова, вихідна, контрольна категорія. Вона є базовою в тому розумінні, що порівняння робляться на основі саме цієї категорії. Таким чином, у моделі (5.33) студент-невідмінник у школі — базова категорія. Перетин α_0 є перетином базової категорії в тому сенсі, що при розгляді регресії стосовно студентів-невідмінників $D=0$ перетин дорівнюватиме α_0 .

По-четверте, коефіцієнт біля dummy -змінної D називається диференціальним коефіцієнтом перетину, тому що він показує, наскільки значення перетину першої категорії відрізняється від значення перетину базової категорії.

5.5.3. Регресія кількісної змінної та однієї якісної змінної з більш ніж двома класами

Припустимо, що, базуючись на міжгалузевій інформації, ми хочемо описати регресійний зв'язок щорічних оздоровчих витрат людини залежно від освіти та доходу. Оскільки освіта є якісною змінною, припустимо, що розглядаємо три взаємовиключні рівні освіти: середня, середня спеціальна і вища. Тепер, на відміну від попереднього випадку, маємо більш ніж дві категорії якісних змінних. Застосовуючи правило, що кількість змінних має бути на одиницю менша від кількості категорій, нам потрібно подати дві *dummy*-змінні, щоб простежити за двома рівнями освіти. Припустимо, три рівні освіти мають однаковий нахил, але різні перетини.

Модель щорічних оздоровчих витрат залежно від доходу матиме вигляд:

$$y_i = \alpha_0 + \alpha_1 D_{1i} + \alpha_2 D_{2i} + \beta_1 x_{1i} + \varepsilon_i, \quad (5.5.9)$$

де y_i — щорічні оздоровчі витрати;

x_i — щорічний дохід;

$D_1=1$ — середня освіта;

=0 — у всіх інших випадках;

$D_2=1$ — вища освіта;

=0 — у всіх інших випадках.

Аналогічно з попереднім завданням ми трактуємо середню спеціальну освіту як базову категорію. Тому α_0 відобразить перетин саме цієї категорії. Диференційні α_1 і α_2 показують, наскільки перетини двох інших категорій відрізняються від перетину базової категорії. Це можна перевірити таким чином: припустимо, $E(\varepsilon_i) = 0$, тоді з (5.38) отримаємо:

$$E(y_i / D_1 = 0, D_2 = 0, x_i) = \alpha_0 + \beta_1 x_{1i}; \quad (5.5.10)$$

$$E(y_i / D_1 = 1, D_2 = 0, x_i) = (\alpha_0 + \alpha_1) + \beta_1 x_{1i}; \quad (5.5.11)$$

$$E(y_i / D_1 = 0, D_2 = 1, x_i) = (\alpha_0 + \alpha_2) + \beta_1 x_{1i}, \quad (5.5.12)$$

які є функціями середніх оздоровчих витрат для трьох рівнів освіти: середньої, середньої спеціальної та вищої.

Після розгляду регресії (5.5.9) можна встановити, чи диференційні перетини α_1 і α_2 статистично значимі, тобто чи відрізняються від базової групи. Гіпотезу $\alpha_1 = \alpha_2 = 0$ можна перевірити за допомогою AOV або *F*-тесту, як показано в розділі 4.

Інтерпретація регресії (5.5.9) зміниться, якщо ми візьмемо іншу схему визначення *dummy*-змінних. Тобто, якщо припустити, що $D_1=1$ — “середня освіта”, і $D_2 = 1$ — “категорія середньої спеціальної освіти”, то базовою категорією буде “вища освіта” і всі порівняння будуть пов'язані з цією категорією.

5.5.4. Регресія однієї кількісної і двох якісних змінних

Техніку *dummy*-змінних можна розширити, маючи справу більше ніж з однією якісною змінною. Повернемося до регресії середнього рейтингу студентів (5.5.3), але тепер до навчання в школі та вступного балу додамо стать — чоловічу або жіночу. Тепер (5.5.3) можна виразити як

$$y_i = \alpha_0 + \alpha_1 D_{1i} + \alpha_2 D_{2i} + \beta_1 x_{1i} + \varepsilon_i, \quad (5.5.13)$$

де y_i та x_i — середній триместровий рейтинг студента та бал вступного іспиту;

$D_1=1$ — якщо студент-відмінник;

=0 — в іншому випадку;

$D_2=1$ — якщо студент-чоловік;

=0 — в іншому випадку.

Зазначимо, що кожна з двох якісних змінних, навчання в школі та стать, мають дві категорії і тому потребують *dummy*-змінну для кожної. Базовою категорією буде “невідмінник-дівчина”. Припускаємо, що $E(\varepsilon_i) = 0$, і з (5.5.13) отримуємо такі регресії:

Середній рейтинг невідмінниці-дівчини:

$$E(y_i / D_1 = 0, D_2 = 0, x_i) = \alpha_0 + \beta_1 x_{1i}. \quad (5.5.14)$$

Середній рейтинг невідмінника-хлопця;

$$E(y_i / D_1 = 1, D_2 = 0, x_i) = (\alpha_0 + \alpha_1) + \beta_1 x_{1i}. \quad (5.5.15)$$

Середній рейтинг відмінниці-дівчини:

$$E(y_i / D_1 = 0, D_2 = 1, x_i) = (\alpha_0 + \alpha_2) + \beta_1 x_{1i}. \quad (5.5.16)$$

Середній рейтинг відмінника-хлопця:

$$E(y_i / D_1 = 1, D_2 = 1, x_i) = (\alpha_0 + \alpha_1 + \alpha_2) + \beta_1 x_{1i}. \quad (5.5.17)$$

Нагадаємо, що попередні регресії відрізняються лише коефіцієнтами перетину, а не нахилу.

За допомогою оцінювання методом найменших квадратів неможливо перевірити всі гіпотези. Але якщо α_2 статистично значимий, це означить, що стать людини впливатиме на середній триместровий рейтинг студентів. Аналогічно, якщо α_1 буде статистично значимим, це означить, що навчання в школі також впливатиме на середній рейтинг студентів. А якщо обидва диференційні перетини статистично значимі, це означить, що навчання в школі та стать є важливими детермінантами середнього рейтингу.

Отже, ми можемо розширити нашу модель, включивши більше ніж одну кількісну змінну та більше ніж одну якісну змінну. Єдине застереження: кількість думпу-змінних для кожної якісної змінної має бути на одиницю меншою від категорії цієї змінної. Наведемо приклад.

Приклад 5.5.1. Додаткові заробітки¹

Людина, яка працює на кількох роботах, відома як заробітчанин. Шіско і Росткер зацікавилися факторами, які визначають зарплати заробітчанин. Базуючись на вибірці щодо 318 заробітчанин, вони отримали таку регресію:

$$w_m = 37.07 + 0.403w_0 \quad (0.062)$$

$$90.06D_1 + 75.51D_2 + 47.33D_3 + 113.64D_4 + 2.26x_1 \quad (5.5.18)$$

(24.47) (21.60) (23.42) (27.62) (0.94)

$$R^2 = 0.34 \quad df = 311,$$

де w_m — зарплата заробітчанина (центів на годину);

w_0 — початкова зарплата (центів на годину);

D_1 — раса = 0, якщо білий, = 1, якщо не білий;

D_2 — місто = 0, якщо не міський, = 1, якщо міський;

D_4 — область = 0, якщо не західний район, = 1, якщо західний район;

D_3 — освіта = 0 — без освіти, = 1 має освіту;

x_1 — вік людини.

У моделі (5.47) є дві кількісні пояснюючі змінні, w_0 та x_1 , а також чотири якісні змінні. Коефіцієнти всіх цих змінних статистично значимі з рівнем похибки 5%. Цікаво, що всі якісні змінні суттєво впливають на зарплату заробітчанин. Наприклад, за інших рівних умов рівень погодинної оплати праці на 47% вищий у осіб з освітою, ніж у тих, хто не має освіти.

З регресії (5.5.18) можна виділити декілька індивідуальних регресій, дві з яких такі: середня зарплата білих, неміських, не із західних регіонів заробітчанин, які не мають вищої освіти (коли всі думпу-змінні дорівнюють нулю):

$$w_m = 37.07 + 0.403w_0 + 2.26x_1. \quad (5.5.19)$$

Середня зарплата чорношкірого, міського, з західного регіону заробітчанина з вищою освітою (коли всі думпу-змінні рівні одиниці):

$$w_m = 183.49 + 0.403w_0 + 2.26x_1. \quad (5.5.20)$$

¹ Shisko R., Rostker B. The Economics of Multiple Job Holding//The Economic Review. — 1976. — Vol. 66. — №3. — June.

5.5.5. Порівняння двох регресійних моделей

Дотепер у розглянутих моделях ми припускали, що якісні змінні впливають лише на перетин, але не на нахил. А що, коли нахили також різні? Якщо нахили відрізняються, то перевірка відмінностей у перетинах фактично мало важлива. Тому потрібно розвинути загальну методологію знаходження відмінностей у двох (або більше) регресіях. Відмінність може бути в перетинах, нахилах або в обох випадках. Щоб простежити, як це відбувається, розглянемо приклад.

Приклад 5.5.2. Заощадження та дохід у Великобританії в 1946 — 1963 рр.¹

У таблиці 5.10 наведено інформацію про доходи та заощадження у Великобританії.

Таблиця 5.10

Особисті заощадження та доходи у Великобританії в 1946 — 1963 рр.
(млн.фунтів)

Період 1	Заощадження	Доходи	Період 2	Заощадження	Доходи
1946	0.36	8.8	1955	0.59	15.5
1947	0.21	9.4	1956	0.90	16.7
1948	0.08	10.0	1957	0.95	17.7
1949	0.20	10.6	1958	0.82	18.6
1950	0.10	11.0	1959	1.004	19.7
1951	0.12	11.9	1960	1.53	21.1
1952	0.41	12.7	1961	1.94	22.8
1953	0.50	13.5	1962	1.75	23.9
1954	0.43	14.3	1963	1.99	25.2

Інформація в таблиці поділена на два періоди, 1946 — 1954 рр. (одразу після Другої світової війни, період реконструкції) і 1955 — 1963 рр. (період після реконструкції). Припустимо, що ми хочемо дослідити, чи змінився взаємозв'язок агрегованих доходів та заощаджень між двома періодами.

Період реконструкції: $y_i = \alpha_0 + \alpha_1 x_i + \varepsilon_{1i}$; $i = 1, 2, \dots, N_1$. (5.5.21)

Після періоду реконструкції: $y_i = \gamma_0 + \gamma_1 x_i + \varepsilon_{2i}$; $i = 1, 2, \dots, N_2$, (5.5.22)

де y — заощадження (мільйони фунтів);

x — доходи (мільйони фунтів);

$\varepsilon_{1i}, \varepsilon_{2i}$ — відхилення в двох регресіях.

Звичайно потрібно, щоб кількість спостережень N_1 та N_2 в обох групах (періодах) була однаковою.

Між регресіями (5.5.21) і (5.5.22) можливі такі чотири варіанти співвідношень.

1. $\alpha_0 = \gamma_0$ і $\alpha_1 = \gamma_1$, тобто дві регресії ідентичні.

2. $\alpha_0 \neq \gamma_0$, але $\alpha_1 = \gamma_1$, тобто дві регресії відрізняються лише за розміщенням (перетинами).

¹ Gyjarati D.N. Basic Econometrics. — 2nd ed. — New York: McGraw-Hill Book Company, 1978.

3. $\alpha_0 = \gamma_0$, але $\alpha_1 \neq \gamma_1$, дві регресії мають однакові перетини, але різні нахили.

4. $\alpha_0 \neq \gamma_0$ і $\alpha_1 \neq \gamma_1$, обидві регресії абсолютно різні.

Використовуючи дані табл. 5.10, можна записати дві окремі регресії (5.5.21) і (5.5.22), а потім застосувати відповідні статистичні засоби для перевірки всіх попередніх варіантів. Одним з таких засобів є Chow-тест, іншим — *dummy*-змінні. Ми обговоримо це в наступних двох частинах.

5.5.5.1. Chow-тест для порівняння регресійних моделей

Одним з відомих методів тестування відмінностей між двома (і більше) регресіями є Chow-тест. Цей тест базується на таких припущеннях.

1. Випадкові величини в обох регресійних моделях — нормально розподілені випадкові величини з математичним сподіванням нуль та постійною дисперсією:

$$\varepsilon_{1i} \approx N(0, \sigma^2),$$

$$\varepsilon_{2i} \approx N(0, \sigma^2).$$

2. Випадкові величини $\varepsilon_{1i}, \varepsilon_{2i}$ незалежно розподілені.

Враховуючи ці припущення, використаємо інформацію табл. 5.10 про доходи та заощадження для застосування Chow-тесту.

Крок 1. Об'єднаємо всі спостереження N_1 та N_2 в обох періодах і розглянемо одну “збірну” регресію (*примітка*: в даному випадку $N_1 = N_2 = 9$).

$$y_t = \alpha + \beta x_t + \varepsilon_t. \quad (5.5.23)$$

З даної регресії отримаємо суму квадратів залишків (SSE) з $df = N_1 - N_2 - k$, де k означає кількість оцінених параметрів. Для нашого прикладу $k = 2$.

Крок 2. Розглянемо окремо дві регресії (5.5.21) і (5.5.22). Отримаємо їх SSE, скажімо, S_2 і S_3 , з $df = N_1 - k$ і $df = N_2 - k$, відповідно; k у даному випадку дорівнює 2. Додамо дві SS, $S_4 = S_2 + S_3$ з $df = N_1 + N_2 - 2k$.

Крок 3. Отримаємо $S_5 = S_1 - S_4$.

Крок 4. Застосуємо F -тест:

$$F = \frac{S_5 / k}{S_4 / (N_1 + N_2 - 2k)}, \quad (5.5.24)$$

де $df = (k; N_1 + N_2 - 2k)$. Якщо обчислене F перевищує критичне F , відхиляємо гіпотезу про те, що дві регресії однакові.

Застосуємо ці кроки до нашого прикладу доходів та заощаджень.

Крок 1

$$\hat{y}_t = -1.0821 + 0.1178x_t \quad (5.5.25)$$

(0.1451)(0.0088)

$$t = (-7.4576)(13.3864) \quad R_2 = 0.9185 \quad S_1 = 0.5722 \quad df = 16.$$

Крок 2

$$\hat{y}_t = -0.266 + 0.0470x_t \quad (5.5.26)$$

(0.3053) (0.0266)

$$t = (-0.8719)(1.7669) \quad R_2 = 0.3092 \quad S_2 = 0.1396 \quad df = 7.$$

Період після реконструкції:

$$\hat{y}_t = -1.7502 + 0.1504x_t \quad (5.5.27)$$

(0.3576) (0.0175)

$$t = (-4.88943)(8.5943) \quad R_2 = 0.9131 \quad S_3 = 0.1931 \quad df = 7$$

$$S_4 = S_2 + S_3 = 0.3327.$$

Крок 3

$$S_5 = S_1 - S_4 = 0.2395.$$

Крок 4

$$F = \frac{0.2395 / 2}{0.3327 / 14} = 5.04.$$

Тоді $F_{2,14}$ з рівнем значимості 5% дорівнює 3.74. Тому обчислене $F = 5.04$ статистично значиме. Можна зробити висновок, що дві регресії різні. Chow-тест є спеціальним застосуванням методу обмежених найменших квадратів: коли ми оцінюємо (5.5.23) замість (5.5.21) і (5.5.22), ми накладаємо обмеження $\alpha_0 = \gamma_0 = \alpha$ і $\alpha_1 = \gamma_1 = \beta$.

Якщо не виконана умова гомоскедастичності, то Chow-тест не можна застосовувати. Тому необхідно спочатку перевірити умову гомоскедастичності ($\sigma_1^2 = \sigma_2^2 = \sigma^2$), застосовуючи один (і більше) методів, обговорюваних раніше в розділі 5. Також має бути достатньо ступенів вільності, щоб розглядати регресії окремо. У винятковому випадку, якщо в прикладі доходів та заощаджень є тільки два спостереження для останнього періоду, буде нуль ступенів вільності $(N - 2) = (2 - 2) = 0$. В таких ситуаціях F -тест у (5.5.24) потрібно модифікувати (як це зробити ми покажемо у вправі 5.5.5).

Приклад, який ми обговорювали, мав дві групи, але Chow-тест можна легко узагальнити і для більше, ніж двох груп.

5.5.6. Порівняння двох регресій: підхід з використанням *dummy*-змінної

Процедуру Chow-тесту можна скоротити, використовуючи *dummy*-змінні. Хоча висновки, які ми отримуємо з Chow-тесту та тесту *dummy*-змінних однакові, застосування *dummy*-змінних дає деякі переваги.

Давайте об'єднаємо всі N_1 та N_2 спостереження і оцінимо таку регресію:

$$y_i = \alpha_1 + \alpha_2 D_i + \beta_1 x_i + \beta_2 (D_i x_i) + \varepsilon_i, \quad (5.5.28)$$

де y_i та x_i — відповідно заощадження і доход, $D_i = 1$ для спостережень у першому періоді та періоді реконструкції і $D_i = 0$ для спостережень у періоді після реконструкції.

Припускаючи, що $E(\varepsilon_i) = 0$, отримаємо:

$$E(y_i / D_i = 0, x_i) = \alpha_1 + \beta_1 x_i; \quad (5.5.29)$$

$$E(y_i / D_i = 1, x_i) = (\alpha_1 + \alpha_2) + (\beta_1 + \beta_2) x_i, \quad (5.5.30)$$

які є відповідно функціями середніх заощаджень для другого (після реконструкції) та першого (реконструкції) періодів і є такими ж як (5.5.22) і (5.5.21) з $\gamma_0 = \alpha_1$, $\gamma_1 = \beta_1$, $\alpha_0 = (\alpha_1 + \alpha_2)$ і $\alpha_1 = (\beta_1 + \beta_2)$. Тому оцінювання (5.5.28) є таким самим, як і оцінювання окремих функцій заощаджень (5.5.21) та (5.5.22).

У (5.5.28) α_2 є диференціальним перетином, а β_2 є диференціальним нахилом, який показує, наскільки нахил функції збережень першого періоду відрізняється від нахилу функції збережень другого періоду. Якщо перемножити D і x , то неможливо зрозуміти відмінність між нахилами двох періодів.

Повертаючись до інформації про доходи та заощадження з табл. 5.10, емпіричну частину (5.5.29) можемо записати таким чином:

$$\hat{y}_i = -1.7502 + 1.4839D_i + 0.1504x_i - 0.1034D_i x_i \quad (5.5.31)$$

(0.3319) (0.4704) (0.0163) (0.0332)

$$t = (-5.2733) (3.1545) (9.2270) (-3.1144) \quad \bar{R}^2 = 0.94425.$$

Як показує регресія, диференціальні коефіцієнти перетину та нахилу статистично значимі, вони чітко показують, що регресії для двох періодів різні. Тоді, розглядаючи (5.5.29) та (5.5.30), можемо отримати дві регресії (примітка: $D=1$ для першого періоду).

Період реконструкції:

$$\hat{y}_i = (-1.77502 + 1.4839) + (0.1504 - 0.1034)x_i = -0.2663 + 0.00470x_i. \quad (5.5.32)$$

Після періоду реконструкції:

$$\hat{y}_i = -1.7502 + 0.1504x_i. \quad (5.5.33)$$

Ці регресії такі ж самі, як і регресії, отримані за допомогою Chow-процедури, відповідно (5.5.26) і (5.5.27).

Тепер можемо продемонструвати переваги техніки *dummy*-змінних (оцінювання (5.5.28)) над Chow-тестом (оцінювання трьох регресій: (5.5.21), (5.5.22) і “зібраної” регресії окремо).

1. Нам потрібно розглядати тільки одну регресію, тому що окремі регресії можна легко вивести з неї, як показано в рівняннях (5.5.29) і (5.5.30).

2. Для множини гіпотез можна застосувати одну регресію. Однак, якщо диференціальний коефіцієнт перетину α_2 статистично незначимий, то можемо припустити гіпотезу, що дві регресії мають однаковий перетин. Відповідно, якщо диференціальний нахил β_2 статистично незначимий, але α_2 — значимий, щонайменше ми можемо не відхилити гіпотезу про те, що дві регресії мають однаковий нахил, тобто лінії двох регресій паралельні. Перевірку стабільності регресії ($\alpha_2 = \beta_2 = 0$) можна здійснити звичайним F -тестом. Якщо гіпотеза прийнята, лінія регресії збігається.

3. Chow-тест чітко не показує, який коефіцієнт — перетину чи нахилу — відмінний або (як у даному прикладі) обидва відмінні в двох періодах. Тобто можна отримати значимий Chow-тест, оскільки відмінний лише нахил або лише перетин, або обидва. Іншими словами, Chow-тест не може дати відповідь на запитання, який з чотирьох варіантів є в даному випадку. В цьому відношенні підхід *dummy*-змінних має чітку перевагу, тут він вказує не лише на відмінність регресій, а й на причини цієї відмінності — чи вона стосується перетину або нахилу, чи відразу обох, що дуже важливо на практиці.

4. Оскільки “складання” збільшує ступені вільності, можна довести відносну точність оцінених параметрів.

5.5.7. Особливий випадок одночасного впливу двох *dummy*-змінних

Розглянемо, що відбувається при одночасному впливі кількох *dummy*-змінних на досліджуваний показник. Для спрощення проілюструємо наші припущення прикладом. Для цього розглянемо модель:

$$y_j = \alpha_0 + \alpha_1 D_{1j} + \alpha_2 D_{2j} + \beta x_j + \varepsilon_j, \quad (5.5.38)$$

де y_j — щорічні витрати на одяг;

x_j — доход;

$D_1 = 1$, якщо це жінка;

$= 0$, якщо це чоловік;

$D_2 = 1$ з вищою освітою;

$= 0$ у всіх інших випадках.

У даній моделі припустимо, що *dummy*-змінна D_1 не залежить від рівня освіти, аналогічно, статева належність особи не впливає на *dummy*-змінну D_2 . Скажімо, якщо середні витрати у жінок на одяг більші, ніж у чоловіків, то це незалежно від того, мають вони вищу освіту чи ні. Відповідно, якщо особи з вищою освітою витрачають на одяг більше, ніж ті, хто її не має, то від статі це не залежить.

У багатьох застосуваннях таке припущення не дійсне. Освічена жінка, порівняно з неосвіченою, може дозволити собі купувати більше одягу. Тобто між двома якісними змінними D_1 і D_2 може бути зв'язок, і тому вони здатні впливати на середнє значення y не тільки як додатки (5.65), а й як множники, що приводить до такої моделі:

$$y_j = \alpha_0 + \alpha_1 D_{1j} + \alpha_2 D_{2j} + \alpha_3 (D_{1j} D_{2j}) + \beta x_j + \varepsilon_j. \quad (5.5.39)$$

З (5.5.39) отримуємо:

$$E(y_j / D_1 = 1, D_2 = 1, x_j) = (\alpha_0 + \alpha_1 + \alpha_2 + \alpha_3) + \beta x_j. \quad (5.5.40)$$

Модель (5.5.40) описує середні витрати на одяг жінок з вищою освітою.

У даній моделі:

α_1 — диференційний ефект витрат на придбання одягу жінкою;

α_2 — диференційний ефект витрат на придбання одягу людиною з вищою освітою;

α_3 — диференційний ефект витрат на придбання одягу жінкою з вищою освітою, який показує, що середні витрати жінок з вищою освітою на одяг відрізняються на величину α_3 від середніх витрат просто жінок або освічених чоловіків. Якщо всі значення α_1 , α_2 і α_3 позитивні, середні витрати на одяг жінок перевищують базову категорію (в даному разі це чоловіки без вищої освіти), але це перевищення набагато більше в освічених жінок. Аналогічно, середні витрати осіб з вищою освітою перевищують базову категорію, але це перевищення набагато більше в освічених жінок. Це приклад того, наскільки взаємозв'язок *dummy*-змінних модифікує ефект двох властивостей, що розглядалися окремо.

За допомогою *t*-тесту можна перевірити статистичну значимість коефіцієнта взаємодії *dummy*-змінних. Якщо він статистично значимий, то спільна дія двох властивостей послабиться або підсилиться.

5.5.8. Використання *dummy*-змінних у сезонному аналізі

Економічні процеси дуже часто підпорядковані сезонним коливанням. Прикладом можуть бути різдвяний розпродаж товарів, попит на гроші домогосподарств під час свят, попит на морозиво і напої влітку, сільськогосподарське виробництво. Крім того, багато первісної інформації (індекси споживчих, оптових цін, цін виробника) підпорядковані також впливу сезонності.

В економічному аналізі інколи виникає проблема вилучення сезонних коливань з метою виявлення тенденцій. Розглянемо один з методів вилучення сезонних коливань. Для прикладу розглянемо квартальну динаміку прибутків деяких приватних фірм України. Необхідні умовні дані містяться в наведеній нижче табл. 5.11. У ній також міститься інформація щодо способу підготовки матриці для об'єднання *dummy*-змінних.

Візуальний аналіз свідчить про наявність сезонності — обсяги прибутку і продажу найбільші в другому кварталі щороку. Використовуючи дані, побудуємо модель:

$$y_t = \alpha_0 + \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \beta x_t + \varepsilon_t, \quad (5.5.41)$$

де $D_1 = 1$ для другого кварталу, $D_1 = 0$ — в усіх інших випадках;

$D_2 = 1$ для третього кварталу, $D_2 = 0$ — в усіх інших випадках;

$D_3 = 1$ для четвертого кварталу, $D_3 = 0$ — в усіх інших випадках.

Припустимо, що змінна “сезон” має чотири класи (чотири квартали). Отже, необхідно використовувати три *dummy*-змінні. За наявності сезонного впливу він буде відображений оцінками диференційованого перетину α_1 , α_2 та α_3 , якщо вони будуть статистично значимі. Можливий випадок статистичної значимості лише декількох диференційованих перетинів.

Використовуючи дані з таблиці, отримуємо такі результати (млн. грн.):

$$y_t = 6688.3789 + 1322.8938 D_{1t} - 217.8037 D_{2t} + 183.8597 D_{3t} + 0.038 x_t.$$

σ_{at}	(1711.377)	(638.47)	(632.25)	(654.29)	(0.0115)
t	(3.9082)	(2.0720)	(-0.3445)	(0.2810)	(3.3313)

$$R^2 = 0.5255. \quad (5.5.42)$$

Результати свідчать, що лише коефіцієнт продаж і диференційний перетин другого кварталу статистично значимі з рівнем помилки 5%. Звідси випливає, що сезонний фактор присутній у другому кварталі щороку. Коефіцієнт продажу 0.0383 показує, що після врахування впливу сезонних коливань збільшення продаж на 1 млн.грн. приведе до підвищення прибутків на 0.04 млн.грн. Середній рівень прибутків у базовому першому кварталі становив 6688 млн. грн., а в другому підвищився на 1323 млн. грн., тобто дорівнював 8011 млн.грн.

Таблиця 5.11

Динаміка прибутків приватних фірм України *

Рік, квартал	Прибуток, млн.грн.	Продаж, млн.грн.	D_1	D_2	D_3
1992 - I	10.53	114.9	0	0	0
- II	12.09	124.0	1	0	0
- III	10.84	121.46	0	1	0
- IV	12.20	131.92	0	0	1
1993 - I	12.25	129.91	0	0	0
- II	14.00	140.98	1	0	0
- III	12.21	137.83	0	1	0
- IV	12.82	145.47	0	0	1
1994 - I	11.34	136.99	0	0	0
- II	12.61	145.13	1	0	0
- III	11.01	141.54	0	1	0
- IV	12.73	151.78	0	0	1
1995 - I	12.54	148.86	0	0	0
- II	14.85	158.91	1	0	0
- III	13.20	155.72	0	1	0
- IV	14.95	168.41	0	0	1
1996 - I	14.15	162.78	0	0	0
- II	15.95	176.06	1	0	0
- III	14.02	172.42	0	1	0
- IV	14.31	183.32	0	0	1

* Дані умовні

Оскільки тільки оцінки другого кварталу статистично значимі, то можна модифікувати (5.5.41), використовуючи лише одну *dummy*-змінну, для того, щоб його виділити:

$$\hat{y}_i = 6515.6 + 1331.4D_1 + 0.0393x_i \quad (5.5.43)$$

(1623.1) (493.02) (0.0106)

$t = (4.0143) (2.7004) (3.7173)$

$R^2=0.5155,$

де $D_1 = 1$ для спостережень у другому кварталі. У всіх інших випадках $D_1 = 0$.

Зрозуміло, що (5.5.43) — обмежена версія (5.5.41); обмеження для першого, третього і четвертого кварталів однакові. Судячи з результатів (5.5.41), можна сподіватися, що ці обмеження дійсні. Тому висновок залишається таким самим: сезонні коливання мають найбільший вплив тільки у другому кварталі.

У формуванні моделі (5.5.41) припускалося, що лише перетин відрізняється між кварталами, нахил змінної продаж залишається таким самим. Однак, це припущення можна перевірити.

5.5.9. Основні висновки щодо моделей з *dummy*-змінними

Завдання цього розділу полягало в тому, щоб показати, як якісні *dummy*-змінні, що набувають значення 1 і 0, можуть бути представлені в моделях лінійних регресій. *Dummy*-змінні класифікують таку інформацію, яка поділяє модель на кілька підгруп (категорій), що базуються на якісних властивостях (стать, сімейний стан, освіта, релігія тощо), і окремо працюють з кожною підгрупою. Якщо в різних підгрупах є відмінності у варіаціях кількісних змінних, вони можуть позначитись на коефіцієнтах перетинів і нахилів у регресіях різних підгруп.

Треба бути уважним, застосовуючи техніку *dummy*-змінних. По-перше, якщо модель регресії містить константу, то кількість *dummy*-змінних має бути на одиницю меншою за кількість категорій кожної якісної змінної. По-друге, коефіцієнти *dummy*-змінних завжди мають інтерпретуватись за відношенням до базової групи, яка набуває значення нуль. Нарешті, якщо модель має кілька якісних змінних з кількома категоріями, введення *dummy*-змінних може призвести до великої кількості ступенів вільності. Тому завжди потрібно зважувати кількість *dummy*-змінних для введення у модель із загальною кількістю спостережень.

ПІДСУМУЄМО ВИВЧЕНЕ

Після вивчення цього розділу ви зможете

1. Виявити порушення основних припущень лінійної регресійної моделі, а також розуміти проблеми, які виникають, коли ці припущення порушуються.
2. Модифікувати первісну регресійну модель для того, щоб позбавитись порушення основних припущень.
3. Вміти оцінити невідомі параметри модифікованих регресійних моделей.
4. Використовувати залежні змінні з урахуванням лагу в регресійній моделі.
5. Використовувати *dummy*-змінні для врахування впливу якісних незалежних змінних.
6. Визначати залежність, за якої дві з незалежних змінних взаємодіють між собою.

Короткий огляд розділу

1. До цього ми вивчили, що прості та багатофакторні лінійні регресійні моделі базуються на деяких припущеннях. Коли ці припущення порушуються, можуть виникати проблеми, пов'язані зі статистичними висновками та оцінюванням невідомих параметрів регресійної моделі.

2. Раніше ми ознайомили вас із концепцією мультиколінеарності. Мультиколінеарність наявна, коли незалежні змінні (фактори) тісно пов'язані між собою. В такому разі неможливо розподілити їхній вплив на залежну змінну.

Іншими словами, мультиколінеарність — ситуація, коли дві чи більше незалежних змінних сильно корелюють між собою.

а) якщо незалежні змінні ідеально корелюють, метод найменших квадратів не може використовуватися для оцінювання параметрів регресії (випадок досконалої мультиколінеарності);

б) якщо незалежні змінні сильно, але не ідеально корелюють, збільшиться середнє квадратичне відхилення параметрів регресії і відповідно до цього значення *t*-статистики буде дуже малим;

в) виявлення мультиколінеарності може бути досить важким. Хоча є два простих методи виявлення мультиколінеарності.

По-перше, ми можемо розрахувати коефіцієнт кореляції між кожною парою незалежних змінних. Якщо якась із пар має коефіцієнт кореляції більший, ніж 0,8, то мультиколінеарність можлива.

По-друге, ми можемо розглянути значення *t*-статистики для коефіцієнтів регресії та значення *F*-критерію регресійної моделі. Якщо *t*-значення є незначимими в той час як *F*-значення є значимими, мультиколінеарність також можлива.

3. Другою проблемою може стати гетероскедастичність. Гетероскедастичність має місце, коли дисперсія випадкової величини не є константою. Наприклад, якщо б ми будували регресію, яка пояснює індивідуальне споживання в його залежності від доходу споживача, було б нормальним припустити, що випадкові величини не будуть мати постійної дисперсії, тому що заможніші люди мають більшу дисперсію в їхньому споживанні.

За наявності гетероскедастичності оцінки параметрів можуть бути без відхилень, залишаючись у той же час неефективними. Гетероскедастичність є загальною проблемою в галузевих моделях регресії.

Одним із методів визначення гетероскедастичності є побудова графіка залежності помилок від незалежних змінних. Якщо не зберігається постійне відношення, наприклад, дисперсія, або розкид значень, помилок стає більшим в міру того, як зростає незалежна змінна, гетероскедастичність може мати місце.

Другий метод виявлення гетероскедастичності: необхідно побудувати регресію, використовуючи квадрати помилок як залежну змінну від незалежних змінних. Гетероскедастичність наявна, якщо хоч одне із значень *t*-статистики для оцінених параметрів такої моделі значиме.

Коли наявна гетероскедастичність, оцінити невідомі параметри можна за допомогою методу зважених найменших квадратів.

4. Автокореляція має місце, коли значення випадкової величини корелюють із своїми попередніми значеннями. При наявності автокореляції оцінки параметрів будуть без відхилень, але неефективними.

Автокореляція може бути виявлена за допомогою використання тесту Дарбіна — Уотсона (DW). Проблема, пов'язана з використанням цього тесту, полягає в тому, що результат може бути невизначеним. Для тестування за критерієм Дарбіна — Уотсона використовуємо значення верхньої d_U та нижньої d_L DW-величини і дотримуємося таких вказівок:

а) однозначне тестування гіпотези. H_0 — автокореляції немає, H_1 — наявна позитивна автокореляція. Ми відкидаємо H_0 , якщо $d < d_L$, і приймаємо H_0 , якщо $d > d_U$. Тест буде безрезультатним, якщо $d_L \leq d \leq d_U$;

б) однозначне тестування гіпотези. H_0 — автокореляції немає, H_1 — наявна негативна автокореляція. Ми відкидаємо H_0 , якщо $d > 4 - d_L$, і приймаємо H_0 , якщо $d < 4 - d_U$. Тест буде безрезультатним, якщо $4 - d_L \leq d \leq 4 - d_U$;

в) двозначне тестування гіпотези. H_0 — автокореляції немає, H_1 — наявна негативна або позитивна автокореляція. Ми відкидаємо H_0 , якщо $d < d_L$ або $d > 4 - d_U$, і приймаємо H_0 , якщо $d_U < d < 4 - d_U$. Тест буде безрезультатним, якщо $d_L \leq d \leq d_U$ або $4 - d_L \leq d \leq 4 - d_U$.

Коли до регресійної моделі включена залежна змінна з лагом (тобто у випадку динамічної моделі), тест Дарбіна — Уотсона не буде дійсним. У цьому разі необхідно користуватися іншим тестом, відомим під назвою *h*-тест Дарбіна.

5. Помилки специфікації трапляються, коли регресійна модель неправильно специфікована. Це може статися, коли не включена доречна змінна, або включена недоречна.

6. Деколи між x та y є нелінійний зв'язок. У цьому разі нелінійна модель може бути кращим вибором при побудові регресії. Найпростішим видом нелінійної моделі є квадратична модель, яка включає значення незалежної змінної, зведеної до квадрата, як незалежну змінну для того, щоб відобразити нелінійне співвідношення.

7. Коли ми будемо регресію на часовому проміжку, можемо вважати, що теперішні значення y залежать від попередніх значень, таких як y_{t-1} або y_{t-2} . У цьому разі можна використати лагову залежну змінну як незалежну змінну для відображення цього ефекту.

8. У цьому розділі також розглядалась проблема поводження із змінними, які є якісними за природою. Наприклад, якщо економіст вважає, що одним із факторів, які впливають на заробітну плату індивіда, є його стать, він може включити в модель бінарну змінну, яка називається атрибутивною змінною (dummy-змінною), і таким чином відобразити вплив даного фактора.

9. Коли припускається, що дві незалежні змінні визначають значення залежної змінної, наприклад при оцінюванні впливу опадів та добрив на урожайність, для відображення цього ефекту може застосовуватися змінна взаємодії. Вона може бути представлена у вигляді добутку двох незалежних змінних та введена в регресію як додаткова незалежна змінна.

Основні формули

Статистика Дарбіна — Уотсона:

$$DW = \frac{\sum_{i=1}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2};$$

h -статистика Дарбіна:

$$h = \left[1 - \frac{d}{2} \right] \sqrt{\frac{n}{1 - n \operatorname{var}(b_1)}}.$$

Квадратична регресійна модель:

$$y = b_0 + b_1 x_1 + b_2 x_1^2 + e.$$

Модель, що включає лаг залежної змінної:

$$y_t = b_0 + b_1 x_{1t} + b_2 x_{2t} + \dots + b_p x_{pt} + \lambda y_{t-1} + e_t.$$

Модель із dummy-змінною:

$$y_t = a_0 + a_1 D_{1t} + \beta_1 x_{1t} + e_t.$$

Модель із змінною взаємодії:

$$y_t = b_0 + b_1 x_{1t} + b_2 x_{2t} + b_3 (x_{1t} \times x_{2t}) + e_t.$$

ВПРАВИ І КОМЕНТАРІ

Вправа 5.1. Статистика Дарбіна — Уотсона

Припустимо, є вибірки з 25 спостережень для двох незалежних змінних та залежної змінної. За цими даними ви побудували двофакторну модель і хочете перевірити її на автокореляцію. Що можна сказати про автокореляцію для кожної з таких статистик Дарбіна — Уотсона:

- $d=1.20$;
- $d=2.00$;
- $d=3.25$;
- $d=1.55$;
- $d=2.55$.

Розв'язок. Для того, щоб визначити, чи є автокореляція, ми повинні проаналізувати таблицю значень Дарбіна — Уотсона. Згадайте, що статистика Дарбіна — Уотсона містить верхнє та нижнє значення і проміжок, де результат тесту є невизначеним. Для 25 спостережень, 2 факторів та 5%-ного рівня значимості за таблицею $d_L = 1.21$, $d_U = 1.54$.

Відповідь:

- негативна автокореляція, тому що $d < 1.21$;
- автокореляції немає, тому що $1.21 < d < 4 - 1.54$;
- негативна автокореляція, тому що $d > 4 - 1.21$;
- автокореляції немає, тому що $1.21 < d < 4 - 1.54$;
- невизначеність, тому що $4 - 1.54 \leq d \leq 4 - 1.21$.

Вправа 5.2. Мультиколінеарність

Вас цікавить відношення між y та трьома можливими незалежними змінними x_1 , x_2 та x_3 . Кореляційну матрицю наведено у табл. 5.12.

Таблиця 5.12

	y	x_1	x_2	x_3
y	1.00	0.55	0.66	0.89
x_1		1.00	0.82	0.75
x_2			1.00	0.65
x_3				1.00

Перевірте можливість наявності мультиколінеарності. Якщо вона є, то між якими змінними?

Розв'язок. У таблиці наявна кореляція між кожною парою змінних. Мультиколінеарність може бути, тому що кореляція між x_1 та x_2 дорівнює 0.82.

Вправа 5.3. Атрибутивна змінна (дитту-змінна)

Припустимо, вас найняв адвокат, який хоче показати, що компанія дискримінує жінок за розміром заробітної плати. Ви складаєте таку регресію:

$$y_i = 19000 + 2000x_{1i} + 1000x_{2i} + 4000D_{1i}$$

(821) (332) (3,400),

де y_i — заробітна плата особи;

x_{1i} — кількість років досвіду;

x_{2i} — кількість років освіти;

D_i — атрибутивна змінна =1 для жінок, =2 для чоловіків.

Стандартні помилки параметрів наводяться в дужках.

1. Проінтерпретуйте значення параметрів при змінних досвіду та освіти.

2. Проінтерпретуйте параметр для статі. Чи є дискримінація?

Розв'язок

1. Коефіцієнти для досвіду та освіти є додатними та значимими (їхні t -значення більше 2), вони показують, що працівники з більшою освітою та/або досвідом отримують більшу заробітну плату.

2. Досліджуючи дискримінацію, розглянемо значимість параметра для атрибутивної змінної, яка відображає стать. Позитивний та значимий коефіцієнт означатиме, що жінки заробляють більше, ніж чоловіки з таким же досвідом та освітою. Негативний та значимий коефіцієнт означатиме, що жінки заробляють менше, ніж чоловіки з таким же досвідом та освітою. Оскільки коефіцієнт для статі незначимий, не можна зробити висновок, що дискримінація наявна.

Вправа 5.4. Розрахунок статистики Дарбіна — Уотсона

Значення помилок регресійної моделі дано в табл. 5.13.

Розрахуйте значення статистики Дарбіна — Уотсона. Чи має тут місце автокореляція?

Таблиця 5.13

Період	Значення e	Період	Значення e
1	21.2	9	-9.6
2	-15.3	10	10.2
3	12.4	11	-13.3
4	-21.0	12	7.4
5	9.4	13	-15.1
6	-25.5	14	12.0
7	-8.4	15	-6.0
8	12.3	16	8.0

Розв'язок. За таблицею знаходимо для $n = 16$ та $\alpha = 0.05$, $d_e = 1.10$ та $d_u = 1.37$. Тобто наявна негативна автокореляція, тому що $d > 4 - d_e$.

Вправа 5.5. Змінні взаємодії

Біолог досліджує вплив температури та вологості на ріст клітин. Отриману інформацію про 6 посівів наведено у табл. 5.14.

Таблиця 5.14

Посів	Температура, °C x_1	Вологість, % x_2	Ріст клітин y
1	4	8	10.000
2	8	12	11.122
3	12	19	14.025
4	16	30	19.022
5	20	50	26.872
6	24	75	42.308

Змоделюйте відношення між кількістю клітин, температурою та вологістю. Використайте змінну взаємодії для оцінки спільного впливу температури та вологості на ріст клітин. Проінтерпретуйте ваші результати.

Розв'язок. Для оцінки загального впливу температури та вологості на ріст клітин ми створюємо третю незалежну змінну за допомогою добутку значень температури та вологості. Коефіцієнт цієї величини буде відображати спільний ефект. Регресія, яку ми хочемо побудувати, має такий вигляд:

$$y_i = b_0 + b_1x_{1i} + b_2x_{2i} + b_3(x_{1i} \times x_{2i}).$$

Створивши змінну взаємодії, ми оцінюємо невідомі параметри побудованої моделі та отримуємо такі результати:

$$y_i = 10916.78 + 447.48x_{1i} - 502.65x_{2i} + 32.39(x_{3i})$$

(268.86) (321.04) (9.53)

$$R^2 = 0.999.$$

Значення середньоквадратичних помилок параметрів (стандартних помилок) наведені в дужках.

Вправа 5.6. Гетероскедастичність

Нехай нас цікавить зв'язок між дивідендом (y) та прибутком (x) на одну акцію різних компаній. Ми збираємо інформацію за двома цими змінними та оцінюємо таку регресію:

$$y_i = b_0 + b_1x_i + e_i.$$

Виходячи з цієї моделі, розрахуємо значення помилок для кожної компанії. Дані наведено в табл. 5.15.

Таблиця 5.15

Компанія	x , дол.	e	Компанія	x , дол.	e
1	0.80	-0.05	7	2.20	-2.12
2	1.10	0.35	8	2.24	2.50
3	1.20	-0.10	9	2.75	-3.00
4	1.30	0.72	10	3.00	2.80
5	1.45	-0.25	11	3.10	-3.50
6	2.05	1.21	12	3.40	4.00

Чи є в даному випадку гетероскедастичність?

Розв'язок. Гетероскедастичність наявна, тому що значення помилок стають більшими при зростанні незалежної змінної.

ТЕСТИ

Вибрати правильну відповідь на запитання

1. Мультиколінеарність наявна, коли:

- а) дві чи більше незалежних змінних мають високу кореляцію;
- б) дисперсія випадкових величин не постійна;
- в) теперішні та лагові значення помилок корелюють;
- г) незалежна змінна виміряна з помилкою;
- д) ми будемо неправильно версію істинної моделі.

2. Гетероскедастичність наявна, коли:

- а) дві чи більше незалежних змінних мають високу кореляцію;
- б) дисперсія випадкової величини не постійна;
- в) теперішні та лагові значення помилок корелюють;
- г) незалежна змінна виміряна з помилкою;
- д) ми будемо неправильно версію істинної моделі.

3. Автокореляція наявна, коли:

- а) дві чи більше незалежних змінних мають високу кореляцію;
- б) дисперсія випадкової величини не постійна;
- в) теперішні та лагові значення випадкової величини корелюють;
- г) незалежна змінна виміряна з помилкою;
- д) ми будемо неправильно версію істинної моделі.

4. Помилка в специфікації наявна, коли:

- а) незалежна змінна виміряна з помилкою;
- б) ми будемо неправильно версію істинної моделі;
- в) дві чи більше незалежних змінних мають високу кореляцію;
- г) дисперсія значень помилки не постійна;
- д) дві чи більше залежних змінних мають високу кореляцію.

5. Гетероскедастичність дає нам:

- а) оцінки параметрів з відхиленням;
- б) найкращі лінійні оцінки (BLUE);
- в) ефективні оцінки параметрів;
- г) проблеми із статистичними висновками;
- д) високий ступінь кореляції між залишками та залежною змінною.

6. Автокореляція дає нам:

- а) оцінки параметрів з відхиленням;
- б) найкращі лінійні оцінки (BLUE);
- в) неефективні оцінки параметрів;
- г) проблеми із статистичними висновками;
- д) високий ступінь кореляції між залишками та незалежною змінною.

7. Мультиколінеарність дає нам:

- а) оцінки параметрів з відхиленням;
- б) найкращі лінійні оцінки (BLUE);
- в) неефективні оцінки параметрів;
- г) проблеми із статистичними висновками;
- д) два залишки, які корелюють один з одним.

8. Помилка специфікації дає нам:

- а) оцінки параметрів з відхиленням;
- б) найкращі лінійні оцінки (BLUE);
- в) неефективні оцінки параметрів;
- г) мультиколінеарність;
- д) автокореляцію.

9. Дитту-змінна може вживатися, коли:

- а) незалежна змінна кількісна;
- б) незалежна змінна якісна;
- в) незалежна змінна гетероскедастична;
- г) немає автокореляції;
- д) проблемою є мультиколінеарність.

10. Якщо ми зацікавлені у використанні дитту-змінних для відображення ефекту різних місяців на віддачу від цінних паперів, ми повинні використати:

- а) 12 дитту-змінних (атрибутивних змінних), одну для кожного місяця року;
- б) 1 атрибутивну змінну;
- в) 11 атрибутивних змінних;
- г) стільки атрибутивних змінних, скільки ми захочемо;
- д) не можна використовувати атрибутивну змінну для відображення цього ефекту.

11. Якщо ми будемо взаємозв'язок, котрий має U-подібний вигляд, (такий, як крива загальних витрат), найкраще відобразити його за допомогою:

- атрибутивної змінної;
- простой регресії;
- залежної змінної з лагом;
- квадратичної регресійної моделі;
- такий зв'язок не може бути відображений за допомогою регресійного аналізу.

12. Для виправлення проблеми гетероскедастичності ми можемо:

- відкинути одну чи більше незалежних змінних;
- використати перехід до логарифмів (*log transformation*);
- використати атрибутивні змінні;
- використати метод зважених найменших квадратів;
- спершу виправити проблему автокореляції.

13. Для виправлення проблеми мультиколінеарності можна:

- використати перехід до логарифмів (*log transformation*);
- відкинути одну чи більше незалежних змінних;
- використати атрибутивні змінні;
- використати метод зважених найменших квадратів;
- використати залежну змінну з лагом.

14. Використання атрибутивних змінних для вимірювання різниці в оплаті праці чоловіків та жінок означає, що:

- нахил в моделі для чоловіків та жінок буде однаковим;
- перетин в моделі для чоловіків та жінок буде однаковим;
- нахил та перетин будуть однаковими;
- нахил та перетин будуть різними;
- наявна мультиколінеарність.

ВІДПОВІСТІ НА ЗАПИТАННЯ: ТАК/НІ; ЯКЩО НІ, ПОЯСНІТЬ ЧОМУ

- Гетероскедастичність призводить до оцінок параметрів з відхиленням.
- Автокореляція призводить до оцінок параметрів з відхиленням.
- Гетероскедастичність може бути визначена за допомогою тесту Дарбіна — Уотсона.
- Коли наявна досконала мультиколінеарність, неможливо застосовувати метод найменших квадратів.
- Тест Дарбіна — Уотсона може використовуватися для визначення наявності автокореляції, коли в регресію включена залежна змінна з лагом.
- Атрибутивні змінні можуть використовуватися, коли в регресію включаються такі якісні змінні (*dummy-змінні*), як, наприклад, стать чи освіта.
- Мультиколінеарність наявна тоді, коли незалежна змінна сильно корелює із залежною змінною.
- Невключення в модель доречної незалежної змінної призводить до помилки специфікації.
- Включення недоречної пояснювальної змінної в регресію призводить до помилки специфікації.
- Використання незалежних змінних, вимірюваних з помилкою, веде до помилки специфікації.
- Включення недоречної пояснювальної змінної в регресію гірше, ніж невключення доречної пояснювальної змінної.
- Проблема гетероскедастичності може бути інколи виявлена за допомогою визначення кореляції між пояснювальними змінними.
- Тест Дарбіна — Уотсона завжди засвідчить наявність автокореляції.
- Змінна взаємодії може бути корисною, коли вплив незалежної змінної на залежну змінну зумовлюється іншою незалежною змінною.
- Один із шляхів визначення гетероскедастичності — аналіз графіка, де на осях відкладені залишки та незалежні змінні або прогнозовані значення.

ЗАПИТАННЯ І ЗАДАЧІ

- Знайдіть d_U та d_L для регресії, яка має k пояснювальних змінних та n спостережень:
 - $k = 1, n = 25$;
 - $k = 2, n = 30$;
 - $k = 3, n = 20$.
- Припустимо, ви моделюєте регресію, використовуючи три незалежні змінні (фактори) та 40 спостережень. Якщо обчислене вами значення статистики Дарбіна — Уотсона дорівнює 1.25, чи наявна автокореляція?

3. Припустимо, ви оцінюєте регресію, використовуючи одну незалежну змінну та 50 спостережень, і встановили, що

$$\sum_{i=2}^{50} (e_i - e_{i-1})^2 = 6.27, \quad \sum_{i=1}^{50} e_i^2 = 2.71.$$

А. Обчисліть статистику Дарбіна — Уотсона.

Б. Чи наявна автокореляція?

4. У табл. 5.16 наведено кореляційну матрицю для регресійної моделі з двома незалежними змінними.

Таблиця 5.16

	y	x_1	x_2
y	1.00	0.98	0.75
x_1		1.00	0.85
x_2			1.00

Чи наявна мультиколінеарність?

5. Які проблеми можуть виникнути у наступних регресійних моделях?

А. Регресійна модель, яка моделює залежність віддачі акцій IBM від нарахування на одну акцію (EPS), за останні 10 років.

Б. Регресійна модель, яка моделює залежність EPS 100 різних компаній від їхнього розміру для 1996 року.

В. Регресійна модель, яка моделює залежність рівня споживання для 5000 чоловік від рівня їхнього індивідуального доходу в 1996 році.

6. Припустимо, ви оцінюєте рівняння з обчислення заробітку залежно від досвіду, освіти та статі індивіда. Воно має такий вигляд:

$$y_i = 12212 + 722x_{1i} + 927x_{2i} + 2211D_i,$$

де $D_i = 1$ для жінок та $= 0$ для чоловіків.

Проінтерпретуйте коефіцієнт біля атрибутивної змінної для статі D_i .

7. Припустимо, що ви, працюючи у вузі, зацікавилися співвідношенням між балами, які студенти отримують на вступному курсі математики, та їхніми оцінками з математики у школі. Якщо ви підозрюєте, що бали з математики можуть різнитися у студентів різної статі незалежно від їхніх оцінок у школі, поясніть, як ви збираєтеся включити таку змінну в модель, яку ви будете?

8. Ви будете регресію для вимірювання взаємозв'язку між бушелями пшениці, рівнем опадів та кількістю використаних добрив.

А. Якщо ви вважаєте, що опади та добрива діють разом, сформулюйте відповідну регресійну модель.

Б. Проінтерпретуйте значення коефіцієнтів регресії в моделі, яку ви запропонували.

ВІДПОВІДІ

Тести

- | | |
|-------|--------|
| 1) а; | 8) а; |
| 2) б; | 9) б; |
| 3) в; | 10) в; |
| 4) б; | 11) г; |
| 5) г; | 12) г; |
| 6) в; | 13) а; |
| 7) в; | 14) а. |

Так/Ні

- Ні. Гетероскедастичність призводить до неефективних оцінок.
- Ні. Автокореляція призводить до неефективних оцінок.
- Ні. Автокореляція визначається за допомогою тесту Дарбіна — Уотсона.
- Так.
- Ні. Коли в регресії використовується лагова залежна змінна, необхідно застосовувати h -тест Дарбіна.
- Так.
- Ні. Мультиколінеарність виникає, коли незалежні змінні корелюють між собою.
- Так.
- Ні. t -тест має визначати недоречність змінної.
- Так.
- Ні. Гірше не включити доречну змінну, аніж включити недоречну.
- Ні. Гетероскедастичність можна виявити, розглянувши графік залишків з регресії.
- Ні. Тест Дарбіна — Уотсона має інтервал, на якому результат невизначений.
- Так.
- Так.

Запитання і проблеми

1. а) $d_L = 1.29, d_U = 1.45$;
 б) $d_L = 1.28, d_U = 1.57$;
 в) $d_L = 1.00, d_U = 1.68$;

2. $d_L = 1.25, d_U = 1.68$.

$0 < d < d_L$, отже, ми відкидаємо гіпотезу про відсутність автокореляції та приймаємо гіпотезу про позитивну автокореляцію.

3. А. $d = 6.27/2.71 = 2.31$.

Б. $d_L = 1.50, d_U = 1.59$.

$2 < d < 4 - d_U$, отже, ми приймаємо нульову гіпотезу про відсутність автокореляції.

4. Кореляція між x_1 та x_2 дорівнює 0.85, отже, можлива мультиколінеарність.

5. А. У таких динамічних рядах можлива наявна автокореляція.

Б. У таких регресіях, як ця, гетероскедастичність може стати проблемою, тому що можна очікувати, що більші компанії матимуть більшу дисперсію у EPS.

В. Гетероскедастичність може стати проблемою, оскільки особи з вищим рівнем доходу скоріше за все матимуть більшу дисперсію в своєму споживанні.

6. Коефіцієнт біля атрибутивної змінної вимірює різницю в заробітках жінок та чоловіків, припускаючи, що вплив досвіду та рівня освіти на заробіток є однаковим. У зв'язку з тим, що коефіцієнт дорівнює -2211, очікується, що жінки будуть отримувати на 2211 гривні менше, ніж чоловіки з таким же досвідом та рівнем освіти.

7. Для вимірювання зв'язку між оцінками з математики у школі та балами з математики у вузі ми можемо застосувати регресійний аналіз, вводючи dummy-змінну, а саме:

$$y_i = b_0 + b_1x_i + b_2D_i,$$

де $D = 1$ для студента та $= 0$ для студентки. Коефіцієнт b_2 буде вимірювати різницю між балами для чоловіків та жінок безвідносно до їхніх оцінок з математики у школі. Позитивний та значимий коефіцієнт b_2 означав би, що чоловіки отримують вищі бали незалежно від оцінок у школі. Негативний та значимий коефіцієнт b_2 означав би, що жінки отримують вищі бали незалежно від оцінок у школі.

8. А. Будуємо модель виду:

$$y_t = b_0 + b_1x_{1t} + b_2x_{2t} + b_3(x_{1t} \times x_{2t}) + e_t,$$

де y — обсяги виробництва пшениці;

x_1 — кількість опадів; x_2 — кількість добрив.

Б. Зміна у виробництві пшениці, обумовлена зміною кількості опадів на одиницю, дорівнює $b_1 + b_3$. Зміна у виробництві пшениці в результаті зміни кількості добрив на одиницю дорівнює $b_2 + b_3$.

ВИКОНАЙТЕ САМОСТІЙНО

Вправи. Мультиколінеарність

Вправа 5.1.1

Припустимо, що моделі

$$y_i = \beta_0 + \beta_1x_{1i} + \beta_2x_{2i} + \varepsilon_i$$

задовольняють дані (див. табл. 5.17).

А. Чи можна оцінити значення трьох невідомих параметрів? Якщо ні, то чому?

Б. Якщо ні, то якою буде лінійна комбінація з цих параметрів? Зробіть потрібні розрахунки.

Вправа 5.1.2

Якщо справджується

$$\lambda_1x_{1i} + \lambda_2x_{2i} + \lambda_3x_{3i} = 0$$

для всіх λ_i , визначте $r_{12.3}, r_{13.2}, r_{23.1}$. Який ступінь кореляції буде в цьому випадку?

Примітка. $R_{1.23}^2$ — коефіцієнт детермінації в регресії x за x_2 і x_3 .

Вправа 5.1.3. Дано модель

$$y_t = \beta_0 + \beta_1x_t + \beta_2x_{t-1} + \beta_3x_{t-2} + \beta_4x_{t-3} + \beta_5x_{t-4} + \varepsilon_t,$$

де y — споживання; x — дохід; t — час. За цією моделлю споживчі витрати в час t є функцією доходу не лише в час t , а й доходу за попередні періоди. Така модель називається **лаговою моделлю розподілу**.

А. Чи очікуєте ви у цьому випадку виникнення проблеми мультиколінеарності? Чому?

Б. Якщо мультиколінеарність наявна, як ви вирішите цю проблему?

Таблиця 5.17

y	x_1	x_2
-10	1	1
-8	2	3
-6	3	5
-4	4	7
-2	5	9
0	6	11
2	7	13
4	8	15
6	9	17
8	10	19
10	11	21

Вправа 5.1.4. Чому серед таких економічних даних, як ВВП, пропозиція грошей, доход, безробіття та інші, очікується мультиколінеарність?

Вправа 5.1.5

Дано модель

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i,$$

де r_{12} — коефіцієнт кореляції між x_1 і x_2 дорівнює нулю. Запропоновано записати такі регресійні моделі:

$$y_i = \alpha_0 + \alpha_1 x_{1i} + \varepsilon_{1i};$$

$$y_i = \gamma_0 + \gamma_2 x_{2i} + \varepsilon_{2i}.$$

- А. Чи буде $\hat{\alpha}_1 = b_1$ і $\hat{\gamma}_2 = b_2$? Чому?
 Б. Чи буде b_0 дорівнювати $\hat{\alpha}_0$ або $\hat{\gamma}_0$, чи комбінаціям їх?
 В. Чи буде $\text{var}(b_1) = \text{var}(\hat{\alpha}_1)$ і $\text{var}(b_2) = \text{var}(\hat{\gamma}_2)$?

Вправа 5.1.6. У прикладі з глави 4 ми застосували функцію Кобба — Дугласа для аналізу стану деревообробного сектора України. Отримані результати свідчать, що фактори праці та капіталу є однаково важливими.

А. Чи є змінні праці та капіталу висококорельованими?

Б. Якщо ви відповіли на (а) ствердно, вилучіть змінну праці з моделі і побудуйте регресію випуску щодо капіталу.

Вправа 5.1.7. Покрокова регресія. Визначаючи, яка множина факторів є найкращою для моделі, дослідники часто вдаються до покрокової регресії. У цьому методі змінні x вводяться на кожному кроці (пряма покрокова регресія) або в модель вводяться одразу всі можливі x , y і відкидаються по одній на кожному кроці (зворотна покрокова регресія). Рішення, які змінні вводити або викидати, залежить від того, як вони впливають на SSE , що визначається за допомогою F -тесту. Чи запропонуєте ви цю процедуру, спираючись на отримані знання про мультиколінеарність? Обґрунтуйте вашу відповідь.

Вправа 5.1.8. Клейн та Голберг намагалися застосувати до економіки США таку модель:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i,$$

де y — споживання; x_1 — зарплата; x_2 — доход, крім зарплати; x_3 — доход від сільського господарства. Оскільки очікується, що x_1 , x_2 , x_3 колінеарні, вони отримали β_2 , β_3 , використовуючи міжгалузевий аналіз: $\beta_2 = 0.75\beta_1$, $\beta_3 = 0.625\beta_1$. На основі цих даних отримали функцію

$$y_i = \beta_0 + \beta_1(x_{1i} + 0.75x_{2i} + 0.625x_{3i}) + \varepsilon_i = \beta_0 + \beta_1 z_i + \varepsilon_i,$$

де $z_i = x_{1i} + 0.75x_{2i} + 0.625x_{3i}$.

Таблиця 5.18

Рік	y	x_1	x_2	x_3
1936	62.8	43.41	17.10	3.96
1937	65.0	46.44	18.65	5.48
1938	63.9	44.35	17.09	4.37
1939	67.5	47.82	19.28	4.51
1940	71.3	51.02	23.24	4.88
1941	76.6	58.71	28.11	6.37
1945	86.3	87.69	30.29	8.96
1946	95.7	76.73	28.26	9.76
1947	98.3	75.91	27.91	9.31
1948	100.3	77.62	32.30	9.85
1949	103.2	78.01	31.39	7.21
1950	108.9	83.57	35.61	7.39
1951	108.5	90.59	37.58	7.98
1952	111.4	95.47	35.17	7.42

- А. Застосуйте цю модель до даних, наведених у табл. 5.18 і обчисліть β_2 і β_3 .
 Б. Як ви проінтерпретуєте змінну Z ?

Вправа 5.1.9. У матричному записі з розділу 4 маємо:

$$b = (X'X)^{-1}X'y.$$

- А. Як вплине на значення b досконала кореляція серед x ?
 Б. Як можна виявити наявність досконалої кореляції?

Вправа 5.1.10. Використаємо матричний запис для дисперсійно-варіаційної матриці:

$$\text{var-cov}(b) = \sigma^2(X'X)^{-1}.$$

Що трапиться з var-cov-матрицею, якщо:

- а) наявна досконала кореляція;
 б) колінеарність висока, але не досконала?

Вправа 5.1.11. Дано таку кореляційну матрицю:

$$R = \begin{matrix} & x_2 & x_3 & \dots & x_k \\ \begin{matrix} x_2 \\ x_3 \\ \dots \\ x_k \end{matrix} & \begin{bmatrix} 1 & r_{23} & \dots & r_{2k} \\ r_{32} & 1 & \dots & r_{3k} \\ \dots & \dots & \dots & \dots \\ r_{k2} & r_{k3} & \dots & 1 \end{bmatrix} \end{matrix}.$$

Як визначити з цієї матриці, чи:

- наявна досконала колінеарність?
- наявна колінеарність нижча, ніж досконала?
- x некорельовані?

Примітка. Використайте $|R|$ для відповіді на запитання, де $|R|$ — визначник R .

Вправа 5.1.12. Ортогональні пояснювальні змінні. Нехай у моделі

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i$$

$x_2 \dots x_k$ — некорельовані. Такі змінні називаються ортогональними змінними. В цьому випадку:

- яку структуру матиме матриця $(X'X)$?
- як ви отримаєте $b = (X'X)^{-1}X'y$?
- яку природу матиме cov-var матриця b ?
- ви побудували регресію і хочете ввести нову змінну x_{k+1} . Чи потрібно переобчислювати коефіцієнти $b_1 \dots b_k$? Так чи ні, поясніть.

Вправа 5.1.13. Нехай дано модель

$$y_t = \beta_0 + \beta_1 M_t + \beta_2 M_{t-1} + \beta_3 (M_t - M_{t-1}) + \varepsilon_t,$$

де y_t — ВВП в час t ; M_t — пропозиція грошей в час t ; M_{t-1} — пропозиції в час $(t-1)$; $(M_t - M_{t-1})$ — зміна грошової пропозиції між моментом t і $t-1$. Ця модель стверджує, що ВВП залежить від грошової пропозиції в час t і $t-1$ та її зміни за цей період.

А. Припустимо, що є необхідні дані для цієї моделі. Чи зможете ви обчислити всі коефіцієнти? Поясніть.

Б. Якщо ні, то які коефіцієнти можна визначити?

В. Припустимо, значення $\beta_2 M_{t-1}$ у моделі немає. Чи буде така сама відповідь для (А)?

Г. Дайте відповідь на запитання (В) ще раз, припускаючи відсутність $\beta_1 M_t$.

Вправа 5.1.14. У табл. 5.19 подано дані щодо імпорту, ВВП, індексу споживчих цін у США за 1970—1983 рр.

Нехай дано таку модель:

$$\ln y_t = \beta_0 + \beta_1 \ln x_{1t} + \beta_2 \ln x_{2t} + \varepsilon_t.$$

А. Визначте параметри моделі, використовуючи дані таблиці.

Б. Чи є мультиколінеарність серед цих даних?

В. Побудуйте регресії:

$$1) \ln y_t = A_0 + A_1 \ln x_{1t};$$

$$2) \ln y_t = B_0 + B_1 \ln x_{2t};$$

$$3) \ln x_{1t} = C_0 + C_1 \ln x_{2t}.$$

Що можна сказати про природу мультиколінеарності даних, базуючись на цих регресіях?

Таблиця 5.19

Рік	Товарний імпорт, млн. дол.	ВВП, млрд. дол.	СРІ (споживчий індекс) 1967=100
1970	39866	992.7	116.3
1971	45579	1077.6	121.3
1972	55797	1185.9	125.3
1973	70499	1326.4	133.1
1974	103811	1434.2	147.7
1975	98185	1549.2	161.2
1976	124228	1718.0	170.5
1977	151907	1918.3	181.5
1978	176020	2163.9	195.4
1979	212028	2417.8	217.4
1980	249781	2631.7	246.8
1981	265086	2957.8	272.4
1982	247667	3069.3	289.1
1983	261312	3304.8	298.4

Джерело: "Economics Report of the President" (1985)

Г. Припустимо, що наявна мультиколінеарність даних, але b_1 , b_2 статистично значимі з 95%-м рівнем значимості; загальний F -тест також значимий. Чи серйозною в такому випадку є проблема мультиколінеарності?

Вправи. Гетероскедастичність

Вправа 5.2.1. У табл. 5.20 наведено дані про співвідношення обсягів продажу і наявних коштів у промисловості України (показники умовні), класифіковані за величиною активів установ за період з I кварталу 1991 р. до IV кварталу 1993 р. Це співвідношення можна розуміти як швидкість обертання доходів у корпоративному секторі.

Таблиця 5.20

Рік, квартал	1—10	10—25	25—50	50—100	100—250	250—1000	1000+
1991—I	3.696	3.929	3.858	3.966	4.819	4.557	4.860
1991—II	3.826	4.311	4.299	4.081	4.907	4.685	4.351
1991—III	3.338	4.035	4.082	4.145	4.691	4.309	4.088
1991—IV	3.272	3.265	3.874	3.485	3.778	4.120	3.765
1992—I	3.692	3.236	4.101	4.060	4.104	4.584	3.717
1992—II	3.818	4.010	4.719	4.009	4.064	4.457	4.280
1992—III	3.783	3.943	4.182	3.923	4.784	4.142	3.619
1992—IV	3.779	3.988	3.531	4.146	4.279	3.928	3.919
1993—I	4.291	4.428	4.272	4.571	4.583	4.053	3.630
1993—II	4.766	6.071	4.818	5.692	5.608	4.571	3.805
1993—III	4.733	5.357	5.090	5.357	4.680	4.654	3.772
1993—IV	5.316	4.621	4.766	4.867	4.666	4.380	4.072

1. Для кожного значення величини активів обчисліть середнє і стандартне відхилення співвідношення продажу і готівки.

2. Використовуючи величину активів як одиницю розрахунків, побудуйте графік для середнього і стандартного значення, розрахованих у пункті 1.

3. За допомогою відповідної регресійної моделі визначте, чи зростатиме стандартне відхилення при зростанні середнього значення. Якщо ні, поясніть, чому.

4. Якщо між стандартним відхиленням і середнім значенням є статистично важливе відношення, як можна трансформувати дані, щоб позбутися гетероскедастичності?

Вправа 5.2.2. Розглянемо тест однорідності дисперсій Бартлета¹. Припустимо, є k незалежних дисперсій $s_1^2, s_2^2, \dots, s_k^2$ з f_1, f_2, \dots, f_k д.ф. (кожна класу нормального розподілу з математичним сподіванням μ і дисперсією σ_i^2). Припустимо далі, що треба перевірити нульову гіпотезу $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2 = \sigma^2$, тобто кожна дисперсія є оцінкою одного і того ж класу дисперсії σ^2 . Якщо нульова гіпотеза правильна, тоді

$$s^2 = \frac{\sum_{i=1}^k f_i s_i^2}{\sum f_i} = \frac{\sum f_i s_i^2}{f} \text{ дає оцінку загальної оцінки класу дисперсій } \sigma^2, \text{ де}$$

$f_i = (N_i - 1)$, N_i — кількість спостережень в i -й групі, а $f = \sum_{i=1}^k f_i$. Бартлет показав, що нульову гіпотезу можна перевірити співвідношенням A/B , яке наближено розподілене за χ^2 розподілом з $(k-1)$ ступенями вільності, де

$$A = f \ln s^2 - \sum (f \ln s_i^2) \quad i \quad B = 1 + \frac{1}{3(k-1)} \left[\sum \left(\frac{1}{f_i} \right) - \frac{1}{f} \right].$$

Для того, щоб проілюструвати тест Бартлета, розглянемо такий приклад. Нехай ми маємо дані, наведені в табл. 5.21, які відображають медіанні зарплати жінок і чоловіків — економістів за напрямками спеціалізації.

Таблиця 5.21

Середня заробітна плата, тис.грн.

Спеціалізація	Жінки	Чоловіки
Фінанси	10.3	12.0
Економісти	9.3	13.0
Аналітики	8.0	11.6
Теоретики	8.7	10.8
Програмісти	12.0	11.5
Трудове право	9.0	12.2

¹ Bartlette M. Properties of Sufficiency and Statistical Tests//Proceedings of the Royal Society of London. — A, 160. — 1937. — P. 268.

1. Знайдіть середню зарплату та відхилення зарплати для жінок і чоловіків — економістів.

2. Чи є значною різниця між цими двома стандартними відхиленнями (можна застосувати тест Бартлета)?

3. Припустимо, що нам треба зробити прогноз про медіанну зарплату чоловіків, виходячи з медіанної зарплати жінок. Побудуйте для цього відповідну лінійну регресійну модель. Як ви діятимете, якщо у вашій моделі передбачається гетероскедастичність?

Вправа 5.2.3. У табл. 5.22 наведено медіанні зарплати економістів відповідно до їхнього наукового ступеня і віку.

Таблиця 5.22

Медіанні зарплати, тис.дол.

Вік, років	M.A.	Ph.D
25—29	8.0	8.8
30—34	9.2	9.6
35—39	11.0	11.0
40—44	12.8	12.5
45—49	14.2	13.6
50—54	14.7	14.3
55—59	14.5	15.0
60—64	13.5	15.0
65—69	12.0	15.0

1. Чи будуть дисперсії медіанних зарплат магістрів і докторів наук однаковими?

2. Якщо так, то як перевірити гіпотезу, що середні медіанні зарплати цих двох груп однакові?

3. Економісти із ступенем магістра віком від 35 до 54 років заробляють більше, ніж їхні ровесники із ступенем доктора. Як пояснити таке спостереження, якщо ви вважаєте, що доктор має заробляти більше магістра?

Вправа 5.2.4. У результаті опитування 10000 економістів щодо їхньої середньої зарплати було отримано такі дані (див. табл. 5.23).

1. Побудуйте регресійну модель, яка відбивала б рівень медіанної зарплати залежно від віку.

Примітка: для регресії припустимо, що медіанна зарплата відповідає середині вікового інтервалу: 7800 гривень відповідає віку 22.5 років тощо. Для останнього вікового інтервалу максимальний вік становить 75 років.

Таблиця 5.23

Вік, років	Зарплата
20—24	7800
25—29	8400
30—34	9700
35—39	11500
40—44	13000
45—49	14800
50—54	15000
55—59	15000
60—64	15000
65—69	14500
70	12000

2. Припустивши, що дисперсія помилки пропорційна квадрату віку, змініть дані таким чином, щоб помилки стали гетероскедастичними.

3. Зробіть завдання 2, припустивши, що дисперсія пропорційна віку. Яка із змін даних виглядає вірогіднішою і правдоподібнішою?

4. Якщо жодна з попередніх змін даних не виглядає правдоподібною, то припустіть, що величина дисперсії пропорційна умовному очікуванню медіанної зарплати для даного віку. Як би ви змінили дані для того, щоб у результаті отримати гомоскедастичну дисперсію?

Вправа 5.2.5. Чи можна оцінити параметри моделей

$$|e_i| = \sqrt{\beta_1 + \beta_2 x_i} + v_i,$$

$$|e_i| = \sqrt{\beta_1 + \beta_2 x_i^2} + v_i$$

методом звичайних найменших квадратів? Чому?

Якщо ні, то чи можете ви запропонувати формальний або неформальний метод оцінки параметрів таких моделей?

Вправа 5.2.6. Хоча логарифмічні моделі часто зменшують гетероскедастичність, треба звернути увагу на властивості помилок таких моделей. Наприклад, модель $y_i = \beta_1 x_i^{\beta_2} \varepsilon_i$ (1) можна розписати як

$$\ln y_i = \ln \beta_1 + \beta_2 \ln x_i + \ln \varepsilon_i. \quad (2)$$

1. Якщо $\ln \varepsilon_i$ повинен мати нульове математичне сподівання, яким має бути розподіл ε_i ?

2. Якщо $E(\varepsilon_i) = 1$, чи буде $E(\ln \varepsilon_i) = 0$? Чому?

3. Якщо $E(\ln \varepsilon_i) \neq 0$, як можна прийти до $E(\ln \varepsilon_i) = 0$?

Вправа 5.2.7

Маємо такі дані:

RSS_1 , засноване на перших 30 спостереженнях, $=55, df=25$;

RSS_2 , засноване на останніх 30 спостереженнях, $=140, df=25$.

Проведіть тест на гетероскедастичність Голдфелда — Квондта при 5%-му рівні значимості.

Вправа 5.2.8. Тест Бройша — Пагана

Розглянемо модель з k змінними:

$$y_i = \beta_1 + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i. \quad (1)$$

Припустимо, що дисперсія помилки $\sigma_i^2 = f(\alpha_1 + \alpha_2 z_{2i} + \dots + \alpha_m z_{mi})$ (2), тобто σ_i^2 є якоюсь функцією нестохастичних змінних, які умовно позначимо через z (відмітимо, що в ролі цих змінних можуть виступати і деякі X). Розглянемо особливий випадок, коли

$$\sigma_i^2 = \alpha_1 + \alpha_2 z_{2i} + \dots + \alpha_m z_{mi}, \quad (3)$$

тобто σ_i^2 є лінійною функцією від z . Якщо $\alpha_1 = \alpha_2 = \dots = \alpha_m = 0$, $\sigma_i^2 = \alpha_1$, яка є константою. Тому, щоб перевірити, чи дисперсія σ_i^2 гомоскедастична, треба перевірити гіпотезу, чи всі $\alpha_1 = \alpha_2 = \dots = \alpha_m = 0$. Це — основна ідея тесту Бройша — Пагана.

Сам тест складається з кількох кроків.

Крок 1. Оцінити невідомі параметри (1) за допомогою методу найменших квадратів, отримати залишки e_1, e_2, \dots, e_N .

Крок 2. Розрахувати $\tilde{\sigma}^2 = \sum e_i^2 / N$. Нагадаємо, що це оцінка максимальної правдоподібності σ^2 .

[Примітка: класична оцінка дисперсії дорівнює $\sum e_i^2 / (N - k)$].

Крок 3. Розрахувати змінні p_i , визначені як $p_i = e_i^2 / \tilde{\sigma}^2$.

Крок 4. Побудувати регресійну модель для p_i :

$$p_i = \alpha_1 + \alpha_2 z_{2i} + \dots + \alpha_m z_{mi} + v_i, \quad (4)$$

де v_i — залишки цієї регресії.

Крок 5. Отримати SSR (суму квадратів, що пояснює регресію) з (4), визначити

$$\theta = \frac{1}{2} (SSR). \quad (5)$$

Можна показати, що якщо є гомоскедастичність, а розмір вибірки N нескінченно зростає, $\theta \approx \chi_{m-1}^2$ (6), тобто θ розподілена за χ^2 -розподілом з $(m-1)$ ступенями вільності.

Звідси, якщо обчислена за тестом $\theta (= \chi^2)$ перевищує критичне значення χ^2 при обраному рівні значимості, можна відкинути припущення про гомоскедастичність; у протилежному разі, якщо θ не перевищує критичного значення, — прийняти його.

Для ближчого ознайомлення з тестом Бройша — Пагана розглянемо таку просту модель: $y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$, де y_t — наявна продукція промислового сектора України; x_t — розмір продажів в цьому секторі за 1977—1988 рр. Припустивши, що $\sigma_i^2 = \alpha_1 + \alpha_2 x_i$, використайте тест Бройша — Пагана [примітка: в цьому прикладі $Z=x$]. Використайте дані з табл. 5.24 і поясніть результат.

Вправа 5.2.9. Маємо таку модель:

$C_t = \beta_0 + \beta_1 x_t + \beta_2 D_t + \varepsilon_t$, де C_t — витрати на

Таблиця 5.25

Промислова продукція та обсяги продаж, трлн. ум. од.

Рік	Продукція	Продаж
1977	41.6	23.3
1978	45.6	23.4
1979	50.3	27.7
1980	51.9	28.7
1981	50.2	27.2
1982	52.9	30.3
1983	53.8	30.9
1984	54.9	30.9
1985	58.2	33.4
1986	60.0	35.0
1987	63.4	37.3
1988	68.2	41.0

сукупне приватне споживання в t -му році; x_t — валовий національний продукт року t ; D_t — національні витрати на оборону в році t . Дослідити вплив витрат на оборону на інші витрати в економіці.

Змінімо модель, прийнявши $\sigma_t^2 = \sigma^2(x_t)^2$. Отримаємо:

$$C_t / x_t = \beta_0(1 / x_t) + \beta_1 + \beta_2(D_t / x_t) + \varepsilon_t / x_t.$$

Емпіричні результати за даними 1967 — 1975 р. були такими (в дужках подано стандартні помилки).

$$\hat{C}_t = 26.19 + 0.6248x_t - 0.4398D_t$$

(2.73) (0.0060) (0.0736) $R^2 = 0.999$.

$$C_t / \hat{x}_t = 25.92(1 / x_t) + 6246 - 0.4315(D_t / x_t)$$

(2.22) (0.0068) (0.0597) $R^2 = 0.875$.

1. Яке припущення зроблено щодо природи гетероскедастичності? Поясніть.
2. Порівняйте результати обох регресій. Чи поліпшила трансформація первісної моделі результат, тобто чи зменшила оцінену стандартну помилку? Чому?
3. Чи можна порівняти два значення R^2 ? Чому? (Підказка: зверніть увагу на залежні змінні).

Вправи. Авторегресивні моделі

Вправа 5.4.1. Розглянемо таку модель:

$$y_t^* = \alpha + \beta_0 X_t + \varepsilon_t,$$

де y_t^* — бажані або довгострокові витрати на новий завод і обладнання; x_t — обсяг продажу; t — час. На підставі даних із табл. 5.25, використовуючи модель часткових пристосувань, оцінити параметри довго- і короткострокової функції попиту на витрати на новий завод і обладнання.

Вправа 5.4.2. Використовуючи дані з вправи 5.4.1, розглянемо таку модель:

$$y_t^* = \beta_0 x_t^\beta e^{\varepsilon_t}.$$

Базуючись на моделі часткових пристосувань, оцінити коротко- і довгострокову еластичність витрат на новий завод та обладнання за обсягом продажу. Порівняйте свої результати з отриманими у вправі 5.4.1.

Таблиця 5.25

Рік	y	x
1982	36.99	52.80
1983	33.60	55.91
1984	35.42	63.03
1985	42.35	72.93
1986	52.48	84.79
1987	53.66	86.60
1988	58.53	98.80
1989	67.48	113.20
1990	78.58	126.90
1991	95.92	143.94
1992	112.33	154.39
1993	126.54	168.13
1994	120.68	159.03
1995	116.20	170.44
1996	138.82	189.58

Вправа 5.4.3. Використовуючи дані з вправи 5.4.1, припустимо, що

$$y_t = \alpha + \beta x_t^* + \varepsilon_t,$$

де x_t^* — бажаний обсяг продажу. Оцініть параметри цієї моделі і порівняйте їх з отриманими у вправі 5.4.1. Яким чином можна визначити, яка модель правильна?

Вправа 5.4.4. Припустимо, вас запевняють, що відношення між витратами на новий завод і обладнання та обсягом продажу має такий вигляд:

$$y_t^* = \alpha + \beta x_t^* + \varepsilon_t,$$

де y_t^* — бажані витрати, а x_t^* — очікуваний обсяг продажу.

Використовуючи дані, отримані у вправі 5.4.1, оцініть цю модель і прокоментуйте результати.

Вправа 5.4.5. Доведіть формулу (5.4.30).

Вправа 5.4.6. Доведіть формулу (5.4.27).

Вправа 5.4.7. Припустимо, що ціни формуються на основі такої гіпотези адаптивних очікувань:

$$P_t^* = \gamma P_{t-1} + (1 - \gamma)P_{t-1}^*,$$

де P_t^* — очікувана ціна, а P — фактична ціна.

Заповніть подану нижче таблицю за умови, що $\gamma = 0.5$.

Період	P^*	P
$t-3$	100	110
$t-2$		125
$t-1$		155
t		185
$t+1$		—

Вправа 5.4.8. Розглянемо модель

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 y_{t-1} + v_t.$$

Припустимо, y_{t-1} і v_t корелюються між собою. Щоб уникнути кореляції, використаємо метод допоміжних змінних. Спочатку знаходимо регресію y_t за x_{1t} і x_{2t} і отримаємо оцінку \hat{y}_t з цього рівняння. Потім знаходимо

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 \hat{y}_{t-1} + v_t,$$

де \hat{y}_{t-1} оцінюється з першого рівняння.

А. Яким чином ця процедура усуває кореляцію між y_{t-1} і v_t в початковій моделі?

Б. Які переваги рекомендованої процедури перед іншими підходами?

Вправа 5.4.9. А. Довести (5.4.13).

Б. Оцінити медіанний лаг для $\lambda = 0.2, 0.4, 0.6, 0.8$.

В. Чи існує якесь постійне співвідношення між значенням λ і значенням медіанного лагу?

Вправа 5.4.10. А. Довести, що для моделі Койка середній лаг обчислюється так, як наведено у (5.4.15).

Б. Якщо λ відносно велике, який це має сенс?

Вправа 5.4.11. Використовуючи формулу для середнього лагу, наведену в (5.4.14), перевірте середній лаг 10.959, використаний в ілюстрації до табл. 5.4.1.

Вправа 5.4.12. Припустимо,

$$M_t = \alpha + \beta_1 y_t^* + \beta_2 R_t^* + \varepsilon_t$$

де M — попит на готівку; y_t — очікуваний дійсний доход; R^* — очікувана відсоткова ставка. Припустимо також, що очікування формуються таким чином:

$$y_t^* = \gamma_1 y_t + (1 - \gamma_1) y_{t-1}^*;$$

$$R_t^* = \gamma_2 R_t + (1 - \gamma_2) R_{t-1}^*,$$

де γ_1 і γ_2 — коефіцієнти очікування, які знаходяться у межах між 0 і 1.

А. Як можна виразити M_t через спостережувані параметри?

Б. Які проблеми оцінювання Ви можете передбачити?

Вправа 5.4.13. Якщо обчислити (5.4.27) за методом найменших квадратів, чи можна отримати оцінки початкових параметрів? Які проблеми можуть виникнути?

Вправа 5.4.14. *Модель серійної кореляції.* Розглянемо таку модель:

$$y_t = \alpha + \beta x_t + \varepsilon_t.$$

Припустимо, ε_t відповідає авторегресійній схемі першого порядку Маркова, розглянутій у попередньому розділі, тобто

$$\varepsilon_t = \rho \varepsilon_{t-1} + \mu_t,$$

де ρ — коефіцієнт автокореляції (першого порядку) і μ_t відповідає класичним припущенням. Як показано раніше, модель

$$y_t = \alpha(1 - \rho) + \beta(x_t - \rho x_{t-1}) + \rho y_{t-1} + \varepsilon_t$$

має серійно незалежну помилку, що, таким чином, дозволяє застосовувати МНК. Але ця модель, що має назву *модель серійної кореляції*, дуже нагадує модель Койка, модель адаптивних очікувань і модель часткових пристосувань. Як визначити, яку саме модель з вищезазначених треба обрати в кожній конкретній ситуації?

Вправа 5.4.15. Розглянемо модель Койка (або в нашому випадку модель адаптивних очікувань), показану в (5.4.12), тобто

$$y_t = \alpha(1 - \lambda) + \beta_0 x_t + \lambda y_{t-1} + (\varepsilon_t - \lambda \varepsilon_{t-1}).$$

Припустимо, в початковій моделі ε_t відповідає авторегресійній схемі першого порядку $\varepsilon_t - \rho \varepsilon_{t-1} = \eta_t$, де ρ — коефіцієнт автокореляції і де η_t відповідає всім припущенням.

А. Якщо $\rho = \lambda$, чи можна модель Койка оцінити за МНК?

Б. Чи будуть отримані оцінки незміщеними?

В. Чи можна взагалі припускати, що $\rho = \lambda$?

Вправа 5.4.16. *Триангулярна, або арифметична дистрибутивно-лагова модель.* Ця модель припускає, що пояснювальна змінна найбільше впливає в поточний період часу, а потім рівномірно спадає до нуля із заглибленням у минуле. Відповідно до цього розподілу припустимо, що ми маємо таку послідовність регресій:

$$y_t = \alpha + \beta \left(\frac{2x_t + x_{t-1}}{3} \right);$$

$$y_t = \alpha + \beta \left(\frac{3x_t + 2x_{t-1} + x_{t-2}}{6} \right);$$

$$y_t = \alpha + \beta \left(\frac{4x_t + 3x_{t-1} + 2x_{t-2} + x_{t-3}}{10} \right) \text{ і т. д.}$$

Оберемо за “кращу” ту регресію, в якій R^2 найбільший. Прокоментуйте цю стратегію.

Вправа 5.4.17. Базуючись на щоквартальних даних за 1950 — 1960 рр., Ф.Бреклінг отримав таку функцію попиту на працю для британської економіки (числа в дужках — стандартні помилки).

$$\hat{E}_t = 14.22 + 0.172Q_t - 0.028t - 0.0007t^2 - 0.297E_{t-1}$$

$$(2.61) \quad (0.014) \quad (0.015) \quad (0.0002) \quad (0.033)$$

$$\bar{R}^2 = 0.76 \quad d = 1.37,$$

де $\hat{E}_t = (E_t - E_{t-1})$; Q — випуск; t — час. Це рівняння базувалося на припу-

щенні, що бажаний рівень зайнятості E_t^* — функція від випуску, часу і квадрату часу, а також на гіпотезі, що $E_t - E_{t-1} = \delta(E_t^* - E_{t-1})$, де δ — коефіцієнт пристосування, $0 < \delta < 1$.

А. Проінтерпретуйте дане рівняння.

Б. Яке значення δ ?

В. З оціненої короткострокової функції попиту знайдіть довгострокову функцію попиту на працю.

Г. Яким чином ви здійснили б тестування серійної кореляції в наведеній моделі?

Вправа 5.4.18. Вивчаючи попит ферм на трактори, Грилічез використав таку модель:

$$T_t^* = \alpha x_{1,t-1}^{\beta_1} x_{2,t-1}^{\beta_2},$$

де T^* — бажана кількість тракторів; x_1 — відносна ціна тракторів; x_2 — відсоткова ставка.

Використовуючи модель часткових пристосувань, він отримав такі результати за період 1921—1957 рр. (числа в дужках — стандартні помилки оцінених параметрів).

$$\log T_t = \text{const} - 0.218 \log x_{1,t-1} - 0.855 \log x_{2,t-2} + 0.864 \log T_{t-1}$$

$$(0.051) \quad (0.170) \quad (0.035)$$

$$R^2 = 0.987.$$

А. Знайдіть оцінений коефіцієнт пристосувань.

Б. Знайдіть коротко- і довгострокову еластичність за ціною.

В. Обрахуйте відповідні еластичності за відсотковою ставкою.

Г. Які причини високого й низького рівня пристосувань у наведеній моделі?

Вправа 5.4.19. Де б не з'явилась лагова залежна змінна серед незалежних змінних, R^2 звичайно вище, ніж якби вона не була включена до моделі. Чому?

Вправа 5.4.20. Проаналізуйте рівняння 5.4.54.

$$\beta_i = a_0 + a_1 i + a_2 i^2 + \dots + a_m i^m.$$

Щоб отримати дисперсію $\hat{\beta}_i$ з дисперсії \hat{a}_i , використаємо таку формулу:

$$\text{var}(\hat{\beta}_i) = \text{var}(\hat{a}_0 + \hat{a}_1 i + \hat{a}_2 i^2 + \dots + \hat{a}_m i^m) =$$

$$= \sum_{j=0}^m i^{2j} \text{var}(\hat{a}_j) + 2 \sum_{j < p} i^{(j+p)} \text{cov}(\hat{a}_j, \hat{a}_p).$$

А. Використовуючи попередню формулу, знайдіть дисперсію $\hat{\beta}_i$, виражену як:

$$\hat{\beta}_i = \hat{a}_0 + \hat{a}_1 i + \hat{a}_2 i^2;$$

$$\hat{\beta}_i = \hat{a}_0 + \hat{a}_1 i + \hat{a}_2 i^2 + \hat{a}_3 i^3.$$

Б. Якщо дисперсії \hat{a}_i великі порівняно із самими значеннями, чи будуть дисперсії $\hat{\beta}_i$ також великими? Чому так або чому ні?

Вправа 5.4.21. Проаналізуйте подану дистрибутивно-лагову модель:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \beta_4 x_{t-4} + \varepsilon_t.$$

Припустимо, що β_i можна адекватно представити поліномом другого ступеня:

$$\beta_i = a_0 + a_1 i + a_2 i^2.$$

Як би Ви оцінили β_i , якщо б захотіли накласти обмеження $\beta_0 = \beta_4 = 0$?

Вправи. Dummy-змінні

Вправа 5.5.1. Якщо є щомісячні дані за кілька років, скільки dummy-змінних ви введете для тестування таких гіпотез:

а) усі 12 місяців року мають вплив сезонних коливань;

б) тільки лютий, квітень, червень, серпень, жовтень і грудень мають вплив сезонних коливань.

Вправа 5.5.2. Що пояснює встановлення погодинної оплати праці заробітчан у регресії (5.5.18). З даного рівняння отримайте рівняння погодинної оплати праці для таких типів заробітчан:

а) білий, неміський, із західних районів, з вищою освітою;

б) чорношкірий, міський, не із західних районів, без вищої освіти;

в) білий, неміський, не із західних районів, з вищою освітою.

Вправа 5.5.3. Базуючись на щорічній інформації за 1972 — 1979 рр., Уільям Нортхауз оцінив таку модель для пояснення поведінки цін на нафту ОПЕК (у круглих дужках стандартні помилки):

$$y_t = 0.3x_{1t} + 5.22x_{2t}$$

$$(0.03) \quad (0.50),$$

де y_t — різниця у цінах поточного і попереднього періодів (доларів за барель);

x_{1t} — різниця в цінах ОПЕК поточного і попереднього періодів;

$x_{2t} = 1$ для 1974 р. і 0 для усіх інших випадків (примітка: в 1973 — 1974 рр. на торгівлю нафтою було накладено ембарго).

Проінтерпретуйте цей результат і покажіть його графічно.

Вправа 5.5.4. Припустимо, що ми маємо модель

$$y_i = a_1 + a_2 D_{2i} + a_3 D_{3i} + a_4 (D_{2i} D_{3i}) + \beta x_i + \varepsilon_i,$$

де y_i — щорічна зарплата викладача вузу;

x_i — роки викладацького досвіду;

$D_2 = 1$ — якщо це чоловік, $i = 0$ у нашому випадку;

$D_3 = 1$ — якщо це жінка, $i = 0$ у нашому випадку.

А. Вираз $(D_{2i} D_{3i})$ відображає вплив взаємозв'язку. Що цей вираз означає?

Б. Знайдіть $E(y_i / D_2=1, D_3=1, x_i)$ і проінтерпретуйте його.

Вправа 5.5.5. Змодифікуйте *Chow*-тест (коли кількість спостережень менша за кількість оцінених параметрів). Це стосується моделей (5.5.21) і (5.5.22). Припустимо, що кількість спостережень у другому періоді N_2 менша або дорівнює кількості оцінюваних параметрів. У цьому випадку *Chow* пропонує таку модифікацію свого тесту: нехай $S_1 = SSE$ із зібраної регресії (5.5.23), $S_2 = SSE$ — з регресії першого періоду (припускаємо, що $N_1 >$ кількість параметрів). Тепер застосуємо *F*-тест:

$$F = \frac{(S_1 - S_2) / N_2}{S_2 / (N_1 - K)},$$

який має N_2 і $(N_1 - K)$ ступенів вільності. Якщо ці *F*-значення статистично значимі, ми відхиляємо гіпотезу про те, що спостереження зібраної регресії першого періоду базуються на N_1 спостереженнях. Якщо значення незначиме, можете не відхиляти цю гіпотезу.

Використовуйте дані таблиці 5.10 для перевірки гіпотези, що останні два спостереження надійшли від тієї ж сукупності, що і 16 перших.

Вправа 5.5.6. Дані таблиці 5.11 відібражають залежність між прибутками та обсягом продажу приватних фірм України і регресією (5.5.41). Розв'янемо модель регресії для того, щоб протестувати гіпотезу, за якою нахил і перетин регресії прибутків на обсяг продажів у другому кварталі відрізняється від обсягів усіх інших кварталів. Зробіть необхідні обчислення.

Вправа 5.5.7. Дані, не залежні від сезонних коливань. Приклад у вправі 5.5.6 показує, як *dumt*-змінні можна використовувати з урахуванням сезонних коливань. Після оцінки регресії (5.5.41) стало відомо, що лише *dumt*-змінна, яка асоціюється з другим кварталом, є статистично значима; це означає, що лише на показники другого кварталу впливали сезонні коливання. Тому один із методів, за допомогою якого можна позбутися впливу сезонних коливань, полягає у тому, щоб відняти від щорічних прибутків і продажів другого кварталу значення 1322.8938 (млн. грн.), коефіцієнт другого кварталу, і оперувати вже трансформованою регресією.

Вправа 5.5.8. Об'єднання міжгалузевої і періодичної інформації. Припустимо, що у нашому розпорядженні є дані про випуск, робочу силу і капіталовкладення для фірми N в промисловості для T періодів, і можна застосувати виробничу функцію такого типу:

$$y_{it} = a + \beta_1 x_{1t} + \beta_2 x_{2it} + \varepsilon_{it}, \quad i = 1, 2, 3, \dots, N \\ t = 1, 2, 3, \dots, T$$

де y — випуск;

x_1 — капіталовкладення;

x_2 — робоча сила.

Припустимо, ви володієте відповідними даними і вас попросили змінити модель таким чином:

а) фірми розрізняються за ефективністю управлінського персоналу, різниця впливає лише на перетин a ; це називається ефектом фірми.

б) усі фірми мають однакову ефективність управлінського персоналу, але перетин a з роками зміщується; це називається часовим ефектом.

в) на перетин попередньої функції впливає як часовий ефект, так і ефект фірми.

г) яке припущення ви робите щодо випадкової величини ε_{it} ?

Вправа 5.5.9. Було оцінено таку регресію, що базується на 77 спостереженнях міжгалузевої інформації:

$$P_i = \beta_0 + \beta_1 D_{ni} + \beta_2 D_{2i} + \beta_3 D_{3i} + u_i,$$

де P_i — ціна за унцію (1 унція = 28.3 г);

$D_{ni} = 001$, якщо магазин за зниженими цінами;

= 010, якщо одностипний магазин однієї фірми;

= 100, якщо універсальний магазин;

$D_{2i} = 10$, якщо фабричний товар;

= 01, якщо нефабричний товар;

$D_{3i} = 0001$ — посудина місткістю 67.6 унцій (2 л);

= 0010 — 28—33.8-унцієві посудини;

= 0100 — посудина місткістю 584 г (16 унцій);

= 1000 — трьохсотграмові пляшки (12 унцій).

Результати такі:

$$\hat{P}_i = 0.0143 - 0.000004 D_{1i} + 0.0090 D_{2i} + 0.0000 D_{3i} \\ (0.00001) (0.00011) (0.00000) \\ t = (-0.3837) (8.3927) (5.8125) \\ R = 0.6033.$$

Примітка: стандартні помилки показано з точністю до сотисячних.

А. Прокоментуйте спосіб, за допомогою якого *dumt*-змінні було б подано у моделі.

Б. Припустимо, що структура дитму-змінних нас влаштовує. Як би ви проінтерпретували результати?

В. Коефіцієнт D_3 додатний і статистично значимий. Як би ви раціоналізували цей результат?

Вправа 5.5.10. Розглянемо відомий з літератури приклад. Для оцінки політики Феда у сфері дерегулювання процентних ставок у липні 1979 р. було оцінено таку модель для періоду 1975-III — 1983-II:

$$\hat{y}_t = 8.5871 - 0.1328P_t - 0.7102U_{nt} - 0.2389M_t + 0.6592Y_{t-1} + 2.5831D_t$$

(1.9563) (0.0992) (0.1909) (0.0727)
(0.1036) (0.7549)

$$R^2 = 0.9156,$$

де y — 3-місячна ставка процента, встановлена Міністерством фінансів;

P — очікуваний рівень інфляції;

U_n — рівень безробіття з урахуванням сезонних коливань;

M — зміни у монетарній базі.

D — дитму-змінна, що набуває значення 1 для спостережень, починаючи з 1 липня 1979 р.

А. Проінтерпретуйте ці результати.

Б. Який ефект дерегулювання процентних ставок? Подайте економічну інтерпретацію.

ЗАСТОСУВАННЯ ЕКОНОМІЧНИХ ЗНАЙ

Визначення ціни на акцію

Позначимо ціну акції через y , дивіденди на акцію — x_1 і отримуваний прибуток на акцію — x_2 . Під “акцією” розумітимемо акції 30 підприємств, котрі використовуються для утворення індексу Доу Джонса (Dow Jones Industrial Average).

Тоді рівняння регресії матиме вигляд:

$$y_{it} = b_0 + b_1x_{1it} + b_2x_{2it} + e_{it},$$

де y_{it} — ціна i -ї акції t -го року;

x_{1it} — дивіденди на акції t -го року;

x_{2it} — отриманий прибуток на i -ту акцію t -го року.

Ми отримали модель залежності ціни акції від дивідендів і одержаного прибутку. Результати розрахунків за цією моделлю¹ наведено у табл. 5.26. Звичайно, ми маємо заперечити наявність мультиколінеарності між факторами x_1 і x_2 . Розглянемо результати, отримані для трьох різних періодів часу.

Таблиця 5.26

	Період часу		
	1	2	3
b_0	22.77 (4.15)	11.33 (2.33)	21.52 (2.81)
b_1	10.73 (4.28)	12.43 (4.41)	12.55 (3.16)
b_2	1.43 (2.3)	3.08 (2.39)	3.01 (1.81)
R^2	0.54	0.70	0.54
F -cm	16.11	31.44	15.97

Результати розрахунків третього періоду використаємо для визначення ступеня мультиколінеарності. Оскільки йдеться про двофакторну модель, то в ролі коефіцієнтів детермінації виступають парні коефіцієнти кореляції, що їх подано в табл. 5.27.

З таблиці видно, що змінні корелюють між собою. Для того, щоб виявити рівень колінеарності для фактора x_1 , значення $r_{1,2} = 0.6$ підставимо у формулу для розрахунку VIF_1 . Отримаємо:

$$VIF_1 = 1/(1 - 0.6^2) = 1.58.$$

Цей показник є значно меншим від 10, тому можна зробити висновок, що рівень колінеарності для цієї регресійної моделі є незначним.

Таблиця 5.27

Кореляційна матриця

Змінні	y	x_1	x_2
y	$r_{yy} = 1$	$r_{yx_1} = 0.69$	$r_{yx_2} = 0.61$
x_1	—	$r_{yx_1} = 1$	$r_{x_1x_2} = 0,6$
x_2	—	—	$r_{x_2x_2} = 1$

¹ Marquardt D.W. You Should Standardize the Prediction Variables in Your Regression Models//Journal of the American Statistical Association. — 1980. — №75. — P. 87—91.

Функція витрат. Тестування мультиколінеарності

Розглянемо функцію річних витрат, необхідних для підтримки торгового судна у робочому стані. Було обстежено 96 суден і побудовано таку регресійну модель¹:

$$\log y = b_0 + b_1 \log x_1 + b_2 \log x_2 + \dots + b_7 \log x_7 + \log e,$$

де y — річні витрати судна, дол.;

x_1 — вік судна (роки);

x_2 — величина судна (тоннажність, тис. т);

x_3 — кількість часу до наступного ремонту (цикл ремонтів, роки);

x_4 — споживання пального;

x_5 — dummy-змінна за типом використання пального (дизельне, ядерне тощо):

$x_5 = 1$, якщо судно використовує дизельне пальне;

$x_5 = 0$, якщо використовується інше пальне.

x_6 — dummy-змінна для визначення складності судна:

$x_6 = 1$, якщо є радарний захист;

$x_6 = 0$, якщо його немає.

x_7 — dummy-змінна, яка описує, чи буде судно перероблено відповідно до FRAM-програми (Fleet Rehabilitation And Modernization):

$x_7 = 1$, якщо використовують FRAM;

$x_7 = 0$, якщо ні.

Після того, як було знайдено параметри за методом найменших квадратів, отримано такі дані:

	b_1	b_2	b_3	b_4	b_5	b_6	b_7
$\hat{\sigma}_{bi}$	0.34 (0.03)	0.4 (0.08)	-0.79 (0.09)	0.05 (0.10)	-0.03 (0.09)	0.11 (0.04)	-0.16 (0.06)

$$R^2_{yx_1x_2\dots x_7} = 0.8 \quad F_{(v_1=k-1=7; v_2=96-8=88; 95\%)} = 56.$$

Далі було розраховано коефіцієнти детермінації допоміжних регресій і відповідні значення F -статистики:

$$R^2_{x_1x_2\dots x_7} = 0.30 \quad F_{x_2} = 19.9;$$

$$R^2_{x_2x_1\dots x_7} = 0.57 \quad F_{x_3} = 6.3;$$

$$R^2_{x_4x_1\dots x_7} = 0.76 \quad F_{x_4} = 47.5;$$

$$R^2_{x_5x_1\dots x_7} = 0.76 \quad F_{x_5} = 47.1;$$

¹ Farar D.E., Glauber R.R. Multicollinearity in Regression Analysis: The Problem Revisited // Review of Economics and Statistics. — Vol. 49. — P. 92—107.

$$R^2_{x_6x_1\dots x_7} = 12.7 \quad F_{x_7} = 4.8;$$

$$R^2_{x_6x_1\dots x_7} = 0.46 \quad F_{x_6} = 12.7;$$

$$R^2_{x_7x_1\dots x_7} = 0.24 \quad F_{x_7} = 4.8.$$

Як бачимо, при високому загальному коефіцієнті детермінації спостерігаються високі значення коефіцієнтів детермінації допоміжних регресій для x_4 (споживання пального) і x_5 (дизельне пальне).

Для того, щоб встановити, які з факторів приводять до мультиколінеарності у факторах x_4 і x_5 , розрахуємо парні коефіцієнти кореляції. Дані наведено у табл. 5.28.

Таблиця 5.28

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	—						
x_2	$r_{12} = 0.13$	—					
x_3	$r_{13} = 0.21$	$r_{23} = 0.27$	—				
x_4	$r_{14} = -0.35$	$r_{24} = 0.34$	$r_{34} = 0.09$	—			
x_5	$r_{15} = -0.27$	$r_{25} = -0.21$	$r_{35} = 0.12$	$r_{45} = -0.68$	—		
x_6	$r_{16} = 0.45$	$r_{26} = 0.08$	$r_{36} = -0.35$	$r_{46} = 0.31$	$r_{56} = 0.5$	—	
x_7	$r_{17} = 0.34$	$r_{27} = -0.13$	$r_{37} = 0.06$	$r_{47} = 0.4$	$r_{57} = 0.27$	$r_{67} = -0.26$	—

Аналіз таблиці показує, що є залежність між факторами x_4 і x_5 , а також між x_5 і x_6 .

Висновки

1. Є досить великий ступінь залежності між факторами у функції витрат, а саме: мультиколінеарність є у факторах x_4 і x_5 .

2. Мультиколінеарність викликана залежністю між x_4 і x_5 , а також між x_5 і x_6 .

РОЗДІЛ 6. ЕКОНОМЕТРИЧНІ СИМУЛЬТАТИВНІ МОДЕЛІ

6.1. ПОНЯТТЯ ПРО ОДНОЧАСНУ ЗАЛЕЖНІСТЬ ЕКОНОМІЧНИХ ЗМІННИХ. ПРИКЛАДИ ЕКОНОМЕТРИЧНИХ СИМУЛЬТАТИВНИХ МОДЕЛЕЙ

Застосування методу найменших квадратів (МНК) до окремого рівняння передбачає, що незалежні змінні (фактори) — справді екзогенні і що є тільки односторонній зв'язок між залежною змінною y та незалежними змінними x . Якщо ці умови не виконуються, тобто, якщо змінні x визначаються через y , то порушується одне з припущень класичного регресійного аналізу, а саме припущення 4 про незалежність факторів та випадкових величин ($\text{cov}(x, \varepsilon) \neq 0$) і використання методу найменших квадратів для знаходження невідомих параметрів моделі призведе до появи неефективних оцінок з відхиленням.

Якщо $y = f(x)$ і, в свою чергу, також $x = f(y)$, недопустимо застосовувати одне регресійне рівняння для опису взаємозв'язку між y та x . У такому разі ми переходимо від регресійної моделі з одним рівнянням до регресійної моделі з багатьма рівняннями, серед яких можуть бути рівняння, які включають x та y у ролі як ендогенних, так і екзогенних змінних. Система, що описує таку взаємну залежність між змінними, називається *системою одночасних або симультативних рівнянь*.

Наведемо декілька прикладів застосування симультативних моделей, які ілюструють значення одночасних зв'язків при описанні економічних явищ, та покажемо, чому при цьому виникає порушення припущення 4.

Приклад 6.1. Припустимо, що нам потрібно оцінити попит на олію. З економічної теорії нам відомо, що попит на будь-який товар залежить від його ціни P , цін на інші товари P_0 , та доходу y . Тому функцію попиту на олію можна записати у вигляді:

$$Q = \beta_0 + \beta_1 P + \beta_2 P_0 + \beta_3 y + \varepsilon,$$

де Q — обсяг попиту, P — середня ціна олії, P_0 — ціни на інші товари, y — доход, ε — випадкова величина.

Якщо ми застосуємо для знаходження невідомих параметрів цього рівняння метод найменших квадратів, то отримаємо зміщені оцінки (або, як ми ще їх називаємо, оцінки з відхиленням) для β_0 та β_1 , оскільки P та ε не незалежні. Попит на будь-який товар є функцією від його ціни, але одночасно і ринкова ціна змінюється під впливом обсягу попиту. Внаслідок

док цього наведене вище одиничне рівняння не можна розглядати як повну модель. Система повинна містити принаймні ще одне рівняння, що вказує на зв'язок між P та Q , наприклад:

$$P = c_0 + c_1 Q + c_2 W + v,$$

де W — може бути, наприклад, індексом погодних умов.

Підставляючи в це рівняння вираз для Q , отримаємо:

$$P = c_0 + c_1(\beta_0 + \beta_1 P + \beta_2 P_0 + \beta_3 y + \varepsilon) + c_2 W + v.$$

Очевидно, що P залежить від ε , а отже, припущення 4 класичної регресії порушується. P не є екзогенною змінною в функції попиту.

Приклад 6.2. Нехай нам потрібно оцінити пропозицію грошової маси. Цей процес регулюється урядом з метою уникнення інфляції. Отже, можна сказати, що основним детермінантом урядового рішення відносно грошової пропозиції буде реальний рівень доходів. Тому функцію грошової пропозиції можна записати у вигляді:

$$M = \beta_0 + \beta_1 y + \varepsilon,$$

де M — грошова пропозиція, y — рівень реального доходу.

Однак рівень реального доходу в свою чергу залежить від грошової пропозиції та інших реальних сил, наприклад, інвестиційних рішень підприємців, соціальної політики уряду та багатьох інших. Внаслідок цього грошову пропозицію не можна розглядати як модель з одного рівняння — y не є справді екзогенною змінною. Між M та y є взаємозалежність. Рівняння цього зв'язку має вигляд:

$$y = \alpha_0 + \alpha_1 M + \alpha_2 I_t + \dots + v.$$

Підставляючи вираз для M , отримаємо:

$$y = \alpha_0 + \alpha_1(\beta_0 + \beta_1 y + \varepsilon) + \alpha_2 I_t + \dots + v.$$

Очевидно, що $y = f(\varepsilon)$, а отже, у функції грошової пропозиції змінна y залежна від випадкової змінної ε .

Приклад 6.3. Кейнсіанська модель визначення доходу

Розглянемо просту кейнсіанську модель визначення доходу.

Функція споживання

$$C = \beta_0 + \beta_1 y + \varepsilon, \quad 0 < \beta_1 < 1. \quad (6.1)$$

Тотожність доходу

$$y_t = C_t + I_t (= S_t), \quad (6.2)$$

де C — витрати на споживання;

y — доход;

I — інвестиції (припускаються екзогенними);

S — заощадження;

t — час;

ε — випадкова величина;

β_0 і β_1 — невідомі параметри, які потрібно оцінити.

Параметр β_1 відомий з економічної теорії як гранична схильність до споживання (MPC — приріст споживчих витрат, пов'язаний зі зростанням доходу, наприклад, на одну гривню), знаходиться на проміжку від 0 до 1. Рівняння (6.1) — це стохастична функція споживання, а рівняння (6.2) — це тотожність національного доходу, яка визначає, що загальний доход дорівнює загальним витратам на споживання та загальним інвестиційним витратам; зрозуміло, що інвестиційні витрати дорівнюють загальним заощадженням.

У функції споживання C та y взаємозалежні, тому y в (6.1) не є незалежною змінною від ε , бо зміщення ε також зміщує й функцію споживання, що у відповідь впливає на y . В такому разі класичний метод найменших квадратів непридатний для оцінки параметрів у рівнянні (6.1). Якщо його застосувати, то отримані оцінки будуть зміщеними.

Приклад 6.4. Модель “зарплата — ціна”

Розглянемо таку модель типу Філіпса “зарплата — ціна”:

$$W_t^0 = a_0 + a_1 UN_t + a_2 P_t^0 = \varepsilon_{1t}; \quad (6.3)$$

$$P_t^0 = \beta_0 + \beta_1 W_t^0 + \beta_2 R_t^0 + \beta_3 M_t^0 + \varepsilon_{2t}, \quad (6.4)$$

де W^0 — норма зміни зарплати в грошовому вираженні;

UN — рівень безробіття, %;

P^0 — норма зміни цін;

R^0 — норма зміни витрат капіталу;

M^0 — норма зміни цін на імпортовану сировину;

t — час;

ε — випадкові величини.

Оскільки змінна ціни P^0 входить у рівняння зарплати, а у рівняння ціни входить змінна зарплати, обидві змінні взаємозалежні. Тому змінні корелюють з відповідними випадковими величинами, що веде до порушення припущення 4 та неможливості застосування методу найменших квадратів при оцінюванні невідомих параметрів моделі.

Приклад 6.5. IS-модель

Відому макроекономічну модель IS , або рівновагу на ринку товарів, можна виразити у нестохастичній формі:

$$\text{функція споживання: } C_t = \beta_0 + \beta_1 y_{dt}, \quad 0 < \beta_1 < 1; \quad (6.5)$$

$$\text{функція податків: } T_t = \alpha_0 + \alpha_1 y_t, \quad 0 < \alpha_1 < 1; \quad (6.6)$$

$$\text{функція інвестицій: } I_t = \gamma_0 + \gamma_1 r_t; \quad (6.7)$$

$$\text{визначення: } y_{dt} = y_t - T_t; \quad (6.8)$$

$$\text{державні витрати: } G_t = \bar{G}; \quad (6.9)$$

$$\text{тотожність національного доходу: } y_{dt} = C_t + I_t + G, \quad (6.10)$$

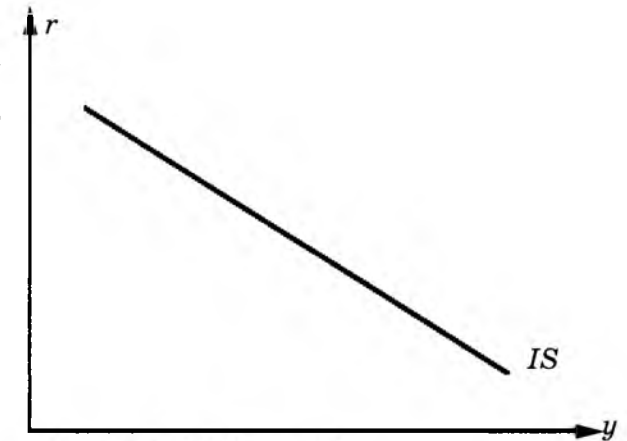
де y — національний доход; C — витрати на споживання; I — заплановані або бажані чисті інвестиції; \bar{G} — витрати державного сектора; T — податки; y_d — наявний доход і r — процентна ставка. Криву IS зображено на мал. 6.1.

Якщо у рівняння (6.5) підставити (6.8) та (6.6), а потім підставити отримане таким чином рівняння для C та рівняння (6.9) у рівняння (6.10.), остаточно отримаємо IS -рівняння:

$$y_t = \pi_0 + \pi_1 r_t, \quad (6.11)$$

$$\pi_0 = \frac{\beta_0 - \alpha_0 \beta_1 + \gamma_0 + \bar{G}}{1 - \beta_1(1 - \alpha_1)};$$

$$\pi_1 = \frac{1}{1 - \beta_1(1 - \alpha_1)}.$$



Малюнок 6.1. Крива IS

Рівняння (6.11) — це рівняння рівноваги на ринку товарів, відоме з макроекономічної теорії як IS -рівняння. Воно дає таку комбінацію процентної ставки та доходу, за якої ринок товарів перебуває у рівновазі.

Що б трапалося, якби ми оцінили, скажімо, функцію споживання (6.5) ізольовано? Чи вдалося б нам отримати ефективні, незміщені оцінки невідомих параметрів? Малоймовірно, тому що споживання залежить від наявного доходу споживача, який у свою чергу залежить від національного доходу y , але останній залежить від r , G , та інших параметрів. Тому, якщо ми не враховуємо всі ці впливи, у простій регресії C від y не буде отримано незміщених оцінок параметрів.

Приклад 6.6. Модель LM

Іншою стороною відомої $IS-LM$ -моделі є LM -модель, або рівновага на ринку грошей, — співвідношення, яке дає таку комбінацію процентної ставки і рівня доходу, за якого досягається рівновага на ринку грошей, тобто

попит на гроші дорівнює пропозиції грошей. Алгебраїчно модель у нестохастичній формі можна виразити так:

функція попиту на гроші:

$$M_t^d = a + by_t - cr_t. \quad (6.12)$$

функція пропозиції грошей:

$$M_t^s = \bar{M}, \quad (6.13)$$

умова рівноваги:

$$M_t^d = M_t^s, \quad (6.14)$$

де y — доход; \bar{M} — середній рівень пропозиції грошей; r — процентна ставка.

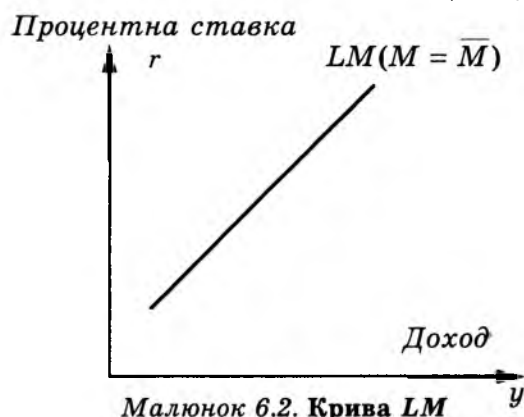
Прирівнявши попит на гроші і функцію пропозиції та спростивши вираз, отримуємо класичне LM -рівняння:

$$y_t = \lambda_0 + \lambda_1 M + \lambda_2 r_t; \quad (6.15)$$

$$\text{де } \lambda_0 = -\left(\frac{a}{b}\right); \lambda_1 = \frac{1}{b}; \lambda_2 = \frac{c}{b}. \quad (6.16)$$

Для даних $M = \bar{M}$ крива LM показує зв'язок (6.15), відображений на мал. 6.2.

Криві IS і LM відповідно показують, що вся множина процентних ставок узгоджується з рівновагою на ринку товарів і на ринку грошей. Звичайно, тільки одна процентна ставка і один рівень доходу будуть одночасно відповідати рівновазі на цих ринках.



Малюнок 6.2. Крива LM

Приклад 6.7. Економетрична модель Клейна

Наведемо одну з перших економетричних моделей. Першовідкривачем у цій галузі був професор Лоренс Клейн. Його початкова модель відома під назвою модель Клейна 1. Вона складається з трьох рівнянь та трьох тотожностей і має такий вигляд:

$$\text{функція споживання: } C = \beta_0 + \beta_1 P_t + \beta_2 (W + W')_t + \beta_3 P_{t-1} + \varepsilon_{1t};$$

$$\text{інвестиційна функція: } I_t = \beta_4 + \beta_5 P_t + \beta_6 P_{t-1} + \beta_7 K_{t-1} + \varepsilon_{2t};$$

попит на працю:

$$W_t = \beta_8 + \beta_9 (y + T - W')_t + \beta_{10} (y + T - W')_{t-1} + \beta_{11} t + \varepsilon_{3t};$$

$$\text{тотожність 1: } y_t + T_t = C_t + I_t + G_t;$$

$$\text{тотожність 2: } y_t = W'_t + W_t + P_t;$$

$$\text{тотожність 3: } K_t = K_{t-1} + I_t,$$

де C — витрати на споживання; I — інвестиційні витрати; G — витрати державного сектора; P — прибутки; W — заробітна плата в приватному секторі; W' — заробітна плата в державному секторі; K — запас капіталу; T — податки; y — доход після сплати податків; t — час; $\varepsilon_1, \varepsilon_2, \varepsilon_3$ — випадкові величини.

У моделі змінні C, I, W, y, P і K приймаються як взаємозалежні або ендогенні змінні, а змінні $P_{t-1}, K_{t-1}, y_{t-1}$ є попередньо визначеними. Всього для вивчення взаємозалежності шести ендогенних змінних, як уже зазначалося, використовується шість рівнянь (включаючи три тотожності).

Взаємозалежність між ендогенними змінними в моделі Клейна робить неможливим застосування методу найменших квадратів до окремого рівняння системи.

Нижче ми розглянемо, як оцінюються невідомі параметри економетричних моделей у такому разі. Поки що зазначимо, що у всіх розглянутих випадках виникають проблеми при застосуванні методу найменших квадратів, що призводить до оцінок з відхиленням (зміщених оцінок). Зміщення, що виникає при застосуванні класичного МНК до системи симульативних рівнянь, називається **зміщенням симульативних рівнянь**. Воно виникає через порушення припущення 4, тобто через залежність пояснюючих змінних та випадкової величини. Це створює певні проблеми. По-перше, виникає проблема ідентифікації параметрів індивідуальних взаємозв'язків. По-друге, виникають проблеми з оцінюванням. Застосування МНК призводить до зміщених та неконсистентних оцінок. Тому для системи симульативних рівнянь потрібно використовувати інші методи оцінювання. Які саме, ми розглянемо далі.

Але перед цим доведемо, що в симульативних рівняннях справді відбувається порушення класичного припущення про незалежність між факторами та випадковими величинами.

6.2. ПОРУШЕННЯ КЛАСИЧНОГО ПРИПУЩЕННЯ В СИМУЛЬАТИВНИХ МОДЕЛЯХ

Як уже зазначалося, коли між x та y існує взаємозалежність, то одним з її наслідків такої залежності буде порушення припущення 4 про незалежність значень факторів і випадкових величин ($\text{cov}(\varepsilon, x) \neq 0$), а через це оцінки невідомих параметрів такої моделі, отримані за методом найменших квадратів будуть зміщеними (оцінками з відхиленням). Розглянемо як ілюстрацію просту кейнсіанську модель доходу з попереднього прикладу 6.3:

$$C = \beta_0 + \beta_1 y + \varepsilon,$$

$$y = C + I.$$

де C — витрати на споживання; y — доход; I — інвестиції.

Припустимо, що ми хочемо оцінити параметри функції споживання (6.1). Нехай $E(\varepsilon_t) = 0$, $E(\varepsilon_t^2) = \sigma^2$, $E(\varepsilon_t \varepsilon_{t+j}) = 0$ (для $j \neq 0$) і $\text{cov}(I_t, \varepsilon_t) = 0$, що є припущеннями класичної лінійної регресійної моделі. Спочатку покажемо, що y_t і ε_t в (6.1) залежні, а потім доведемо, що b_1 є зміщеною оцінкою β_1 .

Моделю математично завершена: вона містить два рівняння з двома ендогенними змінними C та y . Інвестиції I визначаються екзогенно (наприклад, урядом). Для оцінки невідомих параметрів підставимо в друге рівняння вираз для C , отримаємо:

$$y = \beta_0 + \beta_1 y + I + \varepsilon,$$

або

$$y = \frac{\beta_0}{1 - \beta_1} + \frac{1}{1 - \beta_1} I + \frac{1}{1 - \beta_1} \varepsilon = \frac{\beta_0 + I + \varepsilon}{1 - \beta_1}. \quad (6.21)$$

Як бачимо, доход та випадкова величина залежні між собою. Доход не є справді екзогенною змінною у функції споживання.

Можна довести, що коваріація доходу y та ε не дорівнює 0, тобто $\text{cov}(y, \varepsilon) \neq 0$.

Доведення. За означенням коваріація ε та y

$$\text{cov}(\varepsilon, y) = E\{[\varepsilon - E(\varepsilon)][y - E(y)]\}.$$

Але $E(\varepsilon) = 0$. Тому $\text{cov}(\varepsilon, y) = E\{\varepsilon[y - E(y)]\}$.

Оскільки $y = \frac{\beta_0 + I + \varepsilon}{1 - \beta_1}$, а інвестиції визначаються екзогенно, отримаємо:

$$E(y) = \frac{\beta_0}{1 - \beta_1} + \frac{1}{1 - \beta_1} I. \quad (6.22)$$

Віднімемо від (6.21) рівняння (6.22) і отримаємо:

$$y_t - E(y_t) = \frac{\varepsilon_t}{1 - \beta_1}.$$

Беручи до уваги, що

$$\varepsilon_t - E(\varepsilon_t) = \varepsilon_t, \quad (6.23)$$

отримаємо:

$$\text{cov}(Y_t, \varepsilon_t) = E\{[y_t - E(y_t)][\varepsilon_t - E(\varepsilon_t)]\} = \frac{E(\varepsilon_t^2)}{1 - \beta_1} = \frac{\sigma^2}{1 - \beta_1} \neq 0. \quad (6.24)$$

Отже, ми показали, що при одночасній залежності змінних справді порушується припущення класичної регресії про незалежність між випадковою величиною ε та змінною y .

Внаслідок застосування методу найменших квадратів до симультаивних моделей *оцінки коефіцієнтів будуть зміщеними*.

Тобто другим наслідком одночасної залежності змінних буде зміщеність оцінок.

6.2.1. Отримання зміщених оцінок (оцінок з відхиленням) в моделях симультаивних рівнянь

Попередньо зазначалося, що не можна застосовувати метод найменших квадратів для оцінювання невідомих параметрів окремого рівняння, яке входить у систему симультаивних рівнянь. Таке оцінювання призведе до зміщених оцінок параметрів. Щоб довести цей факт, знову повернемося до простої кейнсіанської моделі визначення доходу з прикладу 6.3. Для спрощення розглянемо тільки оцінку параметра β_1 .

Щоб продемонструвати, що МНК-оцінка b_1 є оцінкою з відхиленням β_1 через взаємозалежність між y_t і ε_t , пригадаємо формулу для знаходження b_1 :

$$b_1 = \frac{\sum (C_t - \bar{C})(y_t - \bar{y})}{\sum (y_t - \bar{y})^2} = \frac{\sum \tilde{c}_t \tilde{y}_t}{\sum \tilde{y}_t^2}, \quad (6.25)$$

де маленькими літерами з хвилями позначені відхилення від середніх значень змінних моделі. Підставивши C_t з (6.1), отримуємо:

$$b_1 = \frac{\sum (\beta_0 + \beta_1 y_t + \varepsilon_t) \tilde{y}_t}{\sum \tilde{y}_t^2} = \beta_1 + \frac{\sum \tilde{y}_t \varepsilon_t}{\sum \tilde{y}_t^2}, \quad (6.26)$$

при цьому можна показати, що $\sum \tilde{y}_t = 0$ і $(\sum y_t \tilde{y}_t / \sum \tilde{y}_t^2) = 1$.

Якщо ми знайдемо математичне сподівання обох частин (6.26), то отримаємо:

$$E(b_1) = \beta_1 + E\left[\frac{\sum \tilde{y}_t \varepsilon_t}{\sum \tilde{y}_t^2}\right]. \quad (6.27)$$

На жаль, ми не можемо оцінити $E[\sum \tilde{y}_t \varepsilon_t / \sum \tilde{y}_t^2]$ (зазначте: $E = (A/B) \neq E = ((A)/(E(B)))$). Але інтуїтивно зрозуміло, що якщо вираз $[\sum \tilde{y}_t \varepsilon_t / \sum \tilde{y}_t^2]$ не є нулем, b_1 є зміщеною оцінкою β_1 . Чи правильно це? У (6.24) було пока-

зано, що співвідношення між y та ε не дорівнює нулю, але $\text{cov}(y, \varepsilon)$ є поняттям, яке ми використовуємо для всієї сукупності, а величина $\sum \bar{y}_t \varepsilon_t$ є її вибірковим аналогом, тому нам потрібно знайти границю величини $E[\sum \bar{y}_t \varepsilon_t / \sum \bar{y}_t^2]$, припускаючи, що розмір вибірки зростає до нескінченності. Кажуть, що оцінка зміщена, якщо її збіжність за ймовірністю не дорівнює правильному β_1 . Застосувавши правило збіжності за ймовірністю, отримаємо:

$$\begin{aligned} \text{збіжність за ймовірністю } p \lim(b_1) &= \text{збіжність за ймовірністю } (\beta_1) + \\ &+ \text{збіжність за ймовірністю } \left(\frac{\sum y_t \varepsilon_t}{\sum y_t^2} \right) = p \lim(\beta_1) + p \lim \left(\frac{\sum y_t \varepsilon_t / N}{\sum y_t^2 / N} \right) = \\ &= \beta_1 + p \lim \frac{(\sum y_t \varepsilon_t / N)}{(\sum y_t^2 / N)}, \end{aligned} \quad (6.28)$$

де чисельник у дробі є коваріацією між y і ε , а знаменник — дисперсією y .

Отже, з рівняння (6.28) отримаємо, що границя за ймовірністю b_1 дорівнює правильному β_1 плюс відношення границі за ймовірністю коваріації між y та ε до границі за ймовірністю дисперсії y у моделі. Тепер, коли розмір вибірки зростає до нескінченності, слід очікувати, що коваріація між вибірковими значеннями y і ε майже збігається з коваріацією всієї сукупності $E(y_t - E(y_t))(\varepsilon_t - E(\varepsilon_t))$, яка з рівняння (6.24) дорівнює $[\sigma^2 / (1 - \beta_1)]$. Одночасно, коли N має тенденцію зростати до нескінченності, дисперсія y у вибірці буде майже збігатися з дисперсією генеральної сукупності з σ_y^2 . Тому рівняння (6.28) можна записати як:

$$p \lim(b_1) = \beta_1 + \frac{\sigma^2 / (1 - \beta_1)}{\sigma_y^2} = \beta_1 + \frac{1}{1 - \beta_1} \left(\frac{\sigma^2}{\sigma_y^2} \right). \quad (6.29)$$

Знаючи, що для нашої моделі споживання $0 < \beta_1 < 1$ (оскільки b_1 в нашому прикладі — це МРС, що завжди є додатним, але меншим за 1) і що σ^2 та σ_y^2 додатні, з рівняння (6.29) видно, що гранична ймовірність (b_1) буде завжди більшою за β_1 ; це означає, що b_1 переоцінюватиме β_1 . Інакше кажучи, b_1 є зміщеною оцінкою (оцінкою з відхиленням) β_1 , і відхилення залишається навіть тоді, коли кількість спостережень прямує до нескінченності.

Отже, зміщення не можна усунути, збільшуючи кількість спостережень у вибірці. Оцінки, отримані за допомогою МНК, будуть не тільки зміщеними, а й неконсистентними.

6.2.2. Приклад обчислення оцінок з відхиленнями в симульативних моделях

Щоб продемонструвати деякі поняття, розглянуті вище, знову повернемося до простої кейнсіанської моделі визначення доходу з прикладу 6.3. Припустимо, що значення інвестицій I подано у таблиці 6.1 (колонка 3). Випадкові величини згенеруємо за допомогою *методу Монте-Карло*. Значення ε_t , створеного таким чином, подано у колонці (4). Далі припустимо, що

$$E(\varepsilon_t) = 0, E(\varepsilon_t, \varepsilon_{t+j}) = 0, (j \neq 0), \text{var}(\varepsilon_t) = \sigma^2 = 0.04, \text{cov}(\varepsilon_t, I_t) = 0.$$

Нехай у функції споживання (6.1) значення правильних параметрів відомі і дорівнюють відповідно $\beta_0 = 2$, $\beta_1 = 0.8$.

За допомогою значення β_0 і β_1 та виведеного значення ε_t можемо визначити значення доходу y_t з рівняння (6.1) у колонці (1) таблиці 6.1. Знаючи значення y_t , β_0 , β_1 та ε_t , можна легко знайти значення споживання C_t з рівняння (6.1). Показники C_t наведені у колонці (2).

Оскільки правильні β_0 і β_1 відомі, і оскільки помилки моделі знайдені випадково та незалежно від y , то якщо ми використаємо дані таблиці 6.1 для регресії C_t на y_t , маємо отримати $\beta_0 = 2$, $\beta_1 = 0.8$, за умови, що МНК-оцінки були без відхилення. Але якщо y_t і випадкова величина ε_t залежні, то відхилення наявне, рівняння (6.27) якраз і показує такий випадок. На основі даних табл. 6.1 розрахуємо коваріацію між y_t і ε_t : $\sum \bar{y}_t \varepsilon_t = 3.8$ та дисперсію: $\sum \bar{y}_t^2 = 184$. З рівняння (6.27) отримаємо вираз для коефіцієнта b_1 :

$$b_1 = \beta_1 + \frac{\sum \bar{y}_t \varepsilon_t}{\sum \bar{y}_t^2} = 0.8 + \frac{3.8}{184} = 0.82065. \quad (6.30)$$

Далі, використовуючи дані табл. 6.1, побудуємо регресійну модель залежності споживання (C_t) від доходу (y_t) і розрахуємо невідомі параметри методом МНК. Отримаємо:

$$\begin{aligned} \hat{C} &= 1.4940 + 0.82065y_t \\ &(0.35423)(0.01434) \\ &(4.2188)(57.209) \\ R^2 &= 0.9945. \end{aligned} \quad (6.31)$$

Як і очікувалося, оцінка величини β_1 точно передбачена рівнянням (6.30) і є оцінкою з відхиленням.

Таблиця 6.1

y_t (1)	C_t (2)	I_t (3)	ε_t (4)
18.15697	16.15697	2.000000	-0.3686055
19.59980	17.59980	2.000000	-0.8004084E-01
21.93468	19.73468	2.200000	0.1869357
21.55145	19.35145	2.200000	0.1102906
21.88427	19.48427	2.400000	-0.2314535E-01
22.42648	20.02648	2.400000	0.8529544E-01
25.40940	22.80940	2.600000	0.4818807
22.69523	20.09523	2.600000	-0.6095481E-01
24.36465	21.56465	2.800000	0.7292983E-01
24.39334	21.59334	2.800000	0.7866819E-01
24.09215	21.09215	3.000000	-0.1815703
24.87450	21.87450	3.000000	-0.2509900E-01
25.31580	22.11580	3.200000	-0.1368398
26.30465	23.10465	3.200000	0.6092946E-01
25.78235	22.38235	3.400000	-0.2435298
26.08018	22.68018	3.400000	-0.1839638
27.24440	23.64440	3.600000	-0.1511200
28.00963	24.40963	3.600000	0.1926739E-02
30.89301	27.09301	3.800000	0.3786015
28.98706	25.18706	3.800000	-0.2588852E-02

6.3. МЕТОДИ ОЦІНЮВАННЯ НЕВІДОМИХ ПАРАМЕТРІВ У МОДЕЛЯХ СИМУЛЬТАТИВНИХ РІВНЯНЬ

6.3.1. Загальні поняття про методи оцінювання

Як уже зазначалося, застосування методу найменших квадратів при оцінці невідомих параметрів у системах симультазивних рівнянь призводить до зміщених оцінок. Щоб уникнути цієї неприємної ситуації, необхідне застосування інших методів оцінювання, які давали б кращі оцінки параметрів. Такі методи, справді, є. Ось найуживаніші серед них.

1. *Метод зменшеної форми, або метод непрямих найменших квадратів (ННК).*
2. *Метод інструментальних змінних (МІЗ).*
3. *Двокроковий МНК (2МНК).*
4. *Метод найбільшої вірогідності обмеженої інформації (НВОІ).*
5. *Метод змішаного оцінювання.*
6. *Трикроковий МНК (3МНК).*
7. *Метод найбільшої вірогідності повної інформації (НВПІ).*

Перші 5 методів називають *методами одного рівняння*, оскільки вони застосовуються тільки до одного з рівнянь системи. Шостий та сьомий методи (3МНК та НВПІ) називають *системними методами*, оскільки

вони застосовуються одночасно до всіх рівнянь системи. Вибір методів для оцінювання параметрів конкретної моделі є непростим завданням. Перед його обговоренням потрібно розглянути декілька визначень і торкнутися проблеми ендогенності й екзогенності певних змінних у конкретних симультазивних моделях.

6.3.2. Попередні відомості про структурні моделі.

Ілюстративний приклад

Структурна модель — це повна система рівнянь, що описує структуру взаємозв'язків між економічними змінними. Структурні рівняння виражають ендогенні змінні як функції інших ендогенних змінних, попередньо визначених змінних та випадкових величин.

Щоб проілюструвати побудову структурної моделі, використаємо таку просту модель закритої економіки:

$$C_t = \alpha_0 + \alpha_1 y_t + \varepsilon_1;$$

$$I_t = \beta_0 + \beta_1 y_t + \beta_2 y_{t-1} + \varepsilon_2;$$

$$y_t = C_t + I_t + G_t.$$

Перше рівняння описує функцію споживання, друге — функцію інвестицій, а третє — основну тотожність. Система завершена, оскільки містить 3 рівняння з трьома ендогенними змінними — C_t, I_t, y_t . Модель містить дві попередньо визначені змінні — урядові витрати G та доход із запізненням y_{t-1} .

Параметри структурної моделі відображають *прямий вплив* кожного фактора на залежну змінну. За своїм економічним змістом вони можуть бути коефіцієнтами еластичності, граничними нормами та ін. Непрямі впливи можна обчислити тільки при розв'язуванні структурної системи. Наприклад, зміна в споживанні непрямо впливатиме на інвестиції через викликане змінами споживання зростання доходу y , який визначає інвестиції.

При побудові структурної моделі вважатимемо, що структурні параметри представлені коефіцієнтами β , коли вони відносяться до ендогенних змінних, та коефіцієнтами γ , коли вони стоять біля попередньо визначених змінних. Так само ендогенні змінні позначимо через y , а екзогенні змінні через x (для спрощення будемо розглядати модель без перетину). Тоді структурна модель матиме вигляд:

$$y_1 = \beta_{13} y_3 + \varepsilon_1;$$

$$y_2 = \beta_{23} y_3 + \gamma_{21} x_1 + \varepsilon_2;$$

$$y_3 = y_1 + y_2 + x_2.$$

де $y_1 = C$, $y_2 = I$, $y_3 = y$, $x_1 = y_{t-1}$, $x_2 = G$.

Переносючи всі змінні у ліву сторону, отримуємо таку систему:

$$\begin{aligned} y_1 + 0y_2 - \beta_{13}y_3 + 0x_1 + 0x_2 &= \varepsilon_1; \\ 0y_1 + y_2 - \beta_{23}y_3 - \gamma_{21}x_1 + 0x_2 &= \varepsilon_2; \\ -y_1 - y_2 + y_3 + 0x_1 - x_2 &= 0. \end{aligned}$$

З цієї системи легко отримати таблицю структурних параметрів у загальному вигляді:

$$\begin{pmatrix} 1 & 0 & -\beta_{13} & 0 & 0 \\ 0 & 1 & -\beta_{23} & -\gamma_{21} & 0 \\ -1 & -1 & 1 & 0 & -1 \end{pmatrix}.$$

Значення структурних параметрів можна розрахувати, використовуючи спостереження над змінними моделі та застосовуючи відповідні економетричні методи. Як це робити, ми розглянемо пізніше.

6.3.3. Структурні моделі скороченої форми

Скороченою формою структурної моделі є модель, у якій ендогенні змінні виражені як функції лише попередньо визначених змінних. Скорочена форма записується двома способами. Перший — згорнуте вираження ендогенних змінних як функцій попередньо визначених змінних, наприклад:

$$y_i = \pi_{i1}x_1 + \pi_{i2}x_2 + \dots + \pi_{ik}x_k + v_i, \quad (i = 1, 2, \dots, G).$$

Для нашої простої моделі з трьох рівнянь скорочена форма матиме вигляд:

$$\begin{aligned} C_t &= \pi_{11}y_{t-1} + \pi_{12}G_t + v_1; \\ I_t &= \pi_{21}y_{t-1} + \pi_{22}G_t + v_2; \\ y_t &= \pi_{31}y_{t-1} + \pi_{32}G_t + v_3. \end{aligned}$$

Другий спосіб запису скороченої форми — розгорнуте вираження ендогенних змінних через попередньо визначені змінні, структурні параметри та випадкові величини. У цьому випадку структурна модель нашого прикладу набуває такого вигляду:

$$\begin{aligned} C_t &= \frac{\alpha_1\beta_2}{1-\alpha_1-\beta_1}y_{t-1} + \frac{\alpha_1}{1-\alpha_1-\beta_1}G_t + \frac{\varepsilon_1 + \alpha_1\varepsilon_2 - \beta_1\varepsilon_1}{1-\alpha_1-\beta_1}; \\ I_t &= \frac{\beta_2(1-\alpha_1)}{1-\alpha_1-\beta_1}y_{t-1} + \frac{\beta_1}{1-\alpha_1-\beta_1}G_t + \frac{\varepsilon_2 + \beta_1\varepsilon_1 - \alpha_1\varepsilon_2}{1-\alpha_1-\beta_1}; \end{aligned}$$

$$y_t = \frac{\beta_2}{1-\alpha_1-\beta_1}y_{t-1} + \frac{1}{1-\alpha_1-\beta_1}G_t + \frac{\varepsilon_1 + \varepsilon_2}{1-\alpha_1-\beta_1}.$$

Для збігу двох типів запису скороченої форми необхідно, щоб виконувалось таке співвідношення між числами π та структурними параметрами:

$$\begin{aligned} \pi_{11} &= \frac{\alpha_1\beta_2}{1-\alpha_1-\beta_1}; & \pi_{12} &= \frac{\alpha_1}{1-\alpha_1-\beta_1}; \\ \pi_{21} &= \frac{\beta_2(1-\alpha_1)}{1-\alpha_1-\beta_1}; & \pi_{22} &= \frac{\beta_1}{1-\alpha_1-\beta_1}; \\ \pi_{31} &= \frac{\beta_2}{1-\alpha_1-\beta_1}; & \pi_{32} &= \frac{1}{1-\alpha_1-\beta_1}. \end{aligned}$$

Як бачимо, між параметрами скороченої форми та структурними параметрами є чіткий взаємозв'язок, тобто значення π є функціями структурних параметрів.

6.3.4. Знаходження параметрів скороченої форми та оцінка їхнього впливу на змінні моделі

Тепер детально розглянемо, як усе ж таки знаходяться невідомі параметри скороченої форми. Для цього знову повернемося до нашого прикладу та побудуємо скорочену форму структурної моделі. Для її отримання спочатку в третє структурне рівняння підставимо замість C_t та I_t їхні вирази з перших двох рівнянь:

$$y_t = (\alpha_1y_t + \varepsilon_1) + (\beta_1y_t + \beta_2y_{t-1} + \varepsilon_2) + G_t.$$

За допомогою перестановки та спрощення остаточно матимемо:

$$y_t = \frac{\beta_2}{1-\alpha_1-\beta_1}y_{t-1} + \frac{1}{1-\alpha_1-\beta_1}G_t + \frac{\varepsilon_1 + \varepsilon_2}{1-\alpha_1-\beta_1}.$$

Це і є скорочена форма третього структурного рівняння.

Тепер підставимо отриманий вираз для y_t у функцію споживання:

$$C_t = \alpha_1 \left[\frac{\beta_2}{1-\alpha_1-\beta_1}y_{t-1} + \frac{1}{1-\alpha_1-\beta_1}G_t + \frac{\varepsilon_1 + \varepsilon_2}{1-\alpha_1-\beta_1} \right] + \varepsilon_1;$$

або

$$C_t = \frac{\alpha_1 \beta_2}{1 - \alpha_1 - \beta_1} y_{t-1} + \frac{\alpha_1}{1 - \alpha_1 - \beta_1} G_t + \frac{\varepsilon_1 + \alpha_1 \varepsilon_2 - \beta_1 \varepsilon_1}{1 - \alpha_1 - \beta_1}.$$

Отримаємо скорочену форму для функції споживання. Далі підставимо вираз для y_t у функцію інвестицій:

$$I_t = \beta_1 \left[\frac{\beta_2}{1 - \alpha_1 - \beta_1} y_{t-1} + \frac{1}{1 - \alpha_1 - \beta_1} G_t + \frac{\varepsilon_1 + \varepsilon_2}{1 - \alpha_1 - \beta_1} \right] + \beta_2 y_{t-1} + \varepsilon_2$$

або

$$I_t = \frac{\beta_2(1 - \alpha_1)}{1 - \alpha_1 - \beta_1} y_{t-1} + \frac{\beta_1}{1 - \alpha_1 - \beta_1} G_t + \left(\frac{\varepsilon_2 + \beta_1 \varepsilon_1 - \alpha_1 \varepsilon_2}{1 - \alpha_1 - \beta_1} \right).$$

Отримаємо відповідно скорочену форму для функції інвестицій.

Параметри скороченої форми вимірюють загальний (прямий та непрямий) вплив попередньо визначених змінних на ендогенні змінні, в той час як структурні параметри вимірюють тільки прямий вплив. Наприклад, π_{21} вимірює вплив одиничного зростання y_{t-1} на величину інвестицій. Цей вплив складається з двох частин: по-перше, є прямий вплив на I через коефіцієнт β_2 , який визначається в структурному рівнянні інвестицій; по-друге, наявний непрямий вплив, тому що зростання впливає на I_t , I_t впливає на y_t , яке в свою чергу впливає на I_t ; нарешті значення y_t впливає на C_t , яке в свою чергу впливає на y_t , а отже, і на I_t . Тому загальний вплив y_{t-1} (що вимірюється π_{21}) на I_t можна поділити на такі компоненти:

$$\pi_{21} = \frac{\beta_2(1 - \alpha_1)}{1 - \alpha_1 - \beta_1} = \beta_2 \left(1 + \frac{\beta_1}{1 - \alpha_1 - \beta_1} \right)$$

або

$$\pi_{21} = \beta_2 + \frac{\beta_2 \beta_1}{1 - \alpha_1 - \beta_1};$$

$$\begin{bmatrix} \text{загальний} \\ \text{вплив} \end{bmatrix} = \begin{bmatrix} \text{прямий} \\ \text{вплив} \end{bmatrix} + \begin{bmatrix} \text{непрямий} \\ \text{вплив} \end{bmatrix}.$$

Параметри скороченої форми таким чином можна широко застосовувати для прогнозування та аналізу економічної діяльності, тому що вони дають одночасно оцінку загального, прямого та непрямого впливу екзогенних змінних на залежні змінні.

Як же оцінити параметри скороченої форми? Є два методи отримання оцінок для коефіцієнтів моделі скороченої форми: пряме та непряме оцінювання.

Пряме оцінювання коефіцієнтів скороченої форми. Коефіцієнт ρ можна оцінити методом **необмежених найменших квадратів** (НеНК). У цьому разі всі ендогенні змінні виражаються як функції попередньо визначених змінних системи, і до цих функцій скороченої форми застосовується звичайний МНК. Даний спосіб називається методом необмежених найменших квадратів, оскільки не береться до уваги ніяка інформація про структурні параметри, тобто усуваються обмеження, накладені формою структурної моделі. Даний метод не вимагає знання всіх особливостей структурної моделі, але вимагає знання попередньо визначених змінних.

Непряме оцінювання коефіцієнтів скороченої форми. Як ми вже показували вище, між коефіцієнтами скороченої форми та структурними параметрами є чіткий взаємозв'язок. Тому можна спочатку отримати оцінки структурних параметрів за допомогою будь-якого методу оцінювання, а потім підставити їх у **систему параметричних взаємозв'язків** для отримання значень π . Цей непрямий метод включає три кроки. На першому кроці за допомогою підстановок змінних необхідно отримати скорочену форму для всіх рівнянь моделі, щоб мати **систему параметричних взаємозв'язків** між π , β та γ . На другому кроці отримуємо оцінки структурних параметрів за допомогою будь-якого відомого методу оцінювання. Нарешті, на третьому кроці для отримання оцінок коефіцієнтів скороченої форми підставляємо оцінки β і γ в систему параметричних взаємозв'язків.

Описаний метод складніший, але має ряд переваг перед прямим методом оцінювання.

По-перше, отримання величин π скороченої форми з β та γ ефективніше, оскільки береться до уваги вся інформація (тобто всі початкові обмеження, накладені на параметри), що міститься в структурній моделі. По-друге, структурні зміни відбуваються постійно протягом часу. Якщо ми знаємо інформацію про зміни β і γ , можна легко обчислити зміни π . Якщо ж π обчислювати за допомогою необмеженого методу найменших квадратів, то ми не можемо пов'язати з π зміни, що відбудуться з β та γ . По-третє, з інших досліджень можна отримати додаткову інформацію про деякі структурні параметри, яка стає непотрібною, якщо оцінки для π отримані не через оцінки β і γ .

6.4. ПРОБЛЕМА ОТОТОЖНЕННЯ В СИМУЛЬТАТИВНИХ МОДЕЛЯХ. ЗАГАЛЬНІ ПОНЯТТЯ

Як уже зазначалося, в симультативній моделі є змінні двох типів: ендогенні та попередньо визначені. Ендогенні змінні вважаються стохастичними, тоді як попередньо визначені змінні трактуються як нестохастичні.

Попередньо визначені змінні поділяються на дві категорії: поточні та лагові. Так, наприклад, якщо x_1 є поточною екзогенною змінною, то $x_{1,t-1}$ вважається лаговою екзогенною змінною з одиничним лагом. Якщо y_t є ендогенною змінною, то y_{t-1} — лагова змінна, значення якої відоме в поточний період часу t , отже це значення вважається нестохастичним, а y_{t-1} є також попередньо визначеною змінною. Право визначати, які змінні ендогенні, а які попередньо визначені, належить досліднику, котрий розробляє модель. Виходячи з наведених уточнень, запишемо симультативну модель у загальному випадку:

$$\begin{aligned} y_{1t} &= \beta_{12}y_{2t} + \dots + \beta_{1M}y_{Mt} + \gamma_{11}x_{1t} + \dots + \gamma_{1K}x_{Kt} + \varepsilon_{1t}; \\ y_{2t} &= \beta_{21}y_{1t} + \dots + \beta_{2M}y_{Mt} + \gamma_{21}x_{1t} + \dots + \gamma_{2K}x_{Kt} + \varepsilon_{2t}; \\ &\dots\dots\dots \\ y_{Mt} &= \beta_{M1}y_{1t} + \dots + \beta_{M,M-1}y_{M-1,t} + \gamma_{M1}x_{1t} + \dots + \gamma_{MK}x_{Kt} + \varepsilon_{Mt}, \end{aligned} \quad (6.32)$$

де y_1, y_2, \dots, y_M — ендогенні, або залежні змінні;

x_1, x_2, \dots, x_K — попередньо визначені змінні (одна з цих x -змінних може набувати значення одиниці для отримання перетину в кожному рівнянні),

$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_M$ — випадкові величини;

$t = 1, 2, \dots, N$ — загальна кількість спостережень;

β — невідомі параметри при ендогенних змінних;

γ — невідомі параметри при попередньо визначених змінних.

Параметри β та γ називаються структурними параметрами чи коефіцієнтами.

Зауважимо, що не всі змінні обов'язково мають з'являтися у кожному рівнянні.

З симультативною (структурною) моделі, як ми вже розглядали вище, можна отримати скорочену форму, в якій ендогенні змінні залежать тільки від попередньо визначених змінних та випадкових величин.

Під проблемою оцінювання параметрів симультативних моделей розуміють знаходження оцінок параметрів на основі оцінених коефіцієнтів скороченої форми. Якщо це можна зробити, то ми маємо право стверджувати, що модель ототожнена. І навпаки.

Ототожнена модель може бути як точно ототожненою, так і переототожненою. *Точно ототожнену модель ми маємо в тому разі, коли можна отримати однозначну оцінку її параметрів. Переототожнену модель*

ми маємо у разі, коли для деяких параметрів структурної моделі є можливість отримати більше ніж одне кількісне значення. Обставини, за яких трапляються обидва ці випадки, будуть розглянуті далі.

Крім того, модель може бути і неототожненою. Розглянемо *проблему неототожнення* більш детально. Для прикладу візьмемо модель попиту та пропозиції.

Функція попиту:

$$Q_t^d = \alpha_0 + \alpha_1 P_t + \varepsilon_{1t}, \quad \alpha_1 > 0. \quad (6.39)$$

Функція пропозиції:

$$Q_t^s = \beta_0 + \beta_1 P_t + \varepsilon_{2t}, \quad \beta_1 > 0. \quad (6.40)$$

Умова рівноваги:

$$Q_t^d = Q_t^s. \quad (6.41)$$

За умови рівноваги, коли попит дорівнює пропозиції, отримаємо:

$$\alpha_0 + \alpha_1 P_t + \varepsilon_{1t} = \beta_0 + \beta_1 P_t + \varepsilon_{2t}. \quad (6.42)$$

Враховуючи умову рівноваги (6.41), знайдемо ціну рівноваги:

$$P_t = \pi_0 + v_t, \quad (6.43)$$

де

$$\pi_0 = \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1}; \quad (6.44)$$

$$v_t = \frac{\varepsilon_{2t} - \varepsilon_{1t}}{\alpha_1 - \beta_1}. \quad (6.45)$$

Шляхом підстановки P_t з (6.43) в (6.39) обчислимо попит за умов рівноваги:

$$Q_t = \pi_1 + \omega_t, \quad (6.46)$$

де

$$\pi_1 = \frac{\alpha_1 \beta_0 - \alpha_0 \beta_1}{\alpha_1 - \beta_1}; \quad (6.47)$$

$$\omega_t = \frac{\alpha_1 \varepsilon_{2t} - \beta_1 \varepsilon_{1t}}{\alpha_1 - \beta_1}. \quad (6.48)$$

Помилки v_t і ω_t є лінійними комбінаціями початкових випадкових величин ε_1 та ε_2 . Рівняння (6.43) та (6.46) створюють скорочену форму си-

мультативної моделі попиту та пропозиції, яка має чотири структурних параметри: α_0 , α_1 , β_0 та β_1 . Слід відмітити, що єдиного способу оцінювання їх немає. Чому? Подивимось на формули розрахунку двох коефіцієнтів скороченої форми, наведених у (6.44) та (6.47). Ці коефіцієнти містять усі чотири структурні параметри, але немає способу, за допомогою якого можна було б оцінити чотири невідомих, виходячи лише з двох рівнянь. З шкільного курсу алгебри відомо, що для оцінки чотирьох невідомих має бути чотири незалежні рівняння, і в загальній формі для оцінки k невідомих необхідно k незалежних рівнянь. Наша структурна модель не ототожнена, тобто ми не можемо оцінити її невідомі параметри.

Є й альтернативний спосіб розгляду проблеми ототожнення. Щоб проілюструвати його, повернемося знову до нашого прикладу. Помножимо (6.39) на λ ($0 \leq \lambda \leq 1$) і (6.40) на $(1 - \lambda)$, отримаємо такі рівняння:

$$\lambda Q_t = \lambda \alpha_0 + \lambda \alpha_1 P_t + \lambda \varepsilon_{1t}; \quad (6.49)$$

$$(1 - \lambda) Q_t = (1 - \lambda) \beta_0 + (1 - \lambda) \beta_1 P_t + (1 - \lambda) \varepsilon_{2t}. \quad (6.50)$$

Лінійна комбінація цих рівнянь дорівнює:

$$Q_t = \gamma_0 + \gamma_1 P_t + w_t; \quad (6.51)$$

де

$$\begin{aligned} \gamma_0 &= \lambda \alpha_0 + (1 - \lambda) \beta_0; \\ \gamma_1 &= \lambda \alpha_1 + (1 - \lambda) \beta_1; \\ w_t &= \lambda \varepsilon_{1t} + (1 - \lambda) \varepsilon_{2t}. \end{aligned} \quad (6.52)$$

Змішане рівняння (6.51) наочно не визначене ні з (6.39), ні з (6.40). Тому, коли ми маємо дані тільки про P і Q , будь-яке з рівнянь (6.39), (6.40) і (6.51) може бути прийнятним як для опису функції попиту, так і для функції пропозиції. Для того, щоб ототожнити рівняння, тобто оцінити його параметри, наприклад функцію попиту, потрібно показати, що ми володіємо інформацією, яка дозволяє оцінити саме функцію пропозиції чи змішане рівняння.

6.5. ПРАВИЛЬНЕ, АБО ТОЧНЕ ОТOTOЖНЕННЯ

Причина, через яку ми не могли ототожнити попередні функції попиту та пропозиції, полягала в тому, що ті ж самі змінні P і Q є в обох функціях, а додаткової інформації немає. Припустимо, що ми розглядаємо таку модель попиту та пропозиції.

Функція попиту:

$$Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 D_t + \varepsilon_{1t}, \quad \alpha_1 < 0, \alpha_2 > 0. \quad (6.53)$$

Функція пропозиції:

$$Q_t = \beta_0 + \beta_1 P_t + \varepsilon_{2t}, \quad \beta_1 > 0, \quad (6.54)$$

де D — доход споживача, екзогенна змінна.

Зазначимо, що єдина відмінність між попередньою і нашою поточною моделлю попиту та пропозиції полягає в додатковій змінній D (доход) у функції попиту. З економічної теорії попиту нам відомо, що доход є важливою детермінантою попиту на більшість товарів та послуг. Отже, включення її у функцію попиту дасть нам додаткову інформацію про поведінку споживача. Для більшості товарів доход має позитивний вплив на споживання ($\alpha_2 > 0$).

Використовуючи ринковий механізм рівноваги, коли попит дорівнює пропозиції, маємо:

$$\alpha_0 + \alpha_1 P_t + \alpha_2 D_t + \varepsilon_{1t} = \beta_0 + \beta_1 P_t + \varepsilon_{2t}. \quad (6.55)$$

Розв'язання рівняння (6.55) дасть нам рівноважне значення P_t , яке становить:

$$P_t = \pi_0 + \pi_1 D_t + v_t, \quad (6.56)$$

де коефіцієнти скороченої форми знаходяться за формулами:

$$\begin{aligned} \pi_0 &= \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1}, \\ \pi_1 &= -\frac{\alpha_2}{\alpha_1 - \beta_1}, \\ v_t &= \frac{\varepsilon_{2t} - \varepsilon_{1t}}{\alpha_1 - \beta_1}. \end{aligned} \quad (6.57)$$

Підставивши рівноважне значення P_t в попередню функцію попиту, отримаємо значення Q_t за умови рівноваги:

$$Q_t = \pi_2 + \pi_3 D_t + w_t, \quad (6.58)$$

де

$$\begin{aligned} \pi_2 &= \frac{\alpha_1 \beta_0 - \alpha_0 \beta_1}{\alpha_1 - \beta_1}, \\ \pi_3 &= \frac{-\alpha_2 \beta_1}{\alpha_1 - \beta_1}, \\ w_t &= \frac{\alpha_1 \varepsilon_{2t} - \beta_1 \varepsilon_{1t}}{\alpha_1 - \beta_1}. \end{aligned} \quad (6.59)$$

Оскільки (6.56) і (6.58) є скороченою формою симультавної моделі, для оцінки їхніх параметрів можна застосувати метод найменших квадратів.

Тепер модель попиту та пропозиції містить п'ять структурних параметрів $\alpha_0, \alpha_1, \beta_0, \beta_1$ і β_2 . Але для оцінки їх є лише чотири рівняння із скороченими параметрами π_0, π_1, π_2 і π_3 , наведеними в (6.57) і (6.59). Отже, всі структурні коефіцієнти визначити неможливо. Але при цьому можна оцінити параметри функції пропозиції, тому що:

$$\begin{aligned} \beta_0 &= \pi_2 - \beta_1 \pi_0, \\ \beta_1 &= \frac{\pi_3}{\pi_1}, \end{aligned} \quad (6.60)$$

на відміну від функції попиту, для якої немає способу оцінки невідомих параметрів, отже, вона залишається неототоженою.

Щоб пересвідчитись, що функція попиту (6.53) не може бути ототоженою, помножимо її на λ ($0 < \lambda \leq 1$) та (6.54) та $(1 - \lambda)$ і, додавши їх, отримаємо таке рівняння:

$$Q_t = \gamma_0 + \gamma_1 P_t + \gamma_2 D_t + w_t; \quad (6.61)$$

де

$$\begin{aligned} \gamma_0 &= \lambda \alpha_0 + (1 - \lambda) \beta_0; \\ \gamma_1 &= \lambda \alpha_1 + (1 - \lambda) \beta_1; \\ \gamma_2 &= \lambda \alpha_2; \\ w_t &= \lambda \varepsilon_{1t} + (1 - \lambda) \varepsilon_{2t}. \end{aligned} \quad (6.62)$$

Функція (6.61) не відрізняється від функції попиту (6.53), але відрізняється від функції пропозиції (6.54), яка не містить пояснючої змінної D . Отже, функція попиту залишається неототоженою.

Визначимо цікавий факт: додаткова змінна у функції попиту робить можливим ототожнення функції пропозиції.

Тепер припустимо, що ми розглядаємо таку модель попиту та пропозиції.

Функція попиту:

$$Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 D_t + \varepsilon_{1t}; \quad \alpha_1 < 0, \alpha_2 > 0. \quad (6.63)$$

Функція пропозиції:

$$Q_t = \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + \varepsilon_{2t}; \quad \beta_1 > 0, \beta_2 > 0, \quad (6.64)$$

де функція попиту залишається тією ж самою, але функція пропозиції містить додаткову пояснювальну змінну, ціну попереднього періоду.

Функція пропозиції показує, що кількість запропонованого товару залежить від поточних цін і від цін у попередньому періоді. Така модель часто використовується для пояснення пропозиції сільськогосподарських товарів. Змінна P_{t-1} — попередньо визначена, бо її значення відоме в період часу t .

За умови рівноваги маємо:

$$\alpha_0 + \alpha_1 P_t + \alpha_2 D_t + \varepsilon_{1t} = \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + \varepsilon_{2t}. \quad (6.65)$$

Розв'язавши рівняння відносно P_t , отримаємо:

$$P_t = \pi_0 + \pi_1 D_t + \pi_2 P_{t-1} + v_t, \quad (6.66)$$

де

$$\begin{aligned} \pi_0 &= \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1}, \\ \pi_1 &= -\frac{\alpha_2}{\alpha_1 - \beta_1}, \\ \pi_2 &= \frac{\beta_2}{\alpha_1 - \beta_1}, \\ v_t &= \frac{\varepsilon_{2t} - \varepsilon_{1t}}{\alpha_1 - \beta_1}. \end{aligned} \quad (6.67)$$

Підставивши рівноважну ціну в рівняння попиту чи пропозиції, отримаємо відповідне рівноважне значення Q_t :

$$Q_t = \pi_3 + \pi_4 I_t + \pi_5 P_{t-1} + w_t, \quad (6.68)$$

де коефіцієнти скороченої форми дорівнюють:

$$\begin{aligned}
 \pi_3 &= \frac{\alpha_1\beta_0 - \alpha_0\beta_1}{\alpha_1 - \beta_1}, \\
 \pi_4 &= -\frac{\alpha_2\beta_1}{\alpha_1 - \beta_1}, \\
 \pi_5 &= \frac{\alpha_1\beta_2}{\alpha_1 - \beta_1}, \\
 w_t &= \frac{\alpha_1\varepsilon_{2t} - \beta_1\varepsilon_{1t}}{\alpha_1 - \beta_1}.
 \end{aligned}
 \tag{6.69}$$

Модель попиту та пропозиції (6.53) і (6.64) містить шість структурних параметрів $\alpha_0, \alpha_1, \alpha_2, \beta_0, \beta_1$ і β_2 ; є також шість коефіцієнтів скороченої форми $\pi_0, \pi_1, \pi_2, \pi_3, \pi_4$ і π_5 , щоб їх оцінити. Отже, маємо шість рівнянь з шістьма невідомими і можемо знайти єдині оцінки. Обидва рівняння будуть ототожненими.

Для перевірки можна помножити рівняння попиту (6.53) на λ ($0 \leq \lambda \leq 1$), рівняння пропозиції — на $(1 - \lambda)$ та додати їх, щоб отримати змішане рівняння. Це рівняння міститиме обидві попередньо визначені змінні D_t та P_{t-1} .

6.6. ПРОБЛЕМА ПЕРЕОТOTOЖНЕННЯ

Для певних товарів і послуг доход та багатство споживача є важливою детермінантою попиту. Тому змінимо функцію попиту (6.53), залишивши функцію пропозиції без змін.

Функція попиту:

$$Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 D_t + \alpha_3 R_t + \varepsilon_{1t}; \quad \alpha_1 < 0, \alpha_2 > 0. \tag{6.70}$$

Функція пропозиції:

$$Q_t = \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + \varepsilon_{2t}; \quad \beta_1 > 0, \beta_2 > 0, \tag{6.71}$$

де R — багатство; для більшості товарів та послуг багатство, як і доход, має позитивний ефект на споживання.

Прирівнявши попит і пропозицію, отримаємо Q_t та P_t за умов рівноваги:

$$P_t = \pi_0 + \pi_1 I_t + \pi_2 R_t + \pi_3 P_{t-1} + v_t; \tag{6.72}$$

$$Q_t = \pi_4 + \pi_5 I_t + \pi_6 R_t + \pi_7 P_{t-1} + w_t, \tag{6.72}$$

де

$$\pi_0 = \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1}, \quad \pi_1 = -\frac{\alpha_2}{\alpha_1 - \beta_1},$$

$$\begin{aligned}
 \pi_2 &= -\frac{\alpha_3}{\alpha_1 - \beta_1}, & \pi_3 &= \frac{\beta_2}{\alpha_1 - \beta_1}, \\
 \pi_4 &= \frac{\alpha_1\beta_0 - \alpha_0\beta_1}{\alpha_1 - \beta_1}, & \pi_5 &= -\frac{\alpha_2\beta_1}{\alpha_1 - \beta_1}, \\
 \pi_6 &= -\frac{\alpha_3\beta_1}{\alpha_1 - \beta_1}, & \pi_7 &= \frac{\alpha_2\beta_2}{\alpha_1 - \beta_1}, \\
 v_t &= \frac{\varepsilon_{2t} - \varepsilon_{1t}}{\alpha_1 - \beta_1}, & w_t &= \frac{\alpha_1\varepsilon_{2t} - \beta_1\varepsilon_{1t}}{\alpha_1 - \beta_1}.
 \end{aligned}
 \tag{6.73}$$

Тепер модель попиту та пропозиції містить сім структурних параметрів, а для оцінки їх є вісім рівнянь, тобто вісім коефіцієнтів скороченої форми, даних в (6.73). У даному разі кількість рівнянь більша за кількість невідомих. У результаті єдина оцінка всіх параметрів моделі неможлива. Наприклад:

$$\beta_1 = \frac{\pi_6}{\pi_2}, \tag{6.74}$$

$$\beta_1 = \frac{\pi_5}{\pi_1}, \tag{6.75}$$

тобто для функції пропозиції є дві оцінки параметра β_1 , але немає гарантій, що вони будуть ідентичні. Більше того, оскільки β_1 з'являється у знаменниках коефіцієнтів скороченої форми, двозначність оцінки β_1 може бути також перенесена на інші оцінки.

Чому функція пропозиції була оцінена в моделях (6.53) і (6.64), а в моделі (6.70) переоцінена, незважаючи на те, що в обох випадках вона залишається тією ж самою? Тому що в даному разі ми маємо занадто багато інформації, щоб оцінити криву пропозиції. Це протилежний випадок до неототожнення, коли нам не вистачало інформації.

Підсумувавши розглянуту проблему ототожнення, можна зробити висновок, що **симульативна модель може бути неототожненою, точно ототожненою та переототожненою**. Задамо собі запитання, чи є якісь загальні умови, котрі допомагають виявити всі ці випадки. Відповідь на це запитання ми дамо в наступному параграфі.

6.7. ОСНОВНІ ПРАВИЛА ОТOTOЖНЕННЯ

Розглянемо основні правила ототожнення симультазивних моделей. Введемо таку систему позначень:

M — кількість ендогенних змінних у симультазивній моделі;

m — кількість ендогенних змінних у окремому рівнянні;

K — кількість попередньо визначених змінних у моделі;

k — кількість попередньо визначених змінних у окремому рівнянні.

З врахуванням введеної системи позначень сформулюємо обов'язкову (але не достатню) умову ототожнення, яка має назву "умова порядку" і може бути визначена двома різними, але еквівалентними способами.

Визначення 1. Для ототожнення рівняння в ньому має бути опущено щонайменше $M-1$ змінних, які з'являються в цілому в моделі. Якщо опущено рівно $M-1$ змінних, рівняння буде ототожненим. Якщо опущено більше, ніж $M-1$ змінних, воно буде перетотожненим.

Визначення 2. Для ототожнення рівняння число попередньо визначених змінних, опущених в ньому, має бути не меншим за число включених в нього ендогенних змінних мінус одиниця, тобто

$$K - k \geq m - 1. \quad (6.76)$$

Якщо $K - k = m - 1$, — рівняння ототожнене, але якщо $K - k > m - 1$, — воно перетотожнене.

Щоб проілюструвати умову порядку, звернемося до попередніх прикладів.

Приклад 6.1

Функція попиту:

$$Q_t = \alpha_0 + \alpha_1 P_t + \varepsilon_{1t}.$$

Функція пропозиції:

$$Q_t = \beta_0 + \beta_1 P_t + \varepsilon_{2t}.$$

Ця модель має дві ендогенні змінні P та Q і жодної попередньо визначеної. Для ототожнення в кожному рівнянні має бути опущена щонайменше $M - 1 = 2 - 1 = 1$ змінна. В даному разі жодне з рівнянь не буде ототожненим.

Приклад 6.2

Функція попиту:

$$Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 D_t + \varepsilon_{1t}.$$

Функція пропозиції:

$$Q_t = \beta_0 + \beta_1 P_t + \varepsilon_{2t}.$$

У даній моделі Q і P ендогенні, а D — екзогенна. Застосовуючи умову

порядку (6.76), бачимо, що функція попиту неототожнена. З іншого боку, функція пропозиції ототожнена, бо в ній опущено рівно одну змінну $D_t (M - 1 = 2 - 1 = 1)$.

Приклад 6.3

Функція попиту:

$$Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 D_t + \varepsilon_{1t}.$$

Функція пропозиції:

$$Q_t = \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + \varepsilon_{2t}.$$

P_t і Q_t — ендогенні, а D_t і P_{t-1} — попередньо визначені. У першому рівнянні опущено рівно одну змінну P_{t-1} , у другому рівнянні також опущено рівно одну змінну D_t . Кожне рівняння може бути ототожнене за умовою порядку, а отже, і модель в цілому також може бути ототожненою.

Приклад 6.4

Функція попиту:

$$Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 D_t + \alpha_3 R_t + \varepsilon_{1t}.$$

Функція пропозиції:

$$Q_t = \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + \varepsilon_{2t}.$$

У даній моделі P_t і Q_t — ендогенні, а D_t , R_t і P_{t-1} — попередньо визначені. У функції попиту опущено рівно одну змінну P_{t-1} , за умовою порядку вона точно ототожнена. А у функції пропозиції опущено дві змінні D_t та R_t , отже, вона перетотожнена. Як зазначалось раніше, в даному разі є два способи оцінки β_1 .

Як показують попередні приклади, ототожнення рівнянь симультазивних моделей можливе тоді, коли в окремих рівняннях опущено одну чи більше змінних, які є ще де-небудь у моделі.

Рангова умова ототожнення

Умова порядку, яка обговорювалась раніше, є обов'язковою, але не достатньою умовою ототожнення. Тобто може статися так, що навіть якщо умова порядку $K - k \geq m - 1$ виконана, рівняння може бути неототожненим, тому що попередньо визначені змінні, які опущено в ньому, але є в моделі, можуть бути залежними. Через це відповідність між структурними коефіцієнтами (β) і коефіцієнтами скороченої форми (π) не зберігається. Тобто ми не можемо оцінити структурні параметри за коефіцієнтами скороченої форми. Тому потрібно мати як достатню, так і необхідну умову ототожнення. Такою умовою є рангова умова ототожнення, яка формулюється таким чином.

Рангова умова ототожнення. В симультаивній моделі, яка містить M рівнянь з M ендогенними змінними, рівняння буде ототожненим тоді і тільки тоді, коли ранг матриці, утворений з коефіцієнтів, котрі відповідають опущеним змінним рівняння, що розглядається, у всіх інших рівняннях моделі, крім даного, дорівнює $M-1$.

Як ілюстрацію рангової умови ототожнення розглянемо таку гіпотетичну систему симультаивних рівнянь, у якій y — ендогенні змінні, а x — попередньо визначені:

$$y_{1t} - \beta_{10} - \beta_{12}y_{2t} - \beta_{13}y_{3t} - \gamma_{11}x_{1t} = \varepsilon_{1t}; \quad (6.77)$$

$$y_{2t} - \beta_{20} - \beta_{23}y_{3t} - \gamma_{21}x_{1t} - \gamma_{22}x_{2t} = \varepsilon_{2t}; \quad (6.78)$$

$$y_{3t} - \beta_{30} - \beta_{31}y_{1t} - \gamma_{31}x_{1t} - \gamma_{32}x_{2t} = \varepsilon_{3t}; \quad (6.79)$$

$$y_{4t} - \beta_{40} - \beta_{41}y_{1t} - \gamma_{42}y_{2t} - \gamma_{43}x_{3t} = \varepsilon_{4t}. \quad (6.80)$$

Для полегшення пояснення запишемо попередню систему у вигляді таблиці 6.2.

Таблиця 6.2

№ рівняння	1	y_1	y_2	y_3	y_4	x_1	x_2	x_3
(6.77)	$-\beta_{10}$	1	$-\beta_{12}$	$-\beta_{13}$	0	$-\gamma_{11}$	0	0
(6.78)	$-\beta_{20}$	0	1	$-\beta_{23}$	0	$-\gamma_{21}$	$-\gamma_{22}$	0
(6.79)	$-\beta_{30}$	$-\beta_{31}$	0	1	0	$-\gamma_{31}$	$-\gamma_{32}$	0
(6.80)	$-\beta_{40}$	$-\beta_{41}$	$-\beta_{42}$	0	1	0	0	$-\gamma_{43}$

Використовуючи дані з таблиці 6.2, перевіримо умови порядку для кожного окремого рівняння. Результат перевірки запишемо у вигляді таблиці 6.3.

Таблиця 6.3

№ рівняння	Кількість опущених попередньо визначених змінних ($K - k$)	Кількість включених ендогенних змінних мінус 1 ($m - 1$)	Ототожене
(6.77)	2	2	точно
(6.78)	1	1	точно
(6.79)	1	1	точно
(6.80)	2	2	точно

За умовою порядку кожне рівняння є точно ототожненим. Далі проведемо перевірку за умовою рангу. Розглянемо перше рівняння, у якому опущено змінні x_1 , x_2 і x_3 (нулі в першому рядку таблиці 6.2). Для пере-

вірки цього рівняння побудуємо відповідну матрицю коефіцієнтів при змінних y_4 , x_2 і x_3 , включених в інші рівняння моделі, крім першого.

Для нашого прикладу ця матриця матиме вигляд:

$$A = \begin{pmatrix} 0 & -\gamma_{22} & 0 \\ 0 & -\gamma_{32} & 0 \\ 1 & 0 & -\gamma_{43} \end{pmatrix}. \quad (6.81)$$

Визначник цієї матриці дорівнює нулю:

$$|A| = \begin{vmatrix} 0 & -\gamma_{22} & 0 \\ 0 & -\gamma_{32} & 0 \\ 1 & 0 & -\gamma_{43} \end{vmatrix} = 0. \quad (6.82)$$

Оскільки визначник дорівнює нулю, ранг матриці A менший трьох. Тому рівняння (6.77) не задовольняє умови рангу, тобто не є ототожненим.

Як зазначалося, рангова умова є і необхідною, і достатньою для ототожнення. Тому, хоча умова порядку показує, що рівняння (6.77) ототожене, рангова умова свідчить про протилежне. Стовпці та рядки матриці A (6.81) не є лінійно незалежними. Це означає, що між змінними y_4 , x_2 і x_3 є зв'язок. У результаті ми можемо не мати достатньої кількості інформації для оцінки параметрів рівняння (6.77); рівняння скороченої форми з попередньої моделі покажуть, що неможливо отримати структурні параметри цього рівняння з коефіцієнтів скороченої форми. Самостійно перевірте ранговою умовою рівняння (6.78) та (6.79) і пересвідчитесь, що вони також неототожені, а рівняння (6.80) — навпаки.

Для спрощення перевірку за ранговою умовою можна розбити на такі етапи.

1. Записати систему симультаивних рівнянь у табличній формі (див. табл. 6.2).

2. Викреслити коефіцієнти рядка, в якому з'являється рівняння, що розглядається.

3. Викреслити стовпці, відповідні ненульовим коефіцієнтам, рівняння, що розглядається.

4. Отримаємо необхідну матрицю A . Якщо ранг матриці A точно дорівнює $M-1$, то рівняння ототожене. Якщо ж ранг матриці A менший, ніж $M-1$, рівняння неототожене.

На базі умов порядку та умов рангу можна сформулювати загальні принципи ототожнення структурного рівняння в моделі, яка складається з M симультаивних рівнянь.

1. Якщо $K - k > m - 1$ і ранг матриці A буде дорівнювати $M - 1$, то відповідне рівняння переототожене.

2. Якщо $K - k = m - 1$ і ранг матриці A дорівнює $M - 1$, рівняння точно ототожене.

3. Якщо $K - k \geq m - 1$ і ранг матриці A менший, ніж $M - 1$, рівняння неототожене.

4. Якщо $K - k < m - 1$, структурне рівняння неототожене. Ранг матриці A в даному разі менший за $M - 1$.

6.8. МЕТОДИ ОЦІНЮВАННЯ НЕВІДОМИХ ПАРАМЕТРІВ СИМУЛЬТАТИВНИХ МОДЕЛЕЙ

6.8.1. Оцінювання точно ототоженого рівняння: метод непрямих найменших квадратів (ННК)

Для точно ототоженого структурного рівняння можна отримати структурні параметри з МНК-оцінок коефіцієнтів скороченої форми методом, відомим під назвою *метод непрямих найменших квадратів (ННК)*. Оцінка параметрів за цим методом умовно розбивається на такі три етапи.

1. Спочатку отримуємо рівняння скороченої форми. Для цього виражаємо залежну змінну в кожному рівнянні виключно через попередньо визначені (екзогенні та лагові) змінні та випадкові величини.

2. Окремо до кожного рівняння скороченої форми застосовуємо МНК. Це можливо, оскільки пояснювальні змінні в даних рівняннях попередньо визначені, а отже, некорельовані з випадковими величинами.

3. Отримуємо оцінки початкових структурних параметрів з оцінених на другому етапі коефіцієнтів скороченої форми. Якщо рівняння точно ототожене, є взаємна відповідність між структурними параметрами та коефіцієнтами скороченої форми (див.6.6).

Назва методу “непрямі найменші квадрати” (ННК) пов’язана з тим фактом, що структурні параметри отримуються з МНК-оцінок коефіцієнтів скороченої форми непрямо. Чому це так? Розглянемо ту ж саму модель попиту та пропозиції, яка наводилась раніше у 6.6, але для зручності викладу матеріалу дещо змінимо позначення. Виходячи з цього, функції попиту та пропозиції запишуться у такому вигляді:

функція попиту:

$$Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 x_t + \varepsilon_{1t}; \quad (6.83)$$

функція пропозиції:

$$Q_t = \beta_0 + \beta_1 P_t + \varepsilon_{2t}, \quad (6.84)$$

де Q — кількість; P — ціна; x — дохід або витрати.

Припустимо, що x — екзогенна змінна, а Q та P — ендогенні змінні.

Функція пропозиції — точно ототожнена, тоді як функція попиту — неототожнена.

Рівняннями скороченої форми, відповідними до структурних рівнянь, є:

$$P_t = \pi_0 + \pi_1 x_t + w_t; \quad (6.85)$$

$$Q_t = \pi_2 + \pi_3 x_t + v_t, \quad (6.86)$$

де π є коефіцієнтами скороченої форми і нелінійними комбінаціями структурних параметрів ((6.68) і (6.59)), а w та v є лінійними комбінаціями випадкових величин ε_1 і ε_2 .

Цікаво, що кожне рівняння скороченої форми містить лише одну ендогенну змінну, яка є залежною змінною і функцією виключно екзогенної змінної x (дохід) та випадкових величин. Отже, параметри попередніх рівнянь скороченої форми можуть бути оцінені методом найменших квадратів і матимуть вигляд:

$$\hat{\pi}_1 = \frac{\sum \bar{p}_t \bar{x}_t}{\sum \bar{x}_t^2}; \quad (6.87)$$

$$\hat{\pi}_0 = \bar{P} - \hat{\pi}_1 \bar{x}; \quad (6.88)$$

$$\hat{\pi}_3 = \frac{\sum \bar{Q}_t \bar{x}_t}{\sum \bar{x}_t^2}; \quad (6.89)$$

$$\hat{\pi}_2 = \bar{Q} - \hat{\pi}_3 \bar{x}, \quad (6.90)$$

де маленькими літерами із знаком “~” позначені відхилення від середнього, а \bar{Q} та \bar{P} означають середнє значення Q і P . π_i — оцінки коефіцієнтів скороченої форми, які за припущеннями є ефективними.

Оскільки ми мали на меті визначити структурні коефіцієнти, подивимося, чи можемо ми їх оцінити за коефіцієнтами скороченої форми. Функція пропозиції точно ототожнена (див.6.4). Тому її параметри можуть бути оцінені за коефіцієнтами скороченої форми таким чином:

$$\beta_0 = \pi_2 - \beta_1 \pi_0 \quad \text{і} \quad \beta_1 = \frac{\pi_3}{\pi_1}.$$

Отже, оцінки цих параметрів можна отримати з оцінок коефіцієнтів скороченої форми:

$$b_0 = \hat{\pi}_2 - b_1 \hat{\pi}_0, \quad (6.91)$$

$$b_1 = \frac{\hat{\pi}_3}{\hat{\pi}_1}, \quad (6.92)$$

які є МНК-оцінками. Параметри функції попиту не можуть бути оцінені таким чином.

Розглянемо як ілюстративний приклад залежність обсягу врожаю та цін на нього від доходу населення України. Умовні дані наведено в табл. 6.4.

Таблиця 6.4

Обсяг врожаю, ціни та особисті витрати на споживання на душу населення України в 1980—1995 рр. (в умовн. од. 1992 р.)

Рік	Індекс обсягу врожаю (1987=100)	Індекс цін врожаю (1987=100)	Реальні особисті витрати на споживання на душу населення
	Q	P	x
1980	77	52	7275
1981	86	56	7409
1982	87	60	7726
1983	92	91	7972
1984	84	117	7826
1985	93	105	7926
1986	92	102	8272
1987	100	100	8551
1988	102	105	8808
1989	113	116	8904
1990	101	125	8784
1991	116	134	8798
1992	118	1121	8825
1993	88	127	9148
1994	110	138	9462
1995	117	120	9682

Запишемо відповідно функції попиту та пропозиції:

$$Q_t^\alpha = \alpha_0 + \alpha_1 P_t + \alpha_2 x_t + \varepsilon_1 t,$$

$$Q_t^\beta = \beta_0 + \beta_1 P_t + \varepsilon_2 t,$$

де Q_t , P_t — ендогенні змінні;

x_t — екзогенна змінна.

Оцінимо функцію пропозиції.

Спочатку оцінюємо рівняння скороченої форми, записуючи окремо регресійні рівняння залежності ціни та обсягу врожаю від реальних споживчих витрат. Виходячи з даних таблиці 6.4, отримуємо:

$$\hat{P}_t = -161.0779 + 0.0313x_t \quad (6.93)$$

(48.3806) (0.0057)
(-3.329) (5.504)
 $R^2 = 0.6839$

$$Q_t = -25.3740 + 0.0146x_t \quad (6.94)$$

(25.0873) (0.0029)
(-1.011) (4.954)
 $R^2 = 0.6368$.

Використовуючи (6.91) та (6.92), отримуємо МНК-оцінки:

$$b_0 = 49.7527,$$

$$b_1 = 0.4664.$$

Тому оцінена за методом непрямих найменших квадратів функція пропозиції має вигляд:

$$\hat{Q}_t = 49.7527 + 0.4664P_t.$$

6.8.1.1. Властивості ННК-оцінок

Відомо, що оцінки коефіцієнтів скороченої форми є згідно з класичним припущеннями ефективними. Чи переносяться ці властивості на ННК-оцінки? Можна показати, що ННК-оцінки мають усі властивості оцінок скороченої форми, крім зміщеності. Теоретично можна показати, що ННК-оцінки є оцінками з відхиленнями, які зникають із збільшенням розміру вибірки (тобто оцінки є консистентними).

6.8.2. Оцінка переототоженого рівняння: метод двокрокових найменших квадратів (2МНК)

Для ілюстрації цього методу розглянемо приклад. Нехай ми маємо таку модель:

функція доходу:

$$y_{1t} = \beta_{10} + \beta_{11}y_{2t} + \gamma_{11}x_{1t} + \gamma_{12}x_{2t} + \varepsilon_{1t}; \quad (6.95)$$

функція пропозиції грошей:

$$y_{2t} = \beta_{20} + \beta_{21}y_{1t} + \varepsilon_{2t}, \quad (6.96)$$

де y_1 — дохід; y_2 — запас грошей; x_1 — інвестиційні витрати; x_2 — витрати уряду на товари та послуги.

Змінні x_1 та x_2 є екзогенними, y_1 та y_2 — ендогенними.

Рівняння доходу, яке ми розглядаємо, показує, що дохід визначається пропозицією грошей, інвестиційними витратами та витратами уряду. Рівняння пропозиції грошей показує, що запас грошей визначається відповідно до рівня доходів. Очевидно, що ми маємо симульативну модель.

Застосовуючи умову порядку для її ототожнення, бачимо, що рівняння доходу неототожене, тоді як рівняння пропозиції грошей — переототожене. Переототожена функція пропозиції грошей не може бути оцінена за допомогою методу ННК, тому що ми отримуємо дві різні оцінки β_{21} .

Якщо застосувати метод найменших квадратів для оцінки невідомих параметрів рівняння пропозиції грошей, то отримані оцінки будуть зміщеними через кореляцію між змінною y_1 та випадковою величиною ε_2 . Припустимо, що ми знайшли змінну, близьку до змінної y_1 в тому сенсі, що вона високо корелює з y_1 , але не є корельованою з ε_2 . Така змінна називається допоміжною змінною. Якщо її можна знайти, то МНК можна застосовувати для оцінки функції грошової пропозиції. Але як отримати таку допоміжну змінну? За допомогою методу двокрокових найменших квадратів. З назви видно, що метод складається з двох етапів.

1. Щоб позбавитись кореляції між y_1 і ε_2 , побудуємо спочатку регресійне рівняння залежності y_1 від усіх попередньо визначених змінних:

$$y_{1t} = \hat{\pi}_0 + \hat{\pi}_1 x_{1t} + \hat{\pi}_2 x_{2t} + e_t, \quad (6.97)$$

де e_t є помилками. Невідомі параметри рівняння (6.97) отримаємо за допомогою МНК:

$$\hat{y}_{1t} = \hat{\pi}_0 + \hat{\pi}_1 x_{1t} + \hat{\pi}_2 x_{2t} \quad (6.98)$$

Рівняння (6.97) є нічим іншим, як регресією скороченої форми, тому що в правій частині з'являються тільки екзогенні або попередньо визначені змінні. Його ще можна записати у вигляді:

$$y_{1t} = \hat{y}_{1t} + e_t, \quad (6.99)$$

який показує, що змінна y_1 складається з двох частин: \hat{y}_{1t} — прогновної величини та випадкової компоненти e_t . Виходячи з класичних припущень методу найменших квадратів, \hat{y}_{1t} та e_t — некорельовані між собою.

2. Рівняння пропозиції грошей можна записати таким чином:

$$\begin{aligned} y_{2t} &= \beta_{20} + \beta_{21}(\hat{y}_{1t} + e_t) + \varepsilon_{2t} \\ &= \beta_{20} + \beta_{21}\hat{y}_{1t} + (\varepsilon_{2t} + \beta_{21}e_t), \\ &= \beta_{20} + \beta_{21}\hat{y}_{1t} + \varepsilon_t^*, \end{aligned} \quad (6.100)$$

де $\varepsilon_t^* = \varepsilon_{2t} + \beta_{21}e_t$.

Порівнюючи (6.100) з (6.96) бачимо, що зовні ці рівняння дуже схожі, єдина відмінність полягає в тому, що y_1 замінено на \hat{y}_1 . В чому полягає перевага рівняння у вигляді (6.100)? Хоча y_1 в початковому рівнянні грошової пропозиції корелює з відхиленням ε_2 , \hat{y}_{1t} в (6.100) не корелюється з ε_t^* (у випадку, коли розмір моделі зростає пропорційно). В результаті, МНК можна застосувати до (6.100), з якого можна знайти відповідні оцінки параметрів функції пропозиції грошей.

Для подальшої ілюстрації методу 2МНК видозмінимо модель доходу та пропозиції грошей:

$$y_{1t} = \beta_{10} + \beta_{12}y_{2t} + \gamma_{11}x_{2t} + \varepsilon_{1t} \quad (6.101)$$

$$y_{2t} = \beta_{20} + \beta_{21}y_{1t} + \gamma_{23}x_{3t} + \gamma_{24}x_{4t} + \varepsilon_{2t}, \quad (6.102)$$

де додатково x_3 — дохід у попередньому періоді; x_4 — пропозиція грошей у попередньому періоді (вважаємо, що x_3 та x_4 — попередньо визначені).

Обидва рівняння (6.101) і (6.102) є переототожненими. Для того, щоб застосувати метод 2МНК, на першому етапі побудуємо регресійну модель залежності ендогенних змінних від усіх попередньо визначених змінних:

$$y_{1t} = \hat{\pi}_{10} + \hat{\pi}_{11}x_{1t} + \hat{\pi}_{12}x_{2t} + \hat{\pi}_{13}x_{3t} + \hat{\pi}_{14}x_{4t} + e_{1t}; \quad (6.103),$$

$$y_{2t} = \hat{\pi}_{20} + \hat{\pi}_{21}x_{1t} + \hat{\pi}_{22}x_{2t} + \hat{\pi}_{23}x_{3t} + \hat{\pi}_{24}x_{4t} + e_{2t}. \quad (6.104)$$

На другому етапі заміщуємо y_1 та y_2 в початкових (структурних) рівняннях їхніми оціненими значеннями з двох попередніх регресій:

$$y_{1t} = \beta_{10} + \beta_{12}\hat{y}_{2t} + \gamma_{11}x_{1t} + \gamma_{12}x_{2t} + \varepsilon_{1t}^*, \quad (6.105)$$

$$y_{2t} = \beta_{20} + \beta_{21}\hat{y}_{1t} + \gamma_{23}x_{3t} + \gamma_{24}x_{4t} + \varepsilon_{2t}^*, \quad (6.106)$$

де $\varepsilon_{1t}^* = \varepsilon_{1t} + \beta_{12}e_{2t}$, $\varepsilon_{2t}^* = \varepsilon_{2t} + \beta_{21}e_{1t}$.

Отримані таким чином оцінки будуть спроможними, тобто з розміром вибірки наблизатимуться до BLUE-оцінок.

6.8.3. Характерні особливості методу 2МНК

Можна виділити такі особливості методу 2МНК.

1. Метод можна застосувати до окремого рівняння в системі без врахування інших. Отже, для економетричних моделей, що складаються з великої кількості рівнянь, метод 2МНК є дуже економним, тому він широко використовується на практиці.

2. На відміну від ННК, який дає декілька різних оцінок параметра у переототожнених рівняннях, 2МНК дає лише одну оцінку параметра.

3. Для застосування методу потрібно знати тільки загальну кількість екзогенних або попередньо визначених змінних у системі.

4. Хоча метод 2МНК був спеціально розроблений для переототожнених рівнянь, його також можна застосовувати до точно ототожнених рівнянь. У цьому разі ННК та 2МНК дадуть ідентичні оцінки.

6.8.4. Ілюстративний приклад

Для ілюстрації методу 2МНК розглянемо модель доходу та грошової пропозиції, яка раніше наводилась у рівняннях (6.95) та (6.96).

Рівняння грошової пропозиції переототожене. Для оцінки параметрів цього рівняння ми застосовуємо метод двокрокових найменших квадратів (2МНК).

Необхідні умовні дані наводяться в табл. 6.5, де y_1 — ВВП; y_2 — пропозиція грошей; x_1 — інвестиційні витрати; x_2 — урядові витрати; x_3 — ставка процента і y_3 — ВВП з одиночним лагом. Усі числа, за винятком x_3 , подані в мільйонах гривень, x_3 подано в процентах.

Таблиця 6.5

Вибрані умовні макроекономічні дані (Україна, 1980—1994 рр.)

y_1	y_2	x_1	x_2	x_3	y_3
1015.5	216.6	148.8	218.2	7.29	963.9
1102.7	230.8	172.5	232.4	5.65	1015.5
1212.8	252.0	202.0	250.0	5.72	1102.7
1359.3	265.9	238.8	266.5	6.95	1212.8
1472.8	277.5	240.8	299.1	7.82	1359.3
1598.4	291.1	219.6	335.0	7.49	1472.8
1782.8	310.3	277.7	356.9	6.77	1598.4
1990.5	335.3	344.1	387.3	6.69	1782.8
2249.7	363.0	416.8	425.2	8.29	1990.5
2508.2	389.0	454.8	467.8	9.71	2249.7
2732.0	414.8	437.0	530.3	11.55	2508.2
3052.6	441.8	515.0	588.1	14.44	2732.0
3166.0	480.8	447.3	641.7	12.92	3052.6
3401.6	528.0	501.9	675.7	10.45	3166.0
3774.7	585.5	674.0	736.8	11.89	3401.6
3992.5	624.7	670.4	814.6	9.64	3774.7

1 крок. Спочатку побудуємо регресійну модель залежності доходу y_1 (ВВП) від попередньо визначених приватних інвестицій (x_1) та витрат уряду (x_2):

$$\begin{aligned} \hat{y}_{1t} &= -17.8799 + 1.3529x_{1t} + 3.9627x_{2t} \\ &\quad (34.1809) (0.3002) \quad (0.2626) \\ t &= (-0.523) \quad (4.506) \quad (15.090) \\ R^2 &= 0.9976. \end{aligned} \quad (6.107)$$

2 крок. Оцінимо функцію грошової пропозиції (6.92), замінюючи ендогенну y_1 на \hat{y}_1 з (6.107):

$$\begin{aligned} \hat{y}_2 &= 81.4682 + 0.1292\hat{y}_{1t} \\ &\quad (7.3377)(0.0030) \\ t &= (11.103) (43.416). \\ R^2 &= 0.9926 \end{aligned} \quad (6.108)$$

Оцінені стандартні помилки, наведені в (6.108), потрібно скоригувати (див. додатки 6.1, 6.2), після чого отримаємо:

$$\begin{aligned} \hat{y}_2 &= 81.4682 + 0.1292\hat{y}_{1t} \\ &\quad (9.9081) (0.0040) \\ t &= (8.2223) (32.300) \\ R^2 &= 0.9958. \end{aligned} \quad (6.109)$$

6.9. РЕКУРСИВНІ МОДЕЛІ

Особливим випадком симулятивних моделей є рекурсивні моделі. *Модель називається рекурсивною, якщо її структурні рівняння можна впорядкувати таким чином, що перше містить у правій стороні лише попередньо визначені змінні; друге містить попередньо визначені змінні та першу ендогенну змінну і так далі.* Наприклад:

$$\begin{aligned} y_1 &= f(x_1, x_2, \dots, x_k; \varepsilon_1); \\ y_2 &= f(x_1, x_2, \dots, x_k; y_1; \varepsilon_2); \\ y_3 &= f(x_1, x_2, \dots, x_k; y_1, y_2; \varepsilon_3) \end{aligned}$$

і так далі.

Однією з особливостей рекурсивної моделі є можливість оцінити її рівняння (по одному на кожному кроці) за допомогою МНК. Щоб проілюструвати це, перепишемо вищевказану рекурсивну модель у повній формі. Припустимо, що модель містить m ендогенних змінних та k екзогенних змінних.

$$\begin{aligned} y_1 &= \gamma_{11}x_1 + \gamma_{12}x_2 + \dots + \gamma_{1k}x_k + \varepsilon_1; \\ y_2 &= \gamma_{21}x_1 + \gamma_{22}x_2 + \dots + \gamma_{2k}x_k + \beta_{21}y_1 + \varepsilon_2; \\ y_3 &= \gamma_{31}x_1 + \gamma_{32}x_2 + \dots + \gamma_{3k}x_k + \beta_{31}y_1 + \beta_{32}y_2 + \varepsilon_3; \\ &\dots \\ y_m &= \gamma_{m1}x_1 + \gamma_{m2}x_2 + \dots + \gamma_{mk}x_k + \beta_{m1}y_1 + \beta_{m2}y_2 + \dots + \beta_{m,m-1}y_{m-1} + \varepsilon_m. \end{aligned}$$

Маючи значення екзогенних змінних, можемо застосувати МНК до першого рівняння і отримати оцінене значення y для першої ендогенної

змінної. Потім застосовуємо обчислене y_1 як пояснювальну змінну в другому рівнянні. Знову можна застосувати МНК, оскільки всі змінні незалежні від випадкової величини ε_2 . Зокрема, x та ε_2 незалежні за означенням; y_1 і ε_2 також незалежні, оскільки ε_1 — це єдина випадкова змінна, пов'язана з y_1 в першому рівнянні. Оскільки ми припустили незалежність змінних ε_1 та ε_2 , тоді y_1 та ε_2 також незалежні.

Рекурсивні системи називають ще трикутними системами, оскільки коефіцієнти при ендогенних змінних (β) утворюють трикутну матрицю, головна діагональ якої містить одиниці, а елементи над головною діагоналлю дорівнюють нулю. Для прикладу припустимо, що ми маємо модель з 4 ендогенними та 5 попередньо визначеними змінними:

$$\begin{aligned}y_1 &= \gamma_{11}x_1 + \gamma_{12}x_2 + \varepsilon_1; \\y_2 &= \beta_{21}y_1 + \gamma_{21}x_1 + \gamma_{22}x_2 + \gamma_{23}x_3 + \varepsilon_2; \\y_3 &= \beta_{31}y_1 + \beta_{32}y_2 + \gamma_{31}x_1 + \gamma_{34}x_4 + \varepsilon_3; \\y_4 &= \beta_{41}y_1 + \beta_{43}y_3 + \gamma_{44}x_4 + \gamma_{45}x_5 + \varepsilon_4.\end{aligned}$$

Для того, щоб визначити, чи є ця модель рекурсивною, достатньо дослідити форму матриці коефіцієнтів β . Якщо вона трикутна, то система рекурсивна. Рекурсивність системи дозволяє нам застосувати для оцінки невідомих параметрів метод МНК.

6.10. ОСНОВНІ ВИСНОВКИ ПРО СИМУЛЬТАТИВНІ МОДЕЛІ

У даному розділі було розглянуто симультазивні моделі. Їхньою особливістю є поява залежної змінної одного рівняння у ролі пояснювальної змінної у іншому рівнянні системи. Така змінна стає стохастичною і, як правило, сильно корельована з випадковою величиною рівняння, у якому з'являється як пояснювальна змінна. В такому випадку для оцінки невідомих параметрів не можна застосувати класичний метод найменших квадратів, бо отримані оцінки будуть зміщеними.

Для симультазивних рівнянь важливою є проблема ототожнення. Під проблемою ототожнення мається на увазі можливість отримання кількісних оцінок структурних параметрів симультазивної моделі з оцінених коефіцієнтів скороченої форми. Якщо це можна зробити, то рівняння в симультазивній моделі є ототожненим. І навпаки. Ототожене рівняння може бути як точно ототожненим, так і переототожненим. У першому випадку є єдині значення структурних параметрів, тоді як в останньому випадку структурні параметри можуть набувати більше одного значення.

Для перевірки можливості ототожнення рівняння можна застосувати умови порядку та рангу. Хоча умови порядку легко застосувати, вона є

лише достатньою, але не необхідною умовою ототожнення, тоді як умова рангу є необхідною і достатньою умовою. Якщо умова рангу виконана, умова порядку також виконана, але не навпаки.

Припустимо, що рівняння в симультазивній моделі ототожнено, є декілька методів для оцінки його параметрів. Вони розділяються на дві категорії: методи для одного рівняння та системні методи. Найпопулярнішими є методи оцінки для одного рівняння. Ми розглянули три найчастіше вживані методи оцінки невідомих параметрів окремого рівняння: МНК, ННК та 2МНК. Хоча МНК в загальному випадку не підходить для симультазивних моделей, його можна застосовувати до так званих рекурсивних моделей. Метод ННК прийнятний для точно ототожнених рівнянь. У цьому методі МНК застосовується до рівнянь скороченої форми, а структурні параметри оцінюються за коефіцієнтами скороченої форми. Метод 2МНК був спеціально розроблений для переототожнених рівнянь, хоча його також можна застосувати до точно ототожнених рівнянь. У такому разі оцінки, отримані ННК та 2МНК, будуть однакові. Основна ідея методу 2МНК полягає в заміщенні стохастичної пояснювальної змінної лінійною комбінацією (нестохастичних) попередньо визначених змінних у моделі.

Характерною рисою обох методів (ННК та 2МНК) є те, що отримані за цими методами оцінки є консистентними, тобто із збільшенням кількості спостережень вони наближаються до правильних значень.

ПІДСУМУЄМО ВИВЧЕНЕ

Після вивчення цього розділу ви зможете

1. Розуміти економічні проблеми, дослідження яких вимагає побудови симульативних моделей.
2. Записувати скорочену форму симульативної моделі. Ототожнювати рівняння симульативної моделі.
3. Тестувати за умовами порядку та рангу точну ототожненість, неототожненість та переототожненість.
4. Уміти оцінити невідомі параметри точно отожнених рівнянь методом непрямих найменших квадратів.
5. Уміти оцінити невідомі параметри переототожнених рівнянь методом двокрокових найменших квадратів.
6. Робити розрахунки на базі симульативних моделей.

Короткий огляд розділу

1. До цього ми розглядали окремі регресійні рівняння. В даному розділі було розглянуто симульативні моделі. Їхньою особливістю є поява залежної змінної одного рівняння у ролі пояснювальної змінної в іншому рівнянні системи. У такому разі для оцінки невідомих параметрів не можна застосувати класичний метод найменших квадратів, бо отримані оцінки будуть зміщеними.
2. Для симульативних рівнянь важливою є проблема ототожнення. Під проблемою ототожнення мається на увазі можливість отримання кількісних оцінок структурних параметрів рівнянь симульативної моделі за оціненими коефіцієнтами скороченої форми. Якщо це можна зробити, то рівняння в симульативній системі є ототожненим. І навпаки. Ототожене рівняння може бути як точно ототожненим, так і переототожненим.
3. Для перевірки можливості ототожнення рівняння можна застосувати умови порядку та рангу. Умова порядку є лише достатньою, але не необхідною умовою ототожнення, в той час як умова рангу є необхідною і достатньою умовою.
4. Для ототоженого рівняння симульативної моделі є декілька методів оцінки його параметрів, а саме: МНК, ННК та 2МНК. Хоча МНК в загальному випадку не підходить для симульативних моделей, його можна застосувати до так званих рекурсивних моделей. Метод ННК придатний для точно ототожнених рівнянь. Метод 2МНК був спеціально розроблений для переототожнених рівнянь, хоча його також можна застосувати до точно ототожнених. У такому разі оцінки, отримані ННК та 2МНК, збігатимуться.

Основні формули

Система симульативних рівнянь у загальному вигляді:

$$y_{1t} = \beta_{12}y_{2t} + \dots + \beta_{1M}y_{Mt} + \gamma_{11}x_{1t} + \dots + \gamma_{1K}x_{Kt} + \varepsilon_{1t};$$

$$y_{2t} = \beta_{21}y_{1t} + \dots + \beta_{2M}y_{Mt} + \gamma_{21}x_{1t} + \dots + \gamma_{2K}x_{Kt} + \varepsilon_{2t};$$

$$\dots \dots \dots$$

$$y_{Mt} = \beta_{M1}y_{1t} + \dots + \beta_{MM-1}y_{M-1,t} + \gamma_{M1}x_{1t} + \dots + \gamma_{MK}x_{Kt} + \varepsilon_{Mt}.$$

Система рекурсивних рівнянь:

$$y_1 = \gamma_{11}x_1 + \gamma_{12}x_2 + \dots + \gamma_{1k}x_k + \varepsilon_1;$$

$$y_2 = \gamma_{21}x_1 + \gamma_{22}x_2 + \dots + \gamma_{2k}x_k + \beta_{21}y_1 + \varepsilon_2;$$

$$y_3 = \gamma_{31}x_1 + \gamma_{32}x_2 + \dots + \gamma_{3k}x_k + \beta_{31}y_1 + \beta_{32}y_2 + \varepsilon_3;$$

$$\dots \dots \dots$$

$$y_m = \gamma_{m1}x_1 + \gamma_{m2}x_2 + \dots + \gamma_{mk}x_k + \beta_{m1}y_1 + \beta_{m2}y_2 + \dots + \varepsilon_m.$$

Загальні принципи ототожнення структурного рівняння в моделі, яка складається з M симульативних рівнянь

1. Якщо $K - k > m - 1$ і ранг матриці A дорівнюватиме $M - 1$, то відповідне рівняння переототожене.
2. Якщо $K - k = m - 1$ і ранг матриці A дорівнює $M - 1$, рівняння точно ототожене.
3. Якщо $K - k < m - 1$, рівняння неототожене. Ранг матриці A в даному разі менший за $M - 1$.

ВПРАВИ І КОМЕНТАРІ

Вправа 6.1. Реклама, концентрація та граничні ціни/витрати

Для вивчення взаємозв'язків між рекламою, концентрацією (обчислюється за пропорцією спеціалізації) та цінними витратами Алєнд Стрікланд і Ленорд Вайс сформулювали таку модель*:

функція інтенсивності реклами:

$$AD/S = a_0 + a_1M + a_2(CD/S) + a_3C^2 + a_5GR + a_6Dur; \quad (6.110)$$

функція концентрації:

$$C = b_0 + b_1(AD/S) + b_2(MES/S); \quad (6.111)$$

* Strickland A.D., Weiss L. Advertising, Concentration and Price Margins // Journal of Political Economy. — 1976. — №5.

функція граничних витрат:

$$M = c_0 + c_1(K/S) + c_2Gr + c_3C + c_4GD + c_5(AD/S) + c_6(MES/S), \quad (6.112)$$

де AD — витрати на рекламу; S — вартість перевезень; C — пропорція спеціалізації чотирьох фірм; CD — попит споживачів; MES — мінімально діючий масштаб; M — граничні ціни/витрати; Gr — річний рівень зростання промислового виробництва; Du — dummy-змінна для виробництва товарів багаторазового використання; K — запас капіталу; GD — географічне розповсюдження виробництва.

Розв'язок

Таблиця 6.6

Оцінки, отримані за методом двокрокових найменших квадратів трьох рівнянь (у дужках наведено середньоквадратичні відхилення)

Залежні змінні			
	Ad/S	C	M
	Рівняння (1)	Рівняння (2)	Рівняння (3)
Константа	-0.0245 (-3.86)	0.2591 (21.30)	0.1736 (14.66)
C	0.0737 (2.84)	...	0.0377 (0.93)
C^2	-0.0643 (-2.64)
M	0.0544 (2.01)
CD/S	0.0269 (8.96)
Gr	0.0539 (2.09)	...	0.2336 (2.61)
Dur	-0.0018 (-0.93)
Ad/S	...	1.5347 (2.42)	1.6256 (5.52)
MES/S	...	4.169 (18.84)	0.1720 (0.92)
K/S	0.1165 (7.30)
GD	-0.0003 (-2.79)

Відповідно до необхідних умов ототожнення, рівняння (6.111) є переототожненим, тоді як (6.110) та (6.112) — точно ототожені.

Інформацію для аналізу взято переважно з перепису виробництв за 1983 рік. Вона стосується 408 з 417 галузей промисловості. Невідомі параметри симульативної моделі автори оцінили, використовуючи 2МНК. Відповідні результати подано в табл. 6.6.

Вправа 6.2. Модель оцінювання капітальних активів подана у вигляді рекурсивної системи

У досить незвичайному застосуванні моделювання рекурсивного симульативного рівняння Лі і Лойд оцінили таку модель для нафтової промисловості*:

$$\begin{aligned} R_{1t} &= \alpha_1 && + \gamma_1 M_t + \varepsilon_{1t}, \\ R_{2t} &= \alpha_2 + \beta_{21} R_{1t} && + \gamma_2 M_t + \varepsilon_{2t}, \\ R_{3t} &= \alpha_3 + \beta_{31} R_{1t} + \beta_{32} R_{2t} && + \gamma_3 M_t + \varepsilon_{3t}, \\ R_{4t} &= \alpha_4 + \beta_{41} R_{1t} + \beta_{42} R_{2t} + \beta_{43} R_{3t} && + \gamma_4 M_t + \varepsilon_{4t}, \\ R_{5t} &= \alpha_5 + \beta_{51} R_{1t} + \beta_{52} R_{2t} + \beta_{53} R_{3t} + \beta_{54} R_{4t} && + \gamma_5 M_t + \varepsilon_{5t}, \\ R_{6t} &= \alpha_6 + \beta_{61} R_{1t} + \beta_{62} R_{2t} + \beta_{63} R_{3t} + \beta_{64} R_{4t} + \beta_{65} R_{5t} && + \gamma_6 M_t + \varepsilon_{6t}, \\ R_{7t} &= \alpha_7 + \beta_{71} R_{1t} + \beta_{72} R_{2t} + \beta_{73} R_{3t} + \beta_{74} R_{4t} + \beta_{75} R_{5t} + \beta_{76} R_{6t} && + \gamma_7 M_t + \varepsilon_{7t}, \end{aligned}$$

де R_1 — ставка доходу від застави 1 (“імператорська нафта”);

R_2 — ставка доходу від застави 2 (“сонячна нафта”);

...

R_7 — ставка доходу від застави 7 (“стандарт Індіани”);

M_t — ринкова ставка доходу;

ε_{it} — випадкові величини ($i=1,2,\dots,7$).

Перед тим, як ми представимо результати, постає запитання: як обирати заставу 1, 2 і т.д.? Лі та Лойд відповідають: емпірично. Вони будують регресійне рівняння залежності ставки доходу від застави i за ставками доходу від шести застав, які залишились, і спостерігають за R^2 . Таким чином, буде сім регресій. Потім вони впорядковують оцінені R^2 — від найменшого до найбільшого. Застава, яка має найнижчий R^2 , визначається як застava 1, і відповідно та, що має найвищий R^2 , є заставою 7. Ідея досить проста. Якщо ставка доходу R^2 від, скажімо, “імператорської нафти” найнижча у порівнянні з рештою, це означає, що на дану заставу найменше впливають доходи інших застав. Тобто є односторонній причинно-порядковий зв'язок цієї застави з іншими.

Тепер представимо емпіричні результати, які наводяться у табл. 6.7.

Результати регресії Лі — Лойда повідомляють про важливий міжгалузевий зв'язок між доходами від застави. Наприклад, доход від “стандарту Індіани” залежить не тільки від ринкової ставки доходу, а й від ставок доходу від

* Lee C., Loyd P. The Capital Asset Pricing Model Expressed as a Recursive System: An Empirical Investigation // Journal of Financial and Quantitative.

Таблиця 6.7

Рекурсивна система оцінок для нафтової промисловості

Залежні змінні (лінійна форма)							
	Стандарт Індіани	Нафта Shell	Бензин Філіпа	Союзна нафта	Стандарт Охіо	Сонячна нафта	Нафта імперії
Стандарт Індіани							
Нафта Shell	0.2100 (2.859)						
Бензин Філіпа	0.2293 (2.176)	0.0791 (1.065)					
Союзна нафта	0.1754 (2.472)	0.2171 (3.177)	0.2225 (2.337)				
Стандарт Охіо	-0.0794 (-1.294)	0.0147 (0.235)	0.4248 (5.501)	0.1468 (1.735)			
Сонячна нафта	0.1249 (1.343)	0.1710 (1.843)	0.0472 (0.355)	0.1339 (0.908)	0.0499 (0.271)		
Нафта імперії	-0.1077 (-1.412)	0.0526 (0.6804)	0.0354 (0.319)	0.1580 (1.290)	-0.2541 (-1.691)	0.0828 (0.971)	
Константа	0.0868 (0.681)	-0.0384 (1.296)	-0.0127 (-0.068)	-0.2034 (0.986)	0.3009 (1.204)	0.2013 (1.399)	0.3710 (2.161)
Ринковий індекс	0.3681 (2.165)	0.4997 (3.039)	0.2884 (1.232)	0.7609 (3.069)	0.9089 (3.094)	0.7161 (4.783)	0.6432 (3.774)
R^2	0.5020	0.4658	0.4106	0.2532	0.0985	0.2404	0.1247
Durbin-Watson	2.1083	2.4714	2.2306	2.3468	2.2181	2.3109	1.9592

“нафти Shell”, “бензину Філіпа” та “союзної нафти”. Іншими словами, зміну ставки доходу від “стандарту Індіани” можна пояснити, якщо додатково до ринкового доходу ми також розглянемо ставки доходу від “нафти Shell”, “бензину Філіпа” та “союзної нафти”.

ТЕСТИ

Вибрати правильну відповідь на запитання

1. Симультативна модель складається з:
- рівно n рівнянь та n невідомих;
 - одного лінійного регресійного рівняння;
 - одного нелінійного регресійного рівняння;
 - одного та більше регресійних рівнянь;
 - більше ніж одного регресійного рівняння.

2. Рівняння симультативної моделі:
- завжди можна точно ототожнити;
 - завжди можна переототожнити;
 - завжди можна недоототожнити;
 - можливі будь-які варіанти.

3. Рівняння симультативної моделі точно ототожене, якщо:
- $K - k > t - 1$ і ранг матриці A буде дорівнювати $M - 1$;
 - $K - k = t - 1$ і ранг матриці A дорівнює $M - 1$;
 - $K - k \geq t - 1$ і ранг матриці A менший за $M - 1$;
 - $K - k < t - 1$ і ранг матриці A менший за $M - 1$.

4. Рівняння симультативної моделі переототожене, якщо:
- $K - k > t - 1$ і ранг матриці A буде дорівнювати $M - 1$;
 - $K - k = t - 1$ і ранг матриці A дорівнює $M - 1$;
 - $K - k \geq t - 1$ і ранг матриці A менший за $M - 1$;
 - $K - k < t - 1$, ранг матриці A менший за $M - 1$.

5. Рівняння симультативної моделі неототожене:
- якщо $K - k > t - 1$ і ранг матриці A буде дорівнювати $M - 1$;
 - якщо $K - k = t - 1$ і ранг матриці A дорівнює $M - 1$;
 - якщо $K - k \geq t - 1$ і ранг матриці A менший ніж $M - 1$;
 - якщо $K - k < t - 1$, ранг матриці A менший за $M - 1$.

6. Метод 2МНК можна застосовувати для оцінки параметрів:
- точно ототоженого рівняння;
 - недоототоженого рівняння;
 - переототоженого рівняння;
 - неможливо використовувати в будь-якому випадку;
 - у випадку рекурсивних моделей.

7. Метод МНК можна застосовувати для оцінки параметрів:
- точно ототоженого рівняння;
 - недоототоженого рівняння;
 - переототоженого рівняння;
 - неможливо використовувати в будь-якому випадку;
 - у випадку рекурсивних моделей.

8. Метод НеНК можна застосовувати для оцінки параметрів:
- точно ототоженого рівняння;
 - недоототоженого рівняння;
 - переототоженого рівняння;
 - неможливо використовувати в будь-якому випадку;
 - у випадку рекурсивних моделей.

9. Зробити остаточний висновок про ототожненість рівняння можна за умовою:

- порядку;
- порядку або рангу;
- рангу;
- взагалі неможливо зробити.

ВІДПОВІСТІ НА ЗАПИТАННЯ: ТАК/НІ; ЯКЩО НІ, ПОЯСНІТЬ ЧОМУ

- Чи еквівалентні обидва визначення умови порядку?
- Для моделі

$$\begin{aligned} y_{1t} &= \beta_{10} + \beta_{12}y_{2t} + \gamma_{11}x_{1t} + \varepsilon_{1t}, \\ y_{2t} &= \beta_{20} + \beta_{21}y_{1t} + \gamma_{22}x_{2t} + \varepsilon_{2t} \end{aligned}$$

який вигляд матимуть рівняння скороченої форми: а чи б?

- $$\begin{aligned} y_{1t} &= \pi_{10} + \pi_{11}x_{1t} + \pi_{12}x_{2t} + w_t; \\ y_{2t} &= \pi_{20} + \pi_{21}x_{1t} + \pi_{22}x_{2t} + v_t; \end{aligned}$$
- $$\begin{aligned} y_{1t} &= \pi_{10} + \pi_{12}x_{2t} + w_t; \\ y_{2t} &= \pi_{20} + \pi_{21}x_{1t} + v_t. \end{aligned}$$

- Нехай ми маємо таке визначення умови порядку:

$$K \geq m + k - 1,$$

яке стверджує, що кількість попередньо визначених змінних у системі може бути не меншою, ніж кількість невідомих коефіцієнтів рівняння, яке ототожнюється. Чи буде це визначення еквівалентне двом іншим визначенням умови порядку?

- Чи обов'язково застосовувати метод 2МНК до точно ототожнених рівнянь?

- Чи буде ототожненим у такій моделі хоча б одне рівняння?

Функція попиту:

$$Q_t = \alpha_0 + \alpha_1 P_t + \varepsilon_{1t}.$$

Функція пропозиції:

$$Q_t = \beta_0 + \beta_1 P_t + \varepsilon_{2t},$$

де Q — кількість продукції, що вибирається; P — рівень цін.

- Чи буде в цій моделі ототожнена функція пропозиції?

Функція попиту:

$$Q_t = \alpha_0 + \alpha_1 P_t + \alpha_{2Dt} + \varepsilon_{1t}.$$

Функція пропозиції:

$$Q_t = \beta_0 + \beta_1 P_t + \varepsilon_{2t},$$

де Q — кількість продукції; P — рівень цін; D — доход споживача.

- Чи буде в цій моделі ототожнена функція попиту?

Функція попиту:

$$Q_t = \alpha_0 + \alpha_1 P_t + \alpha_{2Dt} + \varepsilon_{1t}.$$

Функція пропозиції:

$$Q_t = \beta_0 + \beta_1 P_t + \varepsilon_{2t},$$

де Q — кількість продукції; P — рівень цін; D — доход споживача.

- Чи буде в даній моделі ототожнена функція пропозиції?

Функція попиту:

$$Q_t = \alpha_0 + \alpha_1 P_t + \alpha_{2Dt} + \alpha_3 R_t + \varepsilon_{1t}.$$

Функція пропозиції:

$$Q_t = \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + \varepsilon_{2t},$$

де Q — кількість продукції; P — рівень цін; D — доход споживача, R_t — багатство.

- Якщо $K - k > m - 1$ і ранг матриці A буде дорівнювати $M - 1$, то чи буде відповідне рівняння симультаивної моделі переототожене?

- Якщо $K - k = m - 1$ і ранг матриці A дорівнює $M - 1$, чи буде рівняння симультаивної моделі неототожене? Рівняння точно ототожене?

ВІДПОВІДІ

Тест 1

- д; 6) а, в;
- г; 7) д;
- б; 8) а;
- а; 9) в.
- г;

Так/Ні

1. Так.
2. Ні. Рівняння скороченої форми мають вигляд:

$$y_{1t} = \pi_{10} + \pi_{11}x_{1t} + \pi_{12}x_{2t} + w_t;$$

$$y_{2t} = \pi_{20} + \pi_{21}x_{1t} + \pi_{22}x_{2t} + v_t.$$

3. Так.
4. Ні, до точно ототожнених рівнянь можна застосувати метод ННК.
5. Ні. Ця модель має дві ендогенні змінні P та Q і жодної попередньо визначеної. Для ототожнення в кожному рівнянні має бути опущена щонайменше $M - 1 = 2 - 1 = 1$ змінна. У даному разі жодне з рівнянь не буде ототожненим.
6. Так.
7. Ні. В даній моделі Q і P ендогенні, а D — екзогенна. Застосовуючи умову порядку, бачимо, що функція попиту недоототожнена. З іншого боку, функція пропозиції ототожнена, бо вона виключає рівно $M-1=1$ змінну D_t .
8. Ні. В даній моделі P_t і Q_t ендогенні, а D_t , R_t і P_{t-1} попередньо визначені. У функції попиту опущено рівно одну змінну P_{t-1} , за умовою порядку вона точно ототожнена. У функції пропозиції опущено дві змінні D_t та R_t , отже, вона переототожнена. Як зазначалось раніше, в даному разі ми отримуємо декілька оцінок для одного параметра β_1 .
9. Так.
10. Рівняння точно ототожене.

ВИКОНАЙТЕ САМОСТІЙНО

Вправи

Вправа 6.3

У табл. 6.8 наведено дані: y — ВНП; C — власні споживчі витрати; I — валові приватні внутрішні інвестиції (в мільярдах умовних одиниць 1982 р. та сезонно скориговані).

Таблиця 6.8

Роки	y	C	I	Роки	y	C	I
1970	2416.2	1492.0	381.5	1978	3115.2	1961.0	576.9
1971	2484.8	1538.8	419.3	1979	3192.4	2004.4	575.2
1972	2608.5	1621.9	465.4	1980	3187.1	2000.4	509.3
1973	2744.1	1689.6	520.8	1981	3248.8	2024.2	545.5
1974	2729.3	1674.0	481.3	1982	3166.0	2050.7	447.3
1975	2695.0	1711.9	383.3	1983	3277.7	2145.9	503.4
1976	2826.7	1803.9	435.5	1984	3492.0	2239.9	661.3
1977	2958.6	1883.8	521.3	1985	3473.5*	2312.6*	650.6*

Припустимо, що C лінійно залежить від y , як і у простій кейнсіанській моделі визначення доходу у прикладі 6.3. Знайдіть МНК-оцінки параметрів функції споживання. (Збережіть результати для наступних вправ.)

Вправа 6.4

А. Для моделі “попит — пропозиція” вправи 6.3 знайдіть вираз границі імовірності для $\hat{\alpha}_1$.

Б. За яких умов це граничне значення буде дорівнювати правильному α_1 ?

Вправа 6.5

Галавей і Сміт створили просту модель для економіки США, яка має такий вигляд:

$$y_t = G_t + \beta_2 y D_{t-1} + \beta_3 M_t + \varepsilon_{1t},$$

$$I_t = \beta_4 + \beta_5 (y_{t-1} - y_{t-2}) + \beta_6 Z_{t-1} + \varepsilon_{2t},$$

$$G_t = \beta_7 + \beta_8 G_{t-1} + \varepsilon_{3t},$$

де y — валовий національний дохід; C — витрати на власне споживання; I — валові приватні внутрішні інвестиції; G — державні витрати плюс чисті іноземні інвестиції; yD — наявний дохід (дохід після сплати податків), M — гроші, запропоновані на початку кварталу; Z — власний дохід до сплати податків; t — час; $\varepsilon_1, \varepsilon_2, \varepsilon_3$ — випадкові величини.

Використовуючи дані щоквартальників за 1948—1957 рр., автори застосували метод найменших квадратів до кожного рівняння окремо і отримали такі результати:

$$y_t = 0.09 + 0.43YD_{t-1} + 0.23M_t \quad R^2 = 0.23;$$

$$I_t = 0.08 + 0.43(Y_{t-1} - Y_{t-2}) + 0.48Z_{t-1} \quad R^2 = 0.40;$$

$$G_t = 0.13 + 0.67G_{t-1} \quad R^2 = 0.42.$$

А. Як ви оцінюєте використання методу найменших квадратів для одного рівняння?

Б. Чому значення R^2 є досить малими?

Вправа 6.6

Ж. Менгес розробив таку економетричну модель для західнонімецької економіки:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 I_t + \varepsilon_{1t};$$

$$I_t = \beta_3 + \beta_4 y_t + \beta_5 Q_t + \varepsilon_{2t};$$

$$C_t = \beta_6 + \beta_7 y_t + \beta_8 C_{t-1} + \beta_9 P_t + \varepsilon_{3t};$$

$$Q_t = \beta_{10} + \beta_{11} Q_{t-1} + \beta_{12} R_t + \varepsilon_{4t},$$

де y — валовий національний дохід; C — витрати на власне споживання; I — чисте формування капіталу; Q — прибутки; P — витрати індексу життя; R — продукція виробництва; t — час; ε — випадкові величини.

А. Які змінні ми вважаємо ендогенними і які екзогенними?

Б. Чи є у цій системі хоча б одне рівняння, яке можна оцінити за допомогою методу найменших квадратів?

В. Що є причиною введення змінної C_{t-1} у функцію споживання?

Вправа 6.7

Розробіть симультивну модель пропозиції та попиту викладачів вузів в Україні. Визначте ендогенні та екзогенні змінні цієї моделі.

Вправа 6.8

Створіть просту модель попиту та пропозиції грошей в Україні і порівняйте вашу модель з прикладом 6.8.4.

Вправа 6.9

У своїй статті “Модель розподілу якісних продуктів особистого користування, вироблених у Ямайці” Джон У. Фарлей і Гарольд Дж. Левіт створили таку модель (продуктами особистого користування були крем для гоління, крем для шкіри, гігієнічні серветки і зубна паста):

$$\begin{aligned}y_{1i} &= \alpha_1 + \beta_1 y_{2i} + \beta_2 y_{3i} + \beta_3 y_{4i} + \varepsilon_{1i}; \\y_{2i} &= \alpha_2 + \beta_4 y_{1i} + \beta_5 y_{5i} + \gamma_1 x_{1i} + \gamma_2 x_{2i} + \varepsilon_{2i}; \\y_{3i} &= \alpha_3 + \beta_6 y_{2i} + \gamma_3 x_{3i} + \varepsilon_{3i}; \\y_{4i} &= \alpha_4 + \beta_7 y_{2i} + \gamma_4 x_{4i} + \varepsilon_{4i}; \\y_{5i} &= \alpha_5 + \beta_8 y_{2i} + \beta_9 y_{3i} + \beta_{10} y_{4i} + \varepsilon_{5i},\end{aligned}$$

де y_1 — кількість магазинів, у яких зберігається продукція (%); y_2 — місячний обсяг продажу продукції; y_3 — індекс безпосереднього зв'язку з імпортером та виробником продукції; y_4 — індекс активності оптової торгівлі у певному регіоні; y_5 — індекс величини якості продукції, що є у запасі (наприклад, середнє значення якості продукції, запасеної магазином); x_1 — запланована кількість населення, що вживає даний продукт; x_2 — дохід на душу населення в регіоні; x_3 — відстань від населеного центру регіона до його крайньої точки; x_4 — відстань від населеного центру до найближчого місця оптової торгівлі.

А. Чи можна визначити ендогенні та екзогенні змінні у попередній моделі?

Б. Чи можна у цій моделі оцінити одне або декілька рівнянь за допомогою методу найменших квадратів?

Вправа 6.10

Щоб вивчити взаємозв'язок між витратами на рекламу і продажем цигарок, Франк Бас використав таку модель:

$$\begin{aligned}y_{1t} &= \alpha_1 + \beta_1 y_{3t} + \beta_2 y_{4t} + \gamma_1 x_{1t} + \gamma_2 x_{2t} + \varepsilon_{1t}; \\y_{2t} &= \alpha_2 + \beta_3 y_{3t} + \beta_4 y_{4t} + \gamma_3 x_{1t} + \gamma_4 x_{2t} + \varepsilon_{2t}; \\y_{3t} &= \alpha_3 + \beta_5 y_{1t} + \beta_6 y_{2t} + \varepsilon_{3t}; \\y_{4t} &= \alpha_4 + \beta_7 y_{1t} + \beta_8 y_{2t} + \varepsilon_{4t},\end{aligned}$$

де y_1 — логарифм продажу цигарок з фільтром (кількість цигарок, поділена на кількість населення віком старше 20 років); y_2 — логарифм продажу цигарок без фільтра (кількість цигарок, поділена на кількість населення віком старше 20 років); y_3 — логарифм кількості доларів, витрачених на рекламу цигарок з фільтром, поділений на кількість населення віком старше 20 років, поділений на індекс рекламних цін; y_4 — логарифм доларів, витрачених на рекламу цигарок без фільтра, поділений на кількість населення віком старше 20 років, поділений на індекс рекламних цін; x_1 — логарифм наявного власного доходу, поділений на кількість населення віком старше 20 років, поділений на індекс споживчих цін; x_2 — логарифм цін на пакет цигарок без фільтра, поділений на індекс споживчих цін.

А. У попередній моделі y — ендогенні і x — екзогенні змінні. Чому автор припускає, що x_2 — екзогенна змінна?

Б. Якщо x_2 взяти за ендогенну змінну, то яким чином ви будете видозмінювати попередню модель?

В. Чи повинна входити у модель ціна цигарок з фільтром? Якщо так, то вона виступає ендогенною чи екзогенною змінною? Якщо ендогенна, то як видозміниться первісна модель?

Г. Чи можливо у моделі Басса одне або більше рівнянь оцінити за допомогою методу найменших квадратів? Доведіть вашу відповідь.

Вправа 6.11

Покажіть, що обидва визначення умови порядку, розглянуті в моделі 6, еквівалентні.

Вправа 6.12

Знайдіть структурні коефіцієнти за коефіцієнтами скороченої форми, наведеними в (6.67) і (6.69).

Вправа 6.13

Введіть скорочену форму таких моделей і визначте, в якому разі структурні рівняння не-, точно, чи переототоженні:

- (6.61);
- (6.71);
- (6.8).

Вправа 6.14

Перевірте ототожнення моделей вправи 6.13, застосовуючи умови порядку та рангу.

Вправа 6.15

У моделі (6.6) і (6.70) показано, що рівняння пропозиції — неототожене. Які обмеження на структурні параметри зроблять це рівняння точно ототожненим? Поясніть ці обмеження.

Вправа 6.16

З моделі

$$y_{1t} = \beta_{10} + \beta_{12}y_{2t} + \gamma_{11}x_{1t} + \varepsilon_{1t};$$

$$y_{2t} = \beta_{20} + \beta_{21}y_{1t} + \gamma_{22}x_{2t} + \varepsilon_{2t}$$

виведено такі рівняння скороченої форми:

$$y_{1t} = \pi_{10} + \pi_{11}x_{1t} + \pi_{12}x_{2t} + w_t;$$

$$y_{2t} = \pi_{20} + \pi_{21}x_{1t} + \pi_{22}x_{2t} + v_t.$$

А. Чи ототожені структурні рівняння?

Б. Що трапиться з ототоженням, якщо відомо апріорі, що $\gamma_{11} = 0$?

Вправа 6.17

Повернемося до вправи 6.19. Оцінені рівняння скороченої форми мають такий вигляд:

$$y_{1t} = 4 + 3x_{1t} + 8x_{2t};$$

$$y_{2t} = 2 + 6x_{1t} + 10x_{2t}.$$

А. Визначте значення структурних параметрів.

Б. Як би ви тестували нуль-гіпотезу $\gamma_{11} = 0$?

Вправа 6.18

Модель

$$y_{1t} = \beta_{10} + \beta_{12}y_{2t} + \gamma_{11}x_{1t} + \varepsilon_{1t};$$

$$y_{2t} = \beta_{20} + \beta_{21}y_{1t} + \varepsilon_{2t}$$

утворює такі рівняння скороченої форми:

$$y_{1t} = 4 + 8x_{1t};$$

$$y_{2t} = 2 + 12x_{1t}.$$

А. Які структурні параметри (якщо такі є) можуть бути оцінені за коефіцієнтами скороченої форми? Обґрунтуйте вашу точку зору.

Б. Як змінюється відповідь на (А), якщо $\beta_{12} = 0$ і $\beta_{10} = 0$?

Таблиця 6.9

№ рівняння	y_1	y_2	y_3	y_4	y_5	x_1	x_2	x_3	x_4
1	1	β_{12}	0	β_{14}	0	γ_{11}	0	0	γ_{14}
2	0	1	β_{23}	β_{24}	0	0	γ_{22}	γ_{23}	0
3	β_{13}	0	1	β_{34}	β_{35}	0	0	γ_{33}	γ_{34}
4	0	β_{42}	0	1	0	γ_{41}	0	γ_{43}	0
5	β_{51}	0	0	β_{54}	1	0	γ_{52}	γ_{53}	0

Вправа 6.19

Наступна модель (див. табл. 6.9) містить 5 рівнянь з 5 ендогенними змінними y і з 4 екзогенними змінними x .

Визначте ототоженість кожного рівняння за допомогою умов порядку та рангу.

Вправа 6.20

Розглянемо таку розширену кейнсіанську модель визначення доходу: функція споживання:

$$C_t = \beta_1 + \beta_2 y_t - \beta_3 T_t + 1t;$$

функція інвестицій:

$$I_t = \alpha_0 + \alpha_1 y_{t-1} + \varepsilon_{2t};$$

функція оподаткування:

$$T_t = \gamma_0 + \gamma_1 y_t + e_{3t};$$

рівняння доходу:

$$y_t = C_t + I_t + G_t,$$

де C — споживчі витрати; y — доход; I — інвестиції; T — податки; ε — випадкові величини.

У моделі C , I , T , y — ендогенні змінні, а y_{t-1} і G — попередньо визначені.

Застосовуючи умову порядку, перевірте ототоженість кожного рівняння в системі і систему загалом. Що трапилося б, якби праворуч функції інвестицій з'явилася екзогенна змінна, ставка процента r_t ?

Вправа 6.21

Припустимо, ми пропонуємо ще одне визначення умови порядку ототоження:

$$K \geq m + k - 1,$$

яке стверджує, що кількість попередньо визначених змінних у системі може бути не меншою, ніж кількість невідомих коефіцієнтів рівняння, яке ототож-

нюється. Покажіть, що це визначення еквівалентне двом іншим визначенням умови порядку.

Вправа 6.22

Спрощена версія моделі ринку кавунів.

Рівняння попиту:

$$P_t = \alpha_0 + \alpha_1(Q_t/N_t) + \alpha_2(y_t/N_t) + \alpha_3F_t + \varepsilon_{1t}.$$

Функція пропозиції врожаю:

$$Q_t = \beta_0 + \beta_1(P_t/W_t) + \beta_2P_{t-1} + \beta_3C_{t-1} + \beta_4T_{t-1} + \varepsilon_{2t},$$

де P — ціна; (Q/N) — попит на одиницю капіталу; (y/N) — доход на одиницю капіталу; F_t — витрати на перевезення; (P/W) — ціна, пов'язана з рівнем зарплати на фермі; C — ціна бавовни; T — ціна інших овочів.

P і Q є ендогенними змінними.

А. Виведіть скорочену формулу.

Б. Визначте, чи є функція попиту, пропозиції або обидві ототожненими?

Вправа 6.23

Чому необов'язково застосовувати метод 2МНК до точно ототожнених рівнянь?

Вправа 6.24

Розглянемо змінену кейнсіанську модель визначення доходу:

$$C_t = \beta_{10} + \beta_{11}y_t + \varepsilon_{1t};$$

$$I_t = \beta_{20} + \beta_{21}y_t + \beta_{22}y_{t-1} + \varepsilon_{2t};$$

$$y_t = C_t + I_t + G_t,$$

де C — витрати на споживання; I — інвестиційні витрати; y — доход; G — витрати уряду.

G_t і y_{t-1} попередньо визначені.

А. Виведіть рівняння скороченої форми і визначте, які з попередніх рівнянь є точно або переототожненими.

Б. Який метод ви будете використовувати для оцінки параметрів точно або переототоженого рівняння? Відповідь доведіть.

Вправа 6.25

Розглянемо такі результати.

$$\text{МНК: } \dot{W}_t = 0.276 + 0.258\dot{P}_t + 0.046\dot{P}_{t-1} + 4.959\dot{V}_t \quad R^2 = 0.924;$$

$$\text{МНК: } \dot{P}_t = 2.693 + 0.232\dot{W}_t - 0.544\dot{x}_t + 0.247\dot{M}_t + 0.064\dot{M}_{t-1} \quad R^2 = 0.982;$$

$$\text{2МНК: } \dot{W}_t = 0.272 + 0.257\dot{P}_t + 0.046\dot{P}_{t-1} + 4.966\dot{V}_t \quad R^2 = 0.920;$$

$$\text{2МНК: } \dot{P}_t = 2.686 + 0.233\dot{W}_t - 0.544\dot{x}_t + 0.246\dot{M}_t + 0.046\dot{M}_{t-1} \quad R^2 = 0.981,$$

де \dot{W}_t , \dot{P}_t , \dot{M}_t і \dot{x}_t — процентна зміна зарплати, цін, імпорту, продуктивності праці (всі зміни стосуються попереднього року) і де V_t подає незайняті робочі вакансії (процент від загальної кількості зайнятих).

Оскільки результати МНК і 2МНК практично ідентичні, застосування методу 2МНК не має сенсу.

Прокоментуйте.

Вправа 6.26

Припустимо, що продуктивність характеризується виробничою функцією Кобба — Дугласа

$$Q_i = AK_i^\alpha L_i^\beta,$$

де Q — випуск; K — капіталовкладення; L — інтенсивність праці;

A , α , β — параметри i -ї фірми.

Маючи ціну кінцевого продукту P , ціну праці W та ціну капіталу R і припускаючи максимізацію прибутку, отримали таку емпіричну модель виробництва.

Функція виробництва:

$$\ln Q_i = \ln A + \alpha \ln K_i + \beta \ln L_i + \ln \varepsilon_{1i}.$$

Функція MPL — граничного продукту праці:

$$\ln Q_i = -\ln \beta + \ln L_i + \ln \frac{W}{P} + \ln \varepsilon_{2i}.$$

Функція MPC — граничного продукту капіталу:

$$\ln Q_i = -\ln \alpha + \ln K_i + \ln \frac{R}{P} + \ln \varepsilon_{3i},$$

де ε_1 , ε_2 , ε_3 — випадкові величини.

У попередній моделі є три рівняння з трьома ендогенними змінними Q , L та K ; P , R та W — екзогенні.

А. З якими проблемами можна зіткнутися при оцінці моделі, якщо $\alpha + \beta = 1$, тобто за умови постійного масштабу виробництва?

Б. Навіть якщо $\alpha + \beta \neq 1$, чи можна оцінити рівняння? Відповідайте, розглядаючи ототожненість.

В. Якщо система неототожнена, що можна зробити, щоб вона стала ототожненою?

Вправа 6.27

Розглянемо таку модель попиту та пропозиції на гроші.

Попит на гроші:

$$M_t^D = \beta_0 + \beta_1y_t + \beta_2R_t + \beta_3P_t + \varepsilon_{1t}.$$

Пропозиція грошей:

$$M_t^s = \alpha_0 + \alpha_1 y_t + \varepsilon_{2t},$$

де M — гроші; y — доход; R — ставка проценту; P — ціна.

Припустимо, що R і P — попередньо визначені.

А. Чи ототожнена функція попиту?

Б. Чи ототожнена функція пропозиції?

В. Який метод ви б використали для оцінки параметрів ототожнених рівнянь? Чому?

Г. Припустимо, ми модифікуємо функцію пропозиції, додаючи пояснючі змінні y_{t-1} і M_{t-1} . Що відбувається з проблемою ототожнення? Чи будете ви все ще використовувати метод, який застосовували у (В)? Чому так або чому ні?

Вправа 6.28

Розглянемо таку модель:

$$R_t = \beta_0 + \beta_1 M_t + \beta_2 y_t + \varepsilon_{1t};$$

$$y_t = \alpha_0 + \alpha_1 R_t + \varepsilon_{2t},$$

де M_t визначена ендогенно.

А. Як би ви пояснили модель?

Б. Чи ототожнені рівняння?

В. Використовуючи інформацію, наведену в табл. 6.5, оцініть параметри ототожнених рівнянь. Поясніть методи, які ви використовуєте.

Вправа 6.29

Припустимо, ми змінили модель вправи 6.36 таким чином:

$$R_t = \beta_0 + \beta_1 M_t + \beta_2 y_t + \beta_3 y_{t-1} + \varepsilon_{1t};$$

$$y_t = \alpha_0 + \alpha_1 R_t + \varepsilon_{2t}.$$

А. Перевірте, чи ототожнена система.

Б. Використовуючи дані таблиці 6.5, оцініть параметри ототожнених рівнянь.

Вправа 6.30

Розглянемо таку модель:

$$R_t = \beta_0 + \beta_1 M_t + \beta_2 y_t + \varepsilon_{1t};$$

$$y_t = \alpha_0 + \alpha_1 R_t + \alpha_2 I_t + \varepsilon_{2t}.$$

Змінні визначені у вправі 6.36. Вважаючи I (витрати на інвестиції) і M екзогенними, визначте ототожненість системи, використовуючи дані табл. 6.5. Оцініть параметри ототожнених рівнянь.

Вправа 6.31

Припустимо, що модель вправи 6.38 змінено таким чином:

$$R_t = \beta_0 + \beta_1 M_t + \beta_2 y_t + \varepsilon_{1t};$$

$$y_t = \alpha_0 + \alpha_1 R_t + \alpha_2 I_t + \varepsilon_{2t};$$

$$I_t = \gamma_0 + \gamma_1 R_t + \varepsilon_{3t}.$$

Припустимо, що M визначена екзогенно.

А. Визначте, які рівняння ототожнені.

Б. Оцініть параметри ототожнених рівнянь, використовуючи інформацію, наведену в табл. 6.5. Обґрунтуйте ваші методи.

ДОДАТОК 6.1

Зміщеність оцінок, отриманих методом непрямих найменших квадратів (ННК)

Для того, щоб показати, що ННК-оцінки узгоджені, зміщені, але спроможні, використаємо модель попиту та пропозиції, подану в рівняннях (6.83) та (6.84).

З (6.92) отримуємо:

$$b_1 = \frac{\hat{\pi}_3}{\hat{\pi}_1}.$$

Тепер $\hat{\pi}_3 = \frac{\sum \tilde{q}_t \tilde{x}_t}{\sum \tilde{x}_t^2}$ (із 6.89); $\hat{\pi}_1 = \frac{\sum \tilde{p}_t \tilde{x}_t}{\sum \tilde{x}_t^2}$ (із 6.87).

Підставивши, отримуємо:

$$b_1 = \frac{\sum \tilde{q}_t \tilde{x}_t}{\sum \tilde{p}_t \tilde{x}_t}. \quad (1)$$

Використавши (6.85) та (6.86), матимемо:

$$\tilde{p}_t = \pi_1 \tilde{x}_t + (w_t - \bar{w}); \quad (2)$$

$$\tilde{q}_t = \pi_3 \tilde{x}_t + (v_t - \bar{v}), \quad (3)$$

де \bar{w}, \bar{v} — середні значення w_t і v_t відповідно.

Підставляючи (2) і (3) в (1), отримуємо:

$$\begin{aligned} b_1 &= \frac{\pi_3 \sum \tilde{x}_t^2 + \sum (v_t - \bar{v}) \tilde{x}_t}{\pi_1 \sum \tilde{x}_t^2 + \sum (w_t - \bar{w}) \tilde{x}_t} = \\ &= \frac{\pi_3 + \sum (v_t - \bar{v}) \tilde{x}_t / \sum \tilde{x}_t^2}{\pi_1 + \sum (w_t - \bar{w}) \tilde{x}_t / \sum \tilde{x}_t^2}. \end{aligned} \quad (4)$$

Оператор математичного сподівання E є лінійним оператором, тому ми не можемо навести математичне сподівання (4), хоча зрозуміло, що в загальному випадку $b_1 \neq (\pi_3/\pi_1)$. Чому?

Але коли кількість спостережень прямує до нескінченності, ми можемо отримати:

$$p \lim(b_1) = \frac{p \lim \pi_3 + p \lim \sum (v_t - \bar{v})x_t / \sum x_t^2}{p \lim \pi_1 + p \lim \sum (w_t - \bar{w})x_t / \sum x_t^2}, \quad (5)$$

де використовуються такі властивості, як $p \lim(A+B) = p \lim A + p \lim B$ і $p \lim(A/B) = p \lim A / p \lim B$.

Оскільки кількість спостережень прямує до нескінченності, другий вираз у знаменнику і чисельнику прямує до нуля. Чому?

$$p \lim(b_1) = \frac{\pi_3}{\pi_1} \quad (6)$$

показує, що хоча b_1 зміщений, він є консистентною оцінкою β_1 .

ДОДАТОК 6.2

Оцінка стандартних відхилень методом 2МНК

Завдання цього додатка полягає в тому, щоб показати, що стандартні відхилення оцінок, отриманих на другому кроці регресії 2МНК, не є відповідними оцінками "реальних" стандартних відхилень. Щоб пересвідчитись у цьому, використаємо модель доходу та пропозиції грошей, наведену в (6.95) і (6.96). Оцінюємо параметри переототоженої функції пропозиції грошей на другому кроці:

$$y_{2t} = \beta_{20} + \beta_{21}\hat{y}_{1t} + \varepsilon_t^*, \quad (6.100)$$

де

$$\varepsilon_t^* = \varepsilon_{2t} + \beta_{21}e_t. \quad (7)$$

Тепер, коли ми працюємо з регресією (6.100), стандартне відхилення, скажімо $\hat{\beta}_{21}$, отримуємо з такого виразу:

$$\text{var}(\hat{\beta}_{21}) = \frac{\hat{\sigma}_{\varepsilon^*}^2}{\sum \hat{y}_{1t}^2}, \quad (8)$$

де

$$\hat{\sigma}_{\varepsilon^*}^2 = \frac{\sum (\hat{\varepsilon}_t^*)^2}{N-2} = \frac{\sum (y_{2t} - \hat{\beta}_{20} - \hat{\beta}_{21}\hat{y}_{1t})^2}{N-2}. \quad (9)$$

$\hat{\sigma}_{\varepsilon^*}^2$ — не те саме, що $\hat{\sigma}_{\varepsilon_2}^2$, котре є незміщеною оцінкою реальної дисперсії ε_2 . Це можна довести з (7). Щоб отримати реальну $\hat{\sigma}_{\varepsilon_2}^2$, виконаємо такі дії:

$$e_{2t} = y_{2t} - \hat{\beta}_{20} - \hat{\beta}_{21}y_{1t},$$

де $\hat{\beta}_{20}$ і $\hat{\beta}_{21}$ є оцінками регресії на другому кроці. Отже,

$$\hat{\sigma}_{\varepsilon_2}^2 = \frac{\sum (y_{2t} - \hat{\beta}_{20} - \hat{\beta}_{21}y_{1t})^2}{N-2}.$$

Зазначимо відмінність між (9) та (10): у рівнянні (10) ми використовували дійсне y_1 замість оціненого \hat{y}_1 за регресією першого кроку.

Маючи оцінену дисперсію (10), найлегшим способом коригування стандартних відхилень коефіцієнтів, оцінених на другому кроці, є перемноження кожного з них на $\hat{\sigma}_{\varepsilon_2} / \hat{\sigma}_{\varepsilon^*}$. Зазначимо: якщо y_{1t} і \hat{y}_{1t} дуже близькі, то $\hat{\sigma}_{\varepsilon_2} / \hat{\sigma}_{\varepsilon^*}$ буде близьке до 1.5. У цьому разі оцінені стандартні відхилення на другому кроці можуть вважатися правильними оцінками. Але в інших ситуаціях їх потрібно коригувати.

РОЗДІЛ 7. ЕЛЕМЕНТИ ТЕОРІЇ МАТРИЦЬ ТА ОСНОВНІ ЗАКОНИ РОЗПОДІЛУ ВИПАДКОВИХ ВЕЛИЧИН

7.1. ЕЛЕМЕНТИ ТЕОРІЇ МАТРИЦЬ

Цей розділ присвячено математичним і статистичним засобам, що використовуються в економетриці. Використовуючи матричну алгебру, основоположні твердження економетрики можна подати у чіткій і простій формі. Але, що важливіше, стає можливим узагальнення і спрощення результатів.

7.1.1. Основні визначення

Матриця — це упорядкований масив елементів, який позначається:

$$A = [a_{ik}] \text{ або } [A]_{ik}, \quad i = 1, 2, \dots, n, \quad k = 1, 2, \dots, m,$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}. \quad (7.1)$$

Перший індекс (i) елемента a_{ik} вказує на номер рядка, в якому розташований цей елемент, а другий індекс (k) — номер стовпця.

Приклад наведено у таблиці 7.1. У рядках таблиці наводяться параметри за певний рік, у стовпчиках — зміни параметрів у різні роки.

Вектор — це упорядкований набір даних, записаний у вигляді рядка або стовпця.

З огляду на викладене вище, вектор-рядок можна розглядати як матрицю з одним рядком, а вектор-стовпець — як матрицю з одним стовпцем. Тобто 5 величин, зазначених для 1993 року, у таблиці 7.1 становлять вектор-рядок, а величини споживання для конкретних років — вектор-стовпець.

Матрицю також можна розглядати як множину векторів-стовпців, що є природною інтерпретацією наведеного масиву даних, і, звичайно, як множину векторів-рядків.

Розмірністю (типом або порядком) матриці є кількість рядків та стовпців, що її складають. A є матрицею розмірністю $[n \times m]$ означає, що A складається з n рядків і m стовпців.

Таблиця 7.1

Рядок	Рік	Споживання (млн. дол.)	ВНП (млн. дол.)	ВНП дефлятор	Величина дисконтування
1	1987	737.1	1185.9	1.0000	4.50
2	1988	812.0	1326.4	1.0575	6.44
3	1989	808.1	1434.2	1.1508	7.83
4	1990	976.4	1549.2	1.2579	6.25
5	1991	1084.3	1718.0	1.3234	5.50
6	1992	1204.4	1918.3	1.4005	5.46
7	1993	1346.5	2163.9	1.5042	7.46
8	1994	1507.2	2417.8	1.6342	10.28
9	1995	1667.2	2633.1	1.7864	11.77

7.1.2. Особливі види матриць

1. **Прямокутна матриця** A — матриця розмірності $[n \times m]$.
2. **Квадратна матриця** A — матриця розмірності $[n \times n]$.
3. **Симетрична матриця** A — це матриця, у якій $a_{ik} = a_{ki}$ для всіх i та k . Наприклад,

$$A = \begin{bmatrix} 1 & 6 & 7 \\ 6 & 5 & 4 \\ 7 & 4 & 1 \end{bmatrix}.$$

4. **Діагональна матриця** — матриця, у якій всі її ненульові елементи містяться на головній діагоналі, що проходить з лівого верхнього кута у правий нижній.

$$A = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}.$$

5. **Скалярна матриця** — діагональна матриця, у якій всі діагональні елементи рівні між собою.

6. **Одинична матриця** — матриця з одиницями на діагоналі. Вона завжди позначається як I . Індекс іноді вказується, щоб зазначити її розмірність (або порядок).

Наприклад,

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (7.2)$$

7. Трикутна матриця — така квадратна матриця, у якої або нижче, або вище головної діагоналі розташовані всі нулі.

$$B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ 0 & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_{nn} \end{bmatrix}. \quad (7.3)$$

8. Нульова матриця — матриця, всі елементи якої дорівнюють нулю.

7.1.3. Алгебраїчні дії над матрицями

1. Рівність матриць

Матриці (або вектори) A і B називаються рівними, якщо вони мають одну й ту саму розмірність і кожний елемент матриці A дорівнює відповідному елементу матриці B .

2. Транспонування матриць

Транспонованою матрицею щодо матриці A називається матриця A' , яку можна отримати за рахунок заміни в матриці A стовпчиків на рядки і рядків на стовпчики, тобто

$$a'_{ij} = a_{ji}.$$

Транспонована матриця A позначається A' , визначається як матриця, k -ий рядок якої дорівнює k -ому стовпчику початкової матриці. Якщо розмірність матриці A дорівнює $[n \times m]$, то розмірність матриці A' дорівнює $[m \times n]$. Наприклад,

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 1 & 5 \\ 6 & 4 & 5 \\ 3 & 1 & 4 \end{bmatrix}, \quad A' = \begin{bmatrix} 1 & 5 & 6 & 3 \\ 2 & 1 & 4 & 1 \\ 3 & 5 & 5 & 4 \end{bmatrix}.$$

З означення симетричної матриці випливає, що якщо A — симетрична матриця, то

$$A' = A. \quad (7.4)$$

Для будь-якої матриці A

$$(A')' = A. \quad (7.5)$$

Транспонування вектора-стовпця дасть вектор-рядок:

$$a' = [a_1, a_2, \dots, a_n].$$

3. Додавання матриць

Сумою двох матриць однакової розмірності називається матриця такої ж розмірності, елементи якої дорівнюють сумам відповідних елементів матриць A і B , тобто

$$C = A + B = [a_{ik} + b_{ik}]. \quad (7.6)$$

Матриці не можна додавати, якщо вони різної розмірності. У разі однакової розмірності кажуть, що вони узгоджені для додавання.

За аналогією визначається операція віднімання матриць:

$$A - B = [a_{ik} - b_{ik}]. \quad (7.7)$$

Основні властивості операції додавання матриць

$$\bullet A + B = B + A \text{ (комутативність);} \quad (7.8)$$

$$\bullet (A + B) + C = A + (B + C) \text{ (асоціативність);} \quad (7.9)$$

$$\bullet (A + B)' = A' + B' \text{ (транспонування);} \quad (7.10)$$

$$\bullet A + (-A) = 0 \text{ (додавання протилежної матриці);}$$

$$\bullet A + 0 = A \text{ (додавання нульової матриці).}$$

4. Множення матриць

Матриці множаться за допомогою скалярного добутку. Скалярний добуток двох векторів a і b є скаляром і позначається так:

$$a'b = [a_1 b_1 + a_2 b_2 + \dots + a_n b_n]. \quad (7.11)$$

Зауважимо, що проміжний добуток записується як добуток транспонованого вектора a на вектор b , тобто вектора-рядка на вектор-стовпець.

Наприклад,

$$a'b = \begin{bmatrix} 3 & 5 & 4 \end{bmatrix} \begin{bmatrix} 8 \\ 2 \\ 1 \end{bmatrix} = 3(8) + 5(2) + 4(1) = 38. \quad (7.12)$$

У (7.11) кожний добуток $a_j b_j$ дорівнює $b_j a_j$, отже,

$$a'b = b'a. \quad (7.13)$$

Добутком двох матриць $A = [a_{ij}]$ типу $[n \times m]$ і $B = [b_{jk}]$ типу $[m \times p]$, заданих у певному порядку (A — перша, B — друга), називається матриця $C = [c_{ik}]$, елементи якої визначаються за правилом: елемент i -го рядка і k -го стовпчика матриці C дорівнює сумі добутків елементів i -го рядка матриці A і відповідних елементів k -го стовпчика матриці B , тобто

$$c_{ik} = \sum_{s=1}^m a_{is} b_{sk}, \quad i = 1, 2, \dots, n, \quad k = 1, 2, \dots, p.$$

Добуток матриць A і B позначається таким чином:

$$C = AB.$$

Майже у всіх наступних випадках будемо позначати i -ий стовпчик матриці a_i і a^i — i -ий рядок.

Для того, щоб перемножити дві матриці, необхідно, щоб кількість стовпчиків першої матриці дорівнювала кількості рядків другої. У такому випадку кажуть, що вони узгоджені для множення A на B .

Наприклад,

$$\begin{aligned} AB &= \begin{bmatrix} 2 & 3 & 1 \\ 4 & 5 & -1 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 0 & 6 \\ 1 & 5 \end{bmatrix} \\ &= \begin{bmatrix} 2(1) + 3(0) + 1(1) & 2(4) + 3(6) + 1(5) \\ 4(1) + 5(0) + (-1)(1) & 4(4) + 5(6) + (-1)(5) \end{bmatrix} \\ &= \begin{bmatrix} 3 & 31 \\ 3 & 41 \end{bmatrix}. \end{aligned}$$

Множення двох матриць, взагалі, не є комутативною операцією:

$$AB \neq BA.$$

Наприклад, у попередньому обчисленні AB є матриця розмірністю $[2 \times 2]$, у той час як BA має розмірність $[3 \times 3]$:

$$\begin{aligned} BA &= \begin{bmatrix} 1 & 4 \\ 0 & 6 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} 2 & 3 & 1 \\ 4 & 5 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 1(2) + 4(4) & 1(3) + 4(5) & 1(1) + 4(-1) \\ 0(2) + 6(4) & 0(3) + 6(5) & 0(1) + 6(-1) \\ 1(2) + 5(4) & 1(3) + 5(5) & 1(1) + 5(-1) \end{bmatrix} \\ &= \begin{bmatrix} 18 & 23 & -3 \\ 24 & 30 & -6 \\ 22 & 28 & -4 \end{bmatrix}. \end{aligned}$$

В інших випадках AB може існувати, а BA може бути невизначеною, і навіть якщо вони існують, то можуть бути різної розмірності, як у попередньому прикладі. Або, якщо вони і матимуть однакову розмірність, то не будуть рівними. З цієї точки зору вирізняємо множення зліва і множення справа. У добутку AB A буде помножена зліва на B , а B помножена справа на A .

Добуток матриці A і вектора b записується таким чином:

$$c = Ab.$$

Кількість елементів у b повинна дорівнювати кількості стовпців матриці A , а результатом буде вектор з кількістю елементів, що дорівнює кількості рядків у матриці A :

$$c = Ab = b_1 a_1 + b_2 a_2 + \dots + b_k a_k,$$

де a_i — i -й рядок матриці A , $i = 1, 2, \dots, k$.

$$\text{Наприклад: } A = \begin{pmatrix} 0 & 1 \\ 1 & 2 \\ 2 & 3 \end{pmatrix}, \quad b = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \quad c = Ab = \begin{pmatrix} 0(3) + 1(4) \\ 1(3) + 2(4) \\ 2(3) + 3(4) \end{pmatrix} = \begin{pmatrix} 4 \\ 11 \\ 18 \end{pmatrix}.$$

Розглянемо такі **загальні правила множення матриць**:

$$\bullet (AB)C = A(BC) \text{ (асоціативний закон);} \quad (7.14)$$

$$\bullet A(B + C) = AB + AC \text{ (дистрибутивний закон).} \quad (7.15)$$

Зверніть увагу на порядок множення матриць: BA і CA можуть бути невизначеними.

$$\bullet (AB)' = B'A' \text{ (транспонування добутку).} \quad (7.16)$$

Розширюючи це правило, отримаємо:

$$(ABC)' = C'B'A'. \quad (7.17)$$

- $AI = A$ (множення матриці на одиничну матрицю).
- $AO = O$ (множення на нульову матрицю).
- $cA = [ca_{ik}]$ (множення матриці A на скалярну величину). (7.18)

Множення матриці на скалярну величину еквівалентне множенню кожного елемента матриці на цю величину.

5. Запис сум у матричному вигляді

Матриці і вектори можна досить зручно використовувати для представлення сум різних величин. Корисним інструментом при цьому є вектор i , який складається із стовпчика одиниць. Суму елементів вектора x можна подати як:

$$\sum_{i=1}^n x_i = x_1 + x_2 + \dots + x_n = i'x. \quad (7.19)$$

Якщо всі елементи x дорівнюють одній і тій самій константі a , тоді $x = ai$

$$\sum_i x_i = i'(ai) = a(i'i) = na. \quad (7.20)$$

Для будь-якої константи a і вектора x

$$\sum_i ax_i = a \sum_i x_i = a i'x. \quad (7.21)$$

Якщо $a = 1/n$, отримаємо середнє арифметичне:

$$\bar{x} = \frac{1}{n} \sum_i x_i = \frac{1}{n} i'x. \quad (7.22)$$

З цього випливає, що

$$\sum_i x_i = i'x = n\bar{x}.$$

Суму квадратів елементів вектора x можна записати таким чином:

$$\sum_i x_i^2 = x'x, \quad (7.23)$$

а суму добутків елементів векторів x та y :

$$\sum_i x_i y_i = x'y. \quad (7.24)$$

За означенням множення матриць X' і X , елемент добутку

$$[X'X]_{ij} = [x'_i x_j] \quad (7.25)$$

є проміжним добутком i -го рядка та j -го стовпчика матриці X .

6. Рівнопотужна матриця

Основною матрицею статистики є матриця відхилень фактичних значень від середніх (матриця помилок). Введемо такі позначення:

$$i\bar{x} = \begin{bmatrix} \bar{x} \\ \bar{x} \\ \vdots \\ \bar{x} \end{bmatrix} = \frac{1}{n} ii'x. \quad (7.26)$$

Матриця $(1/n)ii'$ є матрицею $[n \times n]$, у якій кожен елемент дорівнює $(1/n)$. Вектор відхилень можна записати у вигляді:

$$\begin{bmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ \vdots \\ x_n - \bar{x} \end{bmatrix} = [x - i\bar{x}] = \left[x - \frac{1}{n} ii'x \right]. \quad (7.27)$$

Оскільки $x = Ix$,

$$\begin{aligned} \left[x - \frac{1}{n} ii'x \right] &= \left[Ix - \frac{1}{n} ii'x \right] \\ &= \left[I - \frac{1}{n} ii' \right] x \\ &= M^0 x. \end{aligned} \quad (7.28)$$

До речі, позначка M^0 буде використовуватись лише для цієї матриці. Діагональні елементи матриці M^0 дорівнюватимуть $(1-1/n)$, а недіагональні — $1/n$.

Матриця M^0 переважно використовується для обчислення суми квадратів відхилень, що дозволяє спростити деякі результати, наприклад:

$$M^0 i = \left[I - \frac{1}{n} i i' \right] i = i - \frac{1}{n} i(i' i) = 0.$$

З цього випливає, що $i' M^0 = 0'$. Тоді сума відхилень дорівнюватиме:

$$\sum_i (x_i - \bar{x}) = i' [M^0 x] = 0' x = 0. \quad (7.29)$$

Для певної змінної x сума квадратів відхилень від середнього дорівнює:

$$\sum_i (x_i - \bar{x})^2 = \sum_i (x_i^2 - 2\bar{x}x_i + \bar{x}^2) = \left(\sum_i x_i^2 \right) - n\bar{x}^2. \quad (7.30)$$

Або у матричному вигляді:

$$\sum_i (x_i - \bar{x})^2 = (x - \bar{x}i)' (x - \bar{x}i) = (M^0 x)' (M^0 x) = x' M^0 M^0 x. \quad (7.31)$$

У цьому разі є важливими такі дві властивості M^0 . По-перше, оскільки всі недиагональні елементи дорівнюють $1/n$, то матриця є симетричною. І по-друге, що можна легко довести, M^0 дорівнює своєму квадрату: $M^0 M^0 = M^0$.

За означенням, рівнопотужна матриця — це така матриця, яка дорівнює своєму квадрату, тобто $M^2 = M M = M$. Якщо M — симетрична рівнопотужна матриця, то $M' M = M$.

Тобто M^0 є рівнопотужною матрицею. Порівнюючи результати, отримаємо:

$$\sum_i (x_i - \bar{x})^2 = x' M^0 x.$$

Сформуємо матрицю квадратів відхилень та перехресних добутоків для двох векторів x та y

$$\sum_i (x_i - \bar{x})(y_i - \bar{y}) = (M^0 x)' (M^0 y). \quad (7.32)$$

Отже,

$$\begin{bmatrix} \sum_i (x_i - \bar{x})^2 & \sum_i (x_i - \bar{x})(y_i - \bar{y}) \\ \sum_i (x_i - \bar{x})(y_i - \bar{y}) & \sum_i (y_i - \bar{y})^2 \end{bmatrix} = \begin{bmatrix} x' M^0 x & x' M^0 y \\ y' M^0 x & y' M^0 y \end{bmatrix}. \quad (7.39)$$

Якщо з векторів x та y утворити матрицю $Z = [x, y]$ розмірністю $[n \times 2]$, то $M^0 Z$ буде матрицею відхилень розмірністю $[n \times 2]$.

Тоді $(M^0 Z)' (M^0 Z) = Z' M^0 M^0 Z = Z' M^0 Z$.

7.2. МАТРИЦІ І ВИЗНАЧНИКИ

7.2.1. Теорія визначників

1. Визначник n -го порядку

Нехай

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

квадратична матриця n -го порядку. Порівняємо матрицю A з певним числом, яке будемо називати визначником (детермінантом) цієї матриці, і позначимо його за допомогою одного з таких символів:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = |A| = \det A = \Delta.$$

Матриця A — певна таблиця чисел a_{ij} . Визначник $\det A$ — це число, яке знаходиться за встановленим правилом. Визначником матриці A назвемо число, яке розраховується на основі елементів цієї матриці за формулою:

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum (-1)^{r+s} a_{i_1 j_1} a_{i_2 j_2} \dots a_{i_n j_n},$$

де сума береться за всіма підстановками $\begin{pmatrix} i_1, i_2, \dots, i_n \\ j_1, j_2, \dots, j_n \end{pmatrix}$ з чисел $1, 2, \dots, n$;

r і s — число інверсій відповідно у верхньому та нижньому рядках підстановки.

Визначник матриці 2-го порядку обчислюється таким чином:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \sum (-1)^{r+s} a_{i_1 j_1} a_{i_2 j_2} = a_{11} a_{22} - a_{12} a_{21},$$

тобто визначник матриці 2-го порядку дорівнює різниці між добутком елементів, розташованих на головній та побічній діагоналях.

Цю формулу можна ілюструвати такою схемою:

$$\begin{vmatrix} \circ & \circ \\ \circ & \circ \end{vmatrix} = \begin{vmatrix} \bullet & \circ \\ \circ & \bullet \end{vmatrix} - \begin{vmatrix} \circ & \bullet \\ \bullet & \circ \end{vmatrix}$$

Визначник матриці 3-го порядку обчислюється так:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \sum (-1)^{r+s} a_{i_1 j_1} a_{i_2 j_2} a_{i_3 j_3} =$$

$$= a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - a_{13} a_{22} a_{31} - a_{12} a_{21} a_{33} - a_{11} a_{23} a_{32}.$$

Для обчислення визначника 3-го порядку слід користуватися правилом Саррюса (правилом трикутників), яке ілюструється схемою, наведеною на мал. 7.1.

Малюнок 7.1. Ілюстрація правила трикутників

Наприклад,

$$\begin{vmatrix} 1 & 3 & -2 \\ 2 & 0 & 3 \\ -2 & 1 & 2 \end{vmatrix} = 1 \cdot 0 \cdot 2 + 3 \cdot 3 \cdot (-2) + \\ + 2 \cdot 1 \cdot (-2) - (-2) \cdot 0 \cdot (-2) - 2 \cdot 3 \cdot 2 - 1 \cdot 3 \cdot 1 = -37.$$

2. Розкладання визначника за рядками або за стовпцями.

Алгебраїчні доповнення

Алгебраїчним доповненням A_{ij} елемента a_{ij} називається алгебраїчна сума добутків $(n-1)$ елементів визначника, взятих по одному з кожного рядка і стовпчика, за винятком того рядка і стовпчика, на перетині яких знаходиться a_{ij} .

З урахуванням поняття алгебраїчного доповнення визначник можна переписати таким чином:

$$\Delta = a_{1j} A_{1j} + a_{2j} A_{2j} + \dots + a_{nj} A_{nj} = \sum_{i=1}^n a_{ij} A_{ij}$$

(формула розкладання визначника Δ за елементами j -го стовпчика) або

$$\Delta = a_{i1} A_{i1} + a_{i2} A_{i2} + \dots + a_{in} A_{in} = \sum_{j=1}^n a_{ij} A_{ij} \quad (7.34)$$

(формула розкладання визначника Δ за елементами i -го рядка).

3. Мінори. Зв'язок доповнюючих мінорів з алгебраїчними доповненнями

Мінором k -го порядку визначника Δ називають визначник матриці, елементи якої знаходяться на перетині певних k ($k \leq n$) рядків і k стовпчиків.

Мінор k -го порядку позначається таким чином:

$$m = m_{j_1 j_2 \dots j_k}^{i_1 i_2 \dots i_k},$$

де i_1, i_2, \dots, i_k — номери вказаних рядків,

j_1, j_2, \dots, j_k — номери вказаних стовпчиків.

Доповнюючими мінорами для мінора m називаються визначники матриці $(n-k)$ -го порядку, який можна отримати з матриці A за рахунок викреслення тих k -стовпчиків і k -рядків, на перетині яких стоїть даний мінор m .

Доповнюючий мінор позначається так:

$$M = M_{j_1 j_2 \dots j_k}^{i_1 i_2 \dots i_k}.$$

Приклад

Нехай дано матрицю п'ятого порядку

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 10 \\ 12 & 1 & 5 & 10 & 9 \\ 13 & 0 & 7 & -11 & 3 \\ 4 & 8 & 2 & 6 & 12 \\ 5 & 1 & 4 & 10 & 1 \end{bmatrix}.$$

Виділяємо в ній 2 стовпчики — 3-й і 5-й, і два рядки — 1-й і 4-й. Тоді елементи матриці A , що знаходяться на перетині вказаних стовпців і рядків, створюють матрицю другого порядку, визначник якої являтиме собою мінор 2-го порядку даної матриці.

$$m = m_{35}^{14} = \begin{vmatrix} 3 & 10 \\ 2 & 12 \end{vmatrix}.$$

Доповнюючим мінором для мінора m буде визначник матриці 3-го порядку, що отримується з даної матриці викресленням 3-го і 5-го стовпців і 1-го та 4-го рядків.

$$M = M_{35}^{14} = \begin{vmatrix} 12 & 1 & 10 \\ 13 & 0 & -11 \\ 5 & 1 & 10 \end{vmatrix}.$$

Доповнюючі мінори для елементів a_{ij} матриці (або її визначника), тобто доповнюючі мінори виду

$$M = M_j^i$$

називаються просто мінорами цих елементів.

Теорема. Алгебраїчне доповнення A_{ij} елемента a_{ij} у визначнику матриці A і мінор M_j^i цього елемента зв'язані формулою:

$$A_{ij} = (-1)^{i+j} M_j^i. \quad (7.35)$$

З урахуванням цієї формули D можна переписати таким чином:

$$\Delta = \sum_{i=1}^n (-1)^{i+j} a_{ij} M_j^i \quad (7.36)$$

(формула розкладання визначника за елементами j -го стовпця);

$$\Delta = \sum_{j=1}^n (-1)^{i+j} a_{ij} M_j^i \quad (7.37)$$

(формула розкладання визначника за елементами i -го рядка).

Ці формули використовуються для практичного обчислення визначників (за рахунок зведення їх до визначників нижчого порядку).

Приклад. Обчислити визначник за формулою (7.36)

$$\Delta = \begin{vmatrix} 1 & -2 & 3 \\ 3 & 5 & -1 \\ 4 & 1 & 2 \end{vmatrix}.$$

Розкладемо його за елементами 2-го стовпця:

$$\begin{aligned} \Delta &= (-1)^{1+2}(-2) \begin{vmatrix} 3 & -1 \\ 4 & 2 \end{vmatrix} + (-1)^{2+2} \cdot 5 \begin{vmatrix} 1 & 3 \\ 4 & 2 \end{vmatrix} + (-1)^{3+2} \cdot 1 \begin{vmatrix} 1 & 3 \\ 3 & -1 \end{vmatrix} = \\ &= 2 \cdot 10 + 5 \cdot (-10) - 1 \cdot (-10) = -20. \end{aligned}$$

7.2.2. Матричні рівняння і характеристичні поліноми

1. Обернена матриця

Розглянемо квадратну матрицю

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

Ця матриця A називається *невиродженою*, або *несингулярною*, якщо її визначник відмінний від нуля.

Матриця називається *виродженою*, або *сингулярною*, якщо її визначник дорівнює нулю.

Квадратна матриця B називається *оберненою* для квадратної матриці A того ж порядку, якщо їх добуток

$$AB = BA = I,$$

де I — одинична матриця того ж порядку, що й матриці A і B .

Теорема. Для того, щоб матриця A мала обернену, необхідно і достатньо, щоб її визначник був відмінним від нуля [1].

Для обчислення оберненої матриці використовується формула:

$$B = \begin{bmatrix} \frac{A_{11}}{\Delta} & \frac{A_{21}}{\Delta} & \dots & \frac{A_{n1}}{\Delta} \\ \frac{A_{12}}{\Delta} & \frac{A_{22}}{\Delta} & \dots & \frac{A_{n2}}{\Delta} \\ \dots & \dots & \dots & \dots \\ \frac{A_{1n}}{\Delta} & \frac{A_{2n}}{\Delta} & \dots & \frac{A_{nn}}{\Delta} \end{bmatrix},$$

де A_{ij} — алгебраїчне доповнення елемента a_{ij} у визначнику матриці A .

Легко перевірити справедливість таких рівностей:

$$1) \det(A^{-1}) = (\det A)^{-1};$$

$$2) (A^{-1})^{-1} = A;$$

$$3) (A_1 A_2)^{-1} = A_2^{-1} A_1^{-1};$$

$$4) (A')^{-1} = (A^{-1})'.$$

Приклад

Знайти обернену матрицю для матриці

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}.$$

Оскільки визначник матриці

$$\det A = \Delta = \begin{vmatrix} 1 & -1 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{vmatrix} = 5 \neq 0,$$

то матриця A — невироджена.

Обчислимо алгебраїчні доповнення A_{ij} елементів a_{ij} :

$$A_{11} = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1, \quad A_{21} = -\begin{vmatrix} -1 & 1 \\ 1 & 2 \end{vmatrix} = 3, \quad A_{31} = \begin{vmatrix} -1 & 1 \\ 1 & 1 \end{vmatrix} = -2,$$

$$A_{12} = -\begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = -3, \quad A_{22} = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1, \quad A_{32} = -\begin{vmatrix} -1 & 1 \\ 2 & 1 \end{vmatrix} = 1,$$

$$A_{13} = \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 1, \quad A_{23} = -\begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} = -2, \quad A_{33} = \begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix} = 3.$$

Тоді обернена матриця дорівнюватиме

$$A^{-1} = \begin{bmatrix} \frac{1}{5} & \frac{3}{5} & -\frac{2}{5} \\ -\frac{3}{5} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{5} & -\frac{2}{5} & \frac{3}{5} \end{bmatrix}.$$

2. Матричні рівняння

Матричними рівняннями називаються рівняння виду $AX = B$, $XA = B$, де A і B — задані квадратні матриці n -го порядку; X — шукана матриця n -го порядку.

Розв'язком матричного рівняння називається матриця відповідного порядку, яка при підстановці в матричне рівняння замість матриці X перетворює рівняння в тотожність.

Матричні рівняння мають єдиний розв'язок, якщо $\det A \neq 0$: $X = A^{-1}B$.

Приклад

Розв'язати матричне рівняння

$$\begin{bmatrix} 3 & 4 \\ 1 & 1 \end{bmatrix} X = \begin{bmatrix} 2 & 9 \\ 1 & 3 \end{bmatrix}.$$

$\det A = -1 \neq 0$, тобто матриця $A = \begin{bmatrix} 3 & 4 \\ 1 & 1 \end{bmatrix}$ — невироджена.

$$A^{-1} = \begin{bmatrix} -1 & 4 \\ 1 & -3 \end{bmatrix} \Rightarrow X = \begin{bmatrix} -1 & 4 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 2 & 9 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ -1 & 0 \end{bmatrix}.$$

Якщо матриця A вироджена, тоді матричне рівняння або має нескінченно багато розв'язків, або взагалі не має розв'язку.

3. Характеристичний поліном

Нехай $A = [a_{ij}]$ — квадратна матриця розмірності $[n \times n]$.

Характеристичною матрицею матриці A називається матриця

$$A - \lambda I = \begin{bmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{bmatrix},$$

де λ — незалежна змінна, I — одинична матриця порядку $[n \times n]$.

Характеристичним поліномом матриці A називається визначник характеристичної матриці, який являє собою поліном n -го ступеня від λ .

Він позначається $\varphi(\lambda)$:

$$\varphi(\lambda) = |A - \lambda E| = (-1)^n \lambda^n + (-1)^{n-1} (a_{11} + a_{22} + \dots + a_{nn}) + \dots |A|.$$

Рівняння $\varphi(\lambda) = 0$, або $|A - \lambda E| = 0$ називається характеристичним рівнянням матриці A . Корені характеристичного рівняння матриці A називаються власними значеннями, або характеристичними коренями.

7.2.3. Ранг матриці і ранг системи векторів

1. Ранг матриці і його головні властивості

Розглянемо матрицю типу $[n \times m]$.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}.$$

Рангом матриці A називається найбільший з порядків мінорів даної матриці, які відмінні від нуля.

Позначається через $r(A)$.

Очевидно, що завжди виконуються співвідношення

$$0 \leq r \leq \min(m, n).$$

Властивості

1. При транспонуванні матриці її ранг не змінюється.
2. Ранг матриці не змінюється:
 - при перестановці її рядків або стовпців;
 - при множенні всіх елементів стовпця (рядка) на число, відмінне від нуля;
 - якщо до одного із стовпців (рядків) додати інший стовпець (рядок), ноживши його на певне число;
 - якщо викреслити з матриці стовпець, створений лише з нулів;
 - якщо викреслити з матриці стовпець, який є лінійною комбінацією інших стовпців.

2. Елементарні перетворення матриці. Обчислення рангу матриці

Елементарними перетвореннями матриці є:

- перестановка двох будь-яких стовпців (рядків);
- множення стовпця (рядка) на число, яке не дорівнює нулеві;
- додавання до одного стовпця (рядка) лінійної комбінації інших стовпців (рядків).

Із властивостей ранга матриці випливає, що **при елементарних перетвореннях матриці її ранг не змінюється.**

Дві матриці A і B називаються еквівалентними, якщо одну з них можна отримати з іншої за допомогою використання елементарних перетворень.

Еквівалентність матриці позначається таким чином: $A \sim B$.

Канонічною матрицею називається матриця, у якій на головній діагоналі стоять підряд декілька одиниць, а всі інші елементи дорівнюють нулю.

Наприклад,

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

За допомогою елементарних перетворень рядків і стовпчиків будь-яку матрицю можна звести до канонічної.

Ранг канонічної матриці дорівнює числу одиниць на її головній діагоналі.

Приклад

Знайти ранг матриці

$$A = \begin{bmatrix} 2 & 3 & 5 & -3 & -2 \\ 3 & 4 & 3 & -1 & -3 \\ 5 & 6 & -1 & 3 & -5 \end{bmatrix}.$$

Виконаємо дії.

1. Віднімемо від 2-го рядка 1-й рядок, переставимо рядки місцями:

$$\begin{bmatrix} 1 & 1 & -2 & 2 & -1 \\ 2 & 3 & 5 & -3 & -2 \\ 5 & 6 & -1 & 3 & -5 \end{bmatrix}.$$

2. Від 2 і 3 рядків відніmemo 1-й рядок, помножений відповідно на 2 і 5:

$$\begin{bmatrix} 1 & 1 & -2 & 2 & -1 \\ 0 & 1 & 9 & -7 & 0 \\ 0 & 1 & 9 & -7 & 0 \end{bmatrix}$$

3. Від 3-го рядка відніmemo 2-й рядок, отримаємо:

$$B = \begin{bmatrix} 1 & 1 & -2 & 2 & -1 \\ 0 & 1 & 9 & -7 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Матриця B еквівалентна матриці A , оскільки вона отримана з неї за допомогою кінцевої множини еквівалентних перетворень.

Ранг матриці B $r(B) = 2$, відповідно $r(A) = 2$.

3. Ранг системи векторів і його зв'язок з рангом матриці

Розглянемо систему n векторів-стовпців

$$X_1 = \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix}, X_2 = \begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{bmatrix}, \dots, X_n = \begin{bmatrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{nn} \end{bmatrix}$$

Теорема

Ранг системи векторів-стовпців (рядків) дорівнює рангу матриці, яка утворена з цих стовпців (рядків).

7.2.4. Комбіновані матриці

У формуванні елементів матриць, наприклад у структуруванні систем рівнянь, інколи корисно згрупувати деякі елементи у підматриці. Наприклад, можна записати:

$$A = \left[\begin{array}{cc|cc} 1 & 4 & 5 & \\ 2 & 9 & 3 & \\ \hline 8 & 9 & 6 & \end{array} \right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

A є комбінованою матрицею. Індекси підматриць визначаються таким же чином, як і елементи матриць. Загальновідомим спеціальним випад-

ком є блочно-діагональна матриця

$$A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix},$$

де A_{11} і A_{22} є квадратними матрицями.

1. Додавання і множення комбінованих матриць

Операції додавання і множення поширюються і на комбіновані матриці. Для комбінованих матриць A і B

$$A + B = \begin{bmatrix} A_{11} + B_{11} & \\ A_{21} + B_{21} & \end{bmatrix} \begin{bmatrix} A_{12} + B_{12} \\ A_{22} + B_{22} \end{bmatrix} \quad (7.38)$$

і

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix} \quad (7.39)$$

При цьому матриці мають бути узгодженими для виконання операцій. Відносно стовпців додавання розмірність A_{ij} і B_{ij} має бути однаковою. Для добутку кількість колонок у A_{ij} має дорівнювати кількості рядків у B_{ik} для всіх пар i та j . Таким чином, усі необхідні матричні добутки підматриць мають бути визначені.

Два випадки, що часто зустрічаються, мають форму:

$$\begin{bmatrix} A_1 \\ A_2 \end{bmatrix}' \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} A_1' & A_2' \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} A_1' A_1 + A_2' A_2 \end{bmatrix} \quad (7.40)$$

і

$$\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}' \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} = \begin{bmatrix} A_{11}' & A_{11} & 0 \\ 0 & A_{22}' & A_{22} \end{bmatrix} \quad (7.41)$$

2. Визначники комбінованих матриць

Визначник блочно-діагональної матриці обчислюється аналогічно до визначника діагональної матриці:

$$\begin{vmatrix} A_{11} & 0 \\ 0 & A_{22} \end{vmatrix} = |A_{11}| \cdot |A_{22}| \quad (7.42)$$

Результатом для загальної комбінованої матриці розмірністю $[2 \times 2]$ буде:

$$\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} = |A_{22}| \cdot |A_{11} - A_{12}A_{22}^{-1}A_{21}| = |A_{11}| \cdot |A_{22} - A_{21}A_{11}^{-1}A_{12}|. \quad (7.43)$$

3. Обернення комбінованих матриць

Оберненою до блочно-діагональної матриці є:

$$\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & A_{22}^{-1} \end{bmatrix}, \quad (7.44)$$

що може бути перевірено шляхом прямого множення.

Для загальної комбінованої матриці розмірністю $[2 \times 2]$ однією формою комбінованого обернення буде:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1}(I + A_{12}F_2A_{21}A_{11}^{-1}) & -A_{11}^{-1}A_{12}F_2 \\ -F_2A_{21}A_{11}^{-1} & F_2 \end{bmatrix}, \quad (7.45)$$

де

$$\Phi_2 = (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}.$$

Це можна перевірити найпростішим способом — шляхом множення A на обернену матрицю. З точки зору симетрії обчислень, верхній лівий блок може також бути записаний як

$$\Phi_1 = (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1}.$$

4. Відхилення від середніх

Застосуємо попередні висновки у подальших обчисленнях. Припустимо, що ми починаємо з вектора-стовпця з n значень x і нехай

$$A = \begin{bmatrix} n & \sum_i x_i \\ \sum_i x_i & \sum_i x_i^2 \end{bmatrix} = \begin{bmatrix} i'i & i'x \\ x'i & x'x \end{bmatrix}.$$

Нас цікавить нижній правий елемент матриці A^{-1} . Використовуючи визначення Φ_2 з (7.45), отримуємо:

$$\Phi_2 = [x'x - (x'i)(i'i)^{-1}(i'x)]^{-1} = \left\{ x' \left[Ix - i \left(\frac{1}{n} \right) i'x \right] \right\}^{-1} = \left\{ x' \left[I - \left(\frac{1}{n} \right) ii' \right] x \right\}^{-1} =$$

$$= [x'M^0x]^{-1}.$$

Таким чином, правий нижній елемент оберненої матриці буде:

$$[x'M^0x]^{-1} = \frac{1}{\sum_i (x_i - \bar{x})^2} = a^{22}.$$

Тепер припустимо, що замість єдиного стовпця x ми маємо матрицю X з декількома стовпцями. Ми шукаємо нижній правий блок $[Z'Z]^{-1}$, де $Z = [i, X]$. Аналогічним результатом буде:

$$(Z'Z)^{22} = [X'X - X'i(i'i)^{-1}i'X]^{-1} = [X'M^0X]^{-1},$$

і це означає, що матриця $[k \times k]$ в нижньому правому кутку $(Z'Z)^{-1}$ є оберненою до матриці розмірністю $[k \times k]$, jk -й елемент якої дорівнює $\sum_i (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)$.

Таким чином, якщо матриця даних має стовпець ls , елементи оберненої матриці з суми квадратів і перехресних добуток будуть підраховані з первинних даних у формі відхилень від відповідних середніх значень стовпців.

7.3. ГЕОМЕТРИЧНА ІНТЕРПРЕТАЦІЯ ВЕКТОРІВ І МАТРИЦЬ

Матрична алгебра надзвичайно корисна у розв'язуванні систем лінійних рівнянь. У той же час алгебраїчні результати мають геометричну основу, необхідну для розуміння розрахунків. Перед розглядом математичних результатів доцільно зробити відступ для геометричної інтерпретації векторів і матриць.

7.3.1. Вектор, операції над векторами, векторний простір

Вектор — це лінійний геометричний об'єкт, який характеризується довжиною і напрямком.

k -елементний вектор-стовпець

$$a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_k \end{bmatrix}$$

може розглядатися як координати точки у k -вимірному просторі (як показано на мал. 7.2 для двовимірного простору).

Основні операції над векторами

Можливі дві основні дії над векторами, а саме добуток на скаляр та додавання.

1. Добуток вектора на скаляр

Добутком вектора a на скаляр λ є інший вектор, нехай це буде a^* , координатами якого є помножені на скаляр координати вектора a :

$$a^* = \lambda a.$$

Таким чином, на мал. 7.2

$$a = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad a^* = 2a = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \quad \text{і} \quad a^{**} = -\frac{1}{2}a = \begin{bmatrix} -\frac{1}{2} \\ -1 \end{bmatrix}.$$

Множиною всіх можливих добутків вектора a на скаляр є пряма, що проходить через точку O та a . Будь-який добуток a на скаляр є відрізком цієї прямої.

2. Додавання векторів

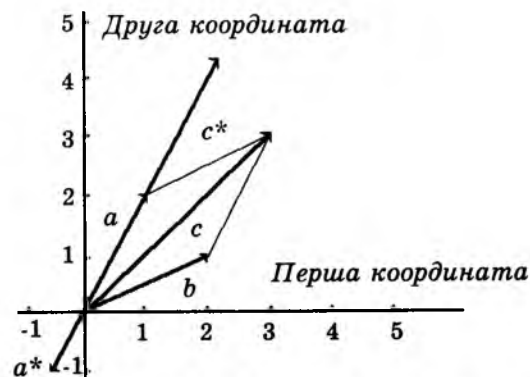
Сумою двох векторів a і b є третій вектор, чії координати є сумами відповідних координат a і b . Наприклад,

$$c = a + b = \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}.$$

З геометричної точки зору, c отримується зміщенням з початку вектора a на величину та у напрямі, визначеному b (додавання є комунікативним: зміщення з початку вектора b на величину і у напрямі, визначеному a). Це і є геометричною інтерпретацією множення вектора на скаляр і додавання векторів.

Означення. Векторний простір — це будь-яка множина векторів, замкнена відносно операцій додавання векторів та множення вектора на скаляр.

Розмірністю простору X називається максимальне число лінійно незалежних векторів простору X .



Малюнок 7.2. Векторний простір

Прикладом векторного простору є двовимірна координатна площина R^2 , тобто множина двох векторів з двома дійсними координатами.

Інший приклад — множина дійсних чисел, тобто R^1 , що є множиною векторів з однією дійсною координатою. Взагалі множина k -координатних векторів, кожна координата яких є дійсним числом, є k -вимірним векторним простором і позначається R^k .

7.3.2. Лінійна комбінація векторів. Базисні вектори. Лінійна залежність векторів

Означення. Вектор x називається лінійною комбінацією векторів a_1, \dots, a_n , якщо може бути поданий у вигляді $x = \alpha_1 a_1 + \alpha_2 a_2 + \dots + \alpha_n a_n$, де $\alpha_1, \alpha_2, \dots, \alpha_n$ — деякі числа.

На мал. 7.3 $c = a + b$ і $d = a^* + b$. Але оскільки $a^* = 2a$, то $d = 2a + b$.

Також $e = a + 2b$ і $f = b + (-a) = b - a$. Як допускає ця вправа, будь-який вектор у R^2 може бути отриманий як лінійна комбінація a і b .

Означення. Множина векторів у векторному просторі є **базисом** для даного векторного простору, якщо будь-який вектор цього векторного простору може бути поданий як лінійна комбінація цих векторів.

Як видно з мал. 7.3, будь-яка пара двоелементних векторів, включаючи a і b , що вказують різні напрямки,

утворюють базис для R^2 . Розглянемо довільну множину векторів у R^2 : a, b та c . Якщо a та b є базисом, то ми можемо знайти числа α_1 та α_2 , такі що $c = \alpha_1 a + \alpha_2 b$.

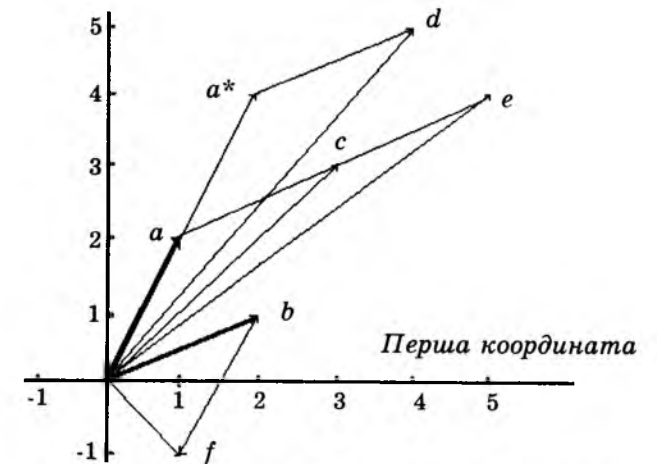
Нехай

$$a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad \text{і} \quad c = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

Тоді

$$\begin{aligned} c_1 &= \alpha_1 a_1 + \alpha_2 b_1, \\ c_2 &= \alpha_1 a_2 + \alpha_2 b_2. \end{aligned} \tag{7.46}$$

Друга координата



Малюнок 7.3. Лінійна комбінація векторів

Розв'язок цієї пари рівнянь:

$$\alpha_1 = \frac{b_2 c_1 - b_1 c_2}{a_1 b_2 - b_1 a_2}; \quad (7.47)$$

$$\alpha_2 = \frac{a_1 c_2 - a_2 c_1}{a_1 b_2 - b_1 a_2}.$$

Розв'язок буде єдиним, якщо $(a_1 b_2 - b_1 a_2) \neq 0$. Якщо $(a_1 b_2 - b_1 a_2) = 0$, тоді $a_1/a_2 = b_1/b_2$, що означає, що вектор b дорівнює вектору a , помноженому на скаляр. Це повертає нас до вихідної умови, що a та b різноспрямовані. Мається на увазі, що коли a та b — будь-яка пара векторів, для яких знаменник у (7.47) відмінний від 0, тоді довільний інший вектор c може бути складений як єдина лінійна комбінація a та b . Базис векторного простору не є єдиним, оскільки ним можуть бути всі множини векторів, що підходять під означення. Але для певного базису є лише одна лінійна комбінація його векторів, що дасть інший певний вектор цього простору.

На малюнку 7.3 вектори a і b створюють базис в R^2 , чого не можна сказати про вектори a і a^* . Це можна пояснити тим, що a і b лінійно незалежні вектори, а вектори a і a^* — лінійно залежні.

Означення. Множина векторів є лінійно залежною, якщо будь-який вектор з цієї множини може бути записаний як лінійна комбінація інших.

Наприклад, якщо

$$a = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, b = \begin{bmatrix} 3 \\ 3 \end{bmatrix}, c = \begin{bmatrix} 10 \\ 14 \end{bmatrix},$$

тоді можна записати:

$$2a + b - \frac{1}{2}c = 0.$$

З цього випливає, що вектори a , b і c — лінійно залежні.

Означення. Множина векторів є лінійно незалежною тоді і тільки тоді, коли єдиним розв'язком системи $\alpha_1 a_1 + \alpha_2 a_2 + \dots + \alpha_k a_k = 0$ є $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$.

Це еквівалентно означенню базису.

7.4. СИСТЕМИ ЛІНІЙНИХ РІВНЯНЬ

Розглянемо систему n лінійних рівнянь

$$Ax = b, \quad (7.48)$$

у якій k елементів матриці X є невідомими. A — відома матриця коефіцієнтів, a b — певний вектор значень. Наша мета — визначити, чи існують розв'язки; якщо вони існують, то як їх визначити; і нарешті, чи є розв'язок даної системи єдиним.

У більшості наших прикладів ми будемо використовувати квадратні системи рівнянь, тобто ті, у яких матриця A є квадратною. Тому далі ми будемо розглядати випадки, коли n дорівнює k . Оскільки кількість рядків дорівнює кількості рівнянь, а кількість стовпців є кількістю змінних, це найпоширеніший випадок “ n рівнянь з n невідомими”.

Є два типи систем.

1. Однорідні системи: $Ax = 0$

За означенням, ненульовий розв'язок такої системи існуватиме тоді і лише тоді, коли A не має повного рангу. Якщо це справджується, то хоча б для одного стовпця A можемо записати таку рівність:

$$a_j = -\sum_{i \neq j} \frac{x_i}{x_j} a_i.$$

Це означає, як ми знаємо, що стовпці A лінійно залежні і що $|A| = 0$.

2. Неоднорідні системи: $Ax = b$

Вектор b є довільним і має бути виражений як лінійна комбінація стовпців A . Оскільки b має k елементів, це буде можливо лише за умови, що лінійною оболонкою стовпців A є k -вимірний простір R^k .

Це еквівалентно вимозі, щоб стовпці A були лінійно незалежними, або щоб визначник матриці A не дорівнював 0.

Щоб розв'язати систему $Ax = b$ відносно x , необхідно виконати дію, подібну до ділення на матрицю.

Розглянемо знаходження матриці, оберненої до даної. Для матриці $[2 \times 2]$ з умови

$$AB = I$$

випливає, що

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

або

$$\begin{aligned} a_{11} b_{11} + a_{12} b_{21} &= 1; \\ a_{11} b_{12} + a_{12} b_{22} &= 0; \end{aligned}$$

$$a_{21}b_{11} + a_{22}b_{21} = 0;$$

$$a_{21}b_{12} + a_{22}b_{22} = 1.$$

Розв'язками є:

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} = \frac{1}{|A|} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}. \quad (7.49)$$

Звернімо увагу на наявність оберненої величини до визначника A . Це не характерне лише для випадку з матрицею розмірністю $[2 \times 2]$. На основі цього ми можемо стверджувати, що коли визначник матриці дорівнює 0, то оберненої до неї матриці немає.

Для неоднорідних систем

$$Ax = b,$$

якщо A несингулярна, єдиним розв'язком буде

$$x = A^{-1}b.$$

7.5. ХАРАКТЕРИСТИЧНІ КОРЕНІ І ВЕКТОРИ, ХАРАКТЕРИСТИЧНЕ РІВНЯННЯ

Корисна група результатів для аналізу квадратних матриць A витікає з розв'язання системи рівнянь

$$Ac = \lambda c. \quad (7.50)$$

Пара розв'язків є *характеристичними числами (векторами)* c і характеристичними коренями λ . Якщо c є будь-яким вектором розв'язків, вектор kc є ним також для будь-якого k . Для усунення невизначеності c *нормалізується*, так що

$$c'c = I.$$

Тоді розв'язок складається з λ і $n-1$ невідомих елементів c .

7.5.1. Характеристичне рівняння

В загальному випадку розв'язок (7.50) можна продовжити таким чином:

по-перше, (7.50) означає, що

$$Ac = \lambda Ic$$

або

$$|A - \lambda I|c = 0.$$

Маємо однорідну систему, що має ненульовий розв'язок c , тільки якщо матриця $(A - \lambda I)$ є виродженою або має нульовий визначник. Таким чином, якщо λ є розв'язком,

$$|A - \lambda I|c = 0. \quad (7.51)$$

Цей поліном від λ є *характеристичним рівнянням* матриці A . Наприклад, якщо

$$A = \begin{bmatrix} 5 & 1 \\ 2 & 4 \end{bmatrix},$$

тоді

$$|A - \lambda I| = \begin{vmatrix} 5 - \lambda & 1 \\ 2 & 4 - \lambda \end{vmatrix} = (5 - \lambda)(4 - \lambda) - 2 = \lambda^2 - 9\lambda + 18.$$

Двома розв'язками є $\lambda = 3$ і $\lambda = 6$.

При розв'язанні характеристичного рівняння немає гарантії, що характеристичні корені будуть дійсними числами. В попередньому прикладі якщо замість 2 у нижньому лівому куті матриці буде -2 , розв'язком буде пара комплексних чисел. Така ж проблема може виникнути в загальному випадку $[n \times n]$. Характеристичні корені симетричної матриці є, однак, дійсними числами.

Для матриці розмірністю $[n \times n]$ характеристичне рівняння є поліномом n -го порядку. Його розв'язками можуть бути n різних значень, як у попередньому прикладі; вони можуть вміщувати повторювані значення λ , як у випадку

$$\begin{vmatrix} 2 - \lambda & 0 \\ 0 & 2 - \lambda \end{vmatrix} = (2 - \lambda)^2 \Rightarrow \lambda_1 = \lambda_2 = 2 \text{ або } -2;$$

або можуть також вміщувати декілька нулів, як у

$$\begin{vmatrix} 1 - \lambda & 2 \\ 2 & 4 - \lambda \end{vmatrix} = \lambda^2 - 5\lambda = 0 \Rightarrow \lambda_1 = 5 \text{ і } \lambda_2 = 0.$$

7.5.2. Характеристичні вектори

Маючи λ , характеристичні вектори можна вивести з первинної задачі

$$Ac = \lambda c,$$

або

$$(A - \lambda I)c = 0.$$

Для першого прикладу маємо:

$$\begin{bmatrix} 5 - \lambda & I \\ 2 & 4 - \lambda \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (7.52)$$

Двома значеннями для $\lambda \in 6$ і 3 . Тоді

- 1) для $\lambda = 6$, $-c_1 + c_2 = 0$ і $2c_1 - 2c_2 = 0$, або $c_1 = c_2$;
- 2) для $\lambda = 3$, $2c_1 + c_2 = 0$ і $2c_1 + c_2 = 0$, або $c_1 = (1/2)c_2$.

Жодна пара не визначає значення c_1 і c_2 . Але цей результат очікувався. Причиною було те, що $c's = I$. Так, наприклад, якщо $\lambda = 6$, будь-який вектор c з рівними елементами задовольнятиме умову (7.52). Однак, додаткове рівняння $c's = I$ призводить до повних розв'язків для обох векторів.

$$1. \text{ Для } \lambda = 6, c = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}.$$

$$2. \text{ Для } \lambda = 3, c = \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}.$$

Ці вектори — єдині, що задовольняють як (7.52), так і умову $c's = I$.

7.5.3. Загальні результати для характеристичних коренів і векторів

Симетрична матриця розмірністю $[k \times k]$ має k різних характеристичних векторів c_1, c_2, \dots, c_k . Відповідні характеристичні корені $\lambda_1, \lambda_2, \dots, \lambda_k$, якщо вони дійсні, не обов'язково відмінні. Характеристичні вектори симетричної матриці є ортогональними. Ортогональність характеристичних векторів означає, що для кожних $i \neq j$ виконується умова $c_i'c_j = 0$. Ця умова не виконується, якщо матриця несиметрична. Для несиметричних матриць є також різниця між "правими" характеристичними векторами $Ac = \lambda c$ і "лівими" характеристичними векторами $d'A = \lambda d'$, які можуть не бути рівними.

Зручно звести характеристичні вектори у матрицю $[k \times k]$, i -й стовпець якої є c_i , що відповідає λ_i :

$$C = [c_1 \ c_2 \ \dots \ c_k].$$

Зведемо характеристичні корені в такому ж порядку в діагональну матрицю

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ & & & \ddots & \\ 0 & 0 & 0 & \dots & \lambda_k \end{bmatrix}.$$

Тоді повна система рівнянь

$$Ac_i = \lambda_i c_i$$

міститься у

$$AC = C\Lambda. \quad (7.53)$$

Через те, що вектори ортогональні і $c_i'c_j = I$, маємо

$$C'C = \begin{bmatrix} c_1'c_1 & c_1'c_2 & \dots & c_1'c_k \\ c_2'c_1 & c_2'c_2 & \dots & c_2'c_k \\ & & \ddots & \\ c_k'c_1 & c_k'c_2 & \dots & c_k'c_k \end{bmatrix} = I. \quad (7.54)$$

Результат (7.54) означає, що

$$C' = C^{-1}. \quad (7.55)$$

Таким же чином

$$CC' = CC^{-1} = I, \quad (7.56)$$

тобто і рядки C також ортогональні.

7.5.4. Приведення матриці до діагональної форми

Шляхом попереднього множення (7.53) на C і використання (7.54) отримуємо *діагональну форму* L :

$$C'AC = C'CL = I\Lambda = \Lambda. \quad (7.54)$$

7.5.5. Ранг характеристичної матриці

За допомогою діагональної форми матриці можна дуже просто отримати ранг матриці A . Щоб зробити це, ми можемо використати таке твердження про ранг добутку матриць.

Для будь-якої матриці A і невироджених матриць B і C ранг BAC дорівнює рангові A .

Оскільки C і C' невироджені, можемо використати їх для застосування цього результату до (7.54). Шляхом очевидних заміни отримаємо

$$\text{rank}(A) = \text{rank}(\Lambda).$$

Знаходження рангу Λ є тривіальним. Оскільки Λ — діагональна матриця, її рангом є кількість ненульових значень на її діагоналі. Тоді загальний висновок такий:

$$\text{рангом симетричної матриці є кількість ненульових характеристичних коренів, які вона має.} \quad (7.55)$$

Здається, що це просте правило не буде корисним, якщо A не квадратна. Але згадаємо, що

$$\text{rank}(A) = \text{rank}(A'A).$$

Оскільки $A'A$ завжди квадратна, ми можемо використовувати її замість A . Звичайно, можна її використовувати навіть, коли A квадратна, що веде до такого загального результату:

$$\text{ранг будь-якої матриці } A \text{ дорівнює кількості ненульових характеристичних коренів у } A'A. \quad (7.56)$$

Цей результат поширений, бо дає простий шлях знаходження рангу несиметричної матриці. З обчислювальної точки зору знаходження характеристичних коренів несиметричних матриць досить складне і потребує використання арифметики комплексних чисел. Але симетричні матриці через свої дійсні корені аналізуються значно легше.

Оскільки ранг рядків і ранг стовпців матриці однакові, ми можемо також застосувати (7.56) до $A'A$. Тоді отримуємо такий результат:

$$\text{ненульові характеристичні корені } AA' \text{ є такі, як у } A'A. \quad (7.57)$$

7.5.6. Власні значення матриці

Формально власним значенням для квадратної матриці A є

$$\gamma = [\text{максимальний корінь/мінімальний корінь}]^{1/2}.$$

Для неквадратної матриці X використовуємо $A = X'X$. Для подальшого вдосконалення, враховуючи, що на характеристичні корені впливає масштабування стовпців X , масштабуємо стовпці так, щоб вони мали одиничну довжину (шляхом ділення кожної колонки на її норму). Припустимо, що для X найбільшим характеристичним коренем A є 4.9255 і найменшим 0.0001543, тобто власне значення 178.67 є надзвичайно великим. Якщо корінь близький до нуля у порівнянні з найбільшим, це означає, що матриця майже вироджена. Матриці з великими власними значеннями важко точно обертати.

7.5.7. Слід матриці

Слідом квадратної матриці є сума її діагональних елементів:

$$\text{tr}(A) = \sum_i a_{ii}. \quad (7.58)$$

Найпростішими результатами є:

$$\text{tr}(cA) = c(\text{tr}(A)), \quad (7.59)$$

$$\text{tr}(A') = \text{tr}(A), \quad (7.60)$$

$$\text{tr}(A+B) = \text{tr}(A) + \text{tr}(B), \quad (7.61)$$

$$\text{tr}(I_k) = k. \quad (7.62)$$

Винятково корисним результатом є слід добутку матриць:

$$\text{tr}(AB) = \text{tr}(BA). \quad (7.63)$$

Вираз (7.63) можна використати для знаходження добутків матриць:

$$a'a = \text{tr}(a'a) = \text{tr}(aa')$$

i

$$\text{tr}(ABCD) = \text{tr}(BCDA) = \text{tr}(CDAB) = \text{tr}(DABC). \quad (7.64)$$

Використовуючи (7.64), отримуємо

$$\text{tr}(C'AC) = \text{tr}(ACC') = \text{tr}(AI) = \text{tr}(A) = \text{tr}(A). \quad (7.65)$$

Оскільки A є діагональною матрицею з коренями A на її діагоналі, загальний результат такий:

$$\text{слід матриці дорівнює сумі її характеристичних коренів.} \quad (7.66)$$

7.5.8. Визначник матриці

Знаючи, яким стомлюючим може бути обчислення детермінанта, ми вважаємо подальші викладки винятково корисними. Оскільки

$$\begin{aligned} C'AC &= A, \\ |C'AC| &= |A|. \end{aligned}$$

Використовуючи декілька попередніх результатів, маємо для ортогональної матриці C :

$$\begin{aligned} |C'AC| &= |C'| |A| |C| = |C'| |C| |A| = |C'C| |A| \\ &= |I| |A| = 1 |A| = |A| = |A|. \end{aligned}$$

Оскільки $|A|$ є якраз добутком її ортогональних елементів, це означає, що визначник матриці дорівнює добутку її характеристичних коренів. (7.67)

Зауважимо, що очікуваний результат буде лише за умови, якщо будь-який з цих коренів нульовий. Оскільки визначник є добутком коренів, це означає, що матриця вироджена тоді і тільки тоді, коли її визначник дорівнює нулю, і в свою чергу тоді й тільки тоді, коли вона має хоча б один нульовий характеристичний корінь.

7.5.9. Спектральна декомпозиція матриць

Починаючи знову з $AC = CA$, маємо $ACC' = SAC'$ або

$$A = SAC'. \quad (7.68)$$

Це називається спектральною декомпозицією матриці A . (7.68) можна записати таким чином:

$$A = \sum_{i=1}^n \lambda_i c_i c_i'. \quad (7.69)$$

7.5.10. Степені матриці

Ми часто використовуємо вирази, що містять степені матриць, такі як $AA = A^2$. Для додатних цілих показників степені це можна обчислити шляхом повторного множення. Однак, таке обчислення не показує, як вирішувати цю проблему, коли, наприклад, треба знайти таке B , щоб $B^2 = A$, тобто знайти квадратний корінь матриці. Характеристичні корені і вектори забезпечують просте розв'язання.

Розглянемо спочатку

$$\begin{aligned} AA &= A^2 = (SAC')(SAC') = SAC'SAC' \\ &= SAIAS' = SAAS' = SA^2C'. \end{aligned} \quad (7.70)$$

Оскільки A^2 є діагональною матрицею, ненульові елементи якої є квадратами таких же елементів у A , це означає, що

$$\begin{aligned} &\text{для будь-якої симетричної матриці } A \\ &\text{характеристичні корені } A^2 \text{ є квадратами} \\ &\text{відповідних коренів } A \text{ і характеристичні вектори} \\ &\text{в них однакові.} \end{aligned} \quad (7.71)$$

Останній рядок у (7.70) є спектральною декомпозицією матриці $B = AA$. Оскільки $A^3 = AA^2$ і т. д., (7.71) поширюється для будь-яких додатних цілих чисел. За домовленістю для будь-якої A , $A^0 = I$. Таким чином, для будь-якої симетричної матриці A , $A^k = SA^kC'$, $k = 0, 1, \dots$. Надалі характеристичними коренями A^k є λ^k , тоді як характеристичні вектори такі ж, як у A . Якщо A невироджена, тобто всі її корені λ_i є ненульовими, це твердження можна поширити також на від'ємні степені. Якщо A^{-1} існує, тоді

$$A^{-1} = (SAC')^{-1} = (C')^{-1}A^{-1}C^{-1} = SA^{-1}C', \quad (7.72)$$

де ми використали отриманий раніше результат $C' = C^{-1}$. Якщо існує A^{-1} , характеристичні корені A^{-1} і характеристичні вектори є відповідними до характеристичних коренів A . Шляхом розширення поняття повторного множення ми маємо:

для будь-якої невиродженої симетричної матриці A

$$A^k = SA^kC', \quad k = \dots, -2, -1, 0, 1, 2, \dots \quad (7.73)$$

Розглянемо загальну проблему обчислення квадратного кореня матриці. У скалярному випадку значення має бути невід'ємним. Аналогом цієї вимоги для матриць будуть невід'ємні всі її характеристичні корені. Розглянемо

$$A^{1/2} = SA^{1/2}C' = C \begin{bmatrix} \sqrt{\lambda_1} & 0 & \dots & 0 \\ 0 & \sqrt{\lambda_2} & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & \sqrt{\lambda_n} \end{bmatrix} C'. \quad (7.74)$$

Це задовольняє вимогу для квадратного кореня, бо

$$A^{1/2}A^{1/2} = SA^{1/2}C'SA^{1/2}C' = SAC' = A.$$

Якщо продовжимо міркування таким шляхом, ми зможемо визначити степені матриці більш загально, все ще припускаючи, що всі характеристичні корені невід'ємні. Наприклад, $A^{1/3} = SA^{1/3}C'$. Якщо всі корені строго додатні, можна зробити ще один крок і розширити результат на будь-який реальний показник степені. Матриця з додатними характеристичними коренями називається додатно визначеною.

$$\text{Для додатно визначеної матриці } A, \quad A^r = SA^rC' \quad (7.75)$$

для будь-якого дійсного числа r . Характеристичні корені A^r є характеристичними коренями A у r -ій степені, те ж саме можна сказати про характеристичні вектори.

Якщо A тільки невід'ємно визначена, тобто має корені, що є нульовими, або додатними, (7.75) виконується тільки для невід'ємних r .

7.5.11. Рівнопотужні матриці

Рівнопотужні матриці — це матриці, які дорівнюють своєму квадратові. Розглянемо деякі результати, що стосуються рівнопотужних матриць, з точки зору їхньої важливості в економетриці. По-перше, з (7.71) випливає, що коли λ — характеристичний корінь рівнопотужної матриці, то $A = \lambda^k$ для всіх цілих невід'ємних k . У цьому разі, оскільки A — симетрична рівнопотужна матриця, всі її корені дорівнюють 1 або 0. Припустимо, що всі її корені дорівнюють одиниці. Тоді $A = I$, $A = SAC = SIC' = SC' = I$. Якщо не всі корені рівні λ , то один або більше дорівнюють 0. Отже, ми маємо такі результати для рівнопотужних симетричних матриць:

лише за повного рангу матриці
симетрична рівнопотужна матриця
є одиничною матрицею. (7.76)

Всі симетричні рівнопотужні матриці
окрім одиничної, є невивродженими. (7.77)

Кінцевий результат рівнопотужної матриці отримується з того факту,
що кількість ненульових коренів A дорівнює їх сумі. Тобто:

ранг симетричної рівнопотужної матриці дорівнює її сліду. (7.78)

7.5.12. Розкладання матриці на множники

У деяких прикладних задачах може знадобитись матриця P , квадрат
якої дорівнює оберненій матриці

$$P'P = A^{-1}.$$

Один з варіантів такий:

$$P = A^{-1/2}C',$$

отже,

$$P'P = (C')(A^{-1/2})A^{-1/2}C' = C^{-1}C',$$

що і було потрібно.

Корисною альтернативою цієї залежності у регресійному аналізі є **розкладання за множниками Чолескі** симетричної додатно визначеної матриці. Будь-яка симетрична додатно визначена матриця A може бути записана як добуток нижньо-трикутної матриці L і транспонованої матриці $L' = U$ (яка є верхньо-трикутною матрицею). Таким чином, $A = LU$. Це є декомпозицією Чолескі матриці A . Квадратні корені квадратів діагональних елементів L , d_{ij} є значеннями Чолескі для A . Звівши їх у діагональну матрицю D , можемо також записати $A = LD^{-1}D^2D^{-1}U = L^*DU^*$. Аналогічно спектральній декомпозиції у (7.69). Корисність цих висновків стає очевидною, коли треба побудувати матрицю, обернену до A . Оскільки L уже підрахована, знаходження $A^{-1} = U^{-1}L^{-1}$ є надзвичайно швидким і точним. Найновіше програмне забезпечення економетрики використовує ці методи для обернення додатно визначених матриць.

7.5.13. Узагальнені обернені матриці

Обернена матриця є фундаментальним поняттям в економетриці. **Узагальненою оберненою матрицею до A** буде матриця A^+ , яка задовольняє такі умови.

1. $AA^+A = A$.
2. $A^+AA^+ = A^+$.
3. A^+A — симетрична.
4. AA^+ — симетрична.

Єдину матрицю A^+ можна знайти для будь-якої матриці незалежно від того, чи вона вивроджена чи ні, і навіть якщо A не квадратна. Ця єдина матриця, що задовольняє усі чотири умови, називається **оберненою матрицею Мура — Пенроуза**, або **псевдооберненою матрицею**. Якщо матриця A є квадратною і невивродженою, то узагальнена обернена буде така сама, як і звичайна обернена матриця. Але якщо A^{-1} не існує, матрицю A^+ можна розрахувати.

Розглянемо ще один важливий випадок перевизначеної системи рівнянь

$$Ab = y,$$

де A має n рядків, $k < n$ стовпців і ранг, рівний $R \leq k$. Припустимо, що $R = k$, тобто матриця $(A'A)^{-1}$ існує. Обернена матриця Мура — Пенроуза буде:

$$A^+ = (A'A)^{-1}A'.$$

Розв'язок системи рівнянь може бути записаний таким чином:

$$b = A^+y.$$

Зауважимо, що цей розв'язок ідентичний розв'язкові, отриманому за методом найменших квадратів.

Тепер припустимо, що матриця A не має повного рангу. Попередній розв'язок неможливий, але можна знайти альтернативний розв'язок. Ми продовжуємо використовувати матрицю $A'A$. У спектральній декомпозиції (розділ 7.5.9), якщо A має ранг R , то у додаванні (7.69) є R операцій. У (7.72) спектральна декомпозиція, що використовує обернені величини характеристичних коренів, використовується для обчислення оберненої матриці. Щоб обчислити обернену матрицю Мура — Пенроуза, застосуємо ці розрахунки до $A'A$, використовуючи лише ненульові корені, потім результат домножимо справа на A' . Нехай C_1 являє собою R характеристичних векторів, що відповідають ненульовим кореням, які ми записуємо як множину діагональних елементів матриці A_1 . Тоді обернена матриця Мура — Пенроуза буде

$$A^+ = C_1A_1^{-1}C_1'A',$$

що дуже схоже на попередній розв'язок.

Якщо A — квадратна матриця з рангом $R \leq K$, то обернену матрицю Мура — Пенроуза можна обчислити аналогічно попередній формулі без множення справа на A' . Тобто для симетричної матриці A

$$A^+ = C_1 A_1^{-1} C_1',$$

де A_1 — діагональна матриця, що включає обернені величини ненульових коренів A .

7.6. КВАДРАТИЧНІ ФОРМИ ТА ВИЗНАЧЕНІ МАТРИЦІ

Багато задач оптимізації включають подвійні суми виду:

$$q = \sum_{i=1}^n \sum_{j=1}^n x_i x_j a_{ij}. \quad (7.79)$$

Квадратична форма може бути записана:

$$q = x'Ax,$$

де A — симетрична матриця. Взагалі q може бути додатним, від'ємним або рівним 0; воно залежить від A та x . Але є деякі матриці, для яких q буде додатним незалежно від x , і такі, для яких q буде завжди від'ємним. Для даної матриці A справджується теорема.

1. Якщо $x'Ax > (<) 0$ для всіх x , не рівних нулю, тоді A — *додатно (від'ємно) визначена*.

2. Якщо $x'Ax \geq (\leq) 0$ для всіх x , не рівних нулю, тоді A — *невід'ємно (недодатно) визначена* або *додатно (від'ємно) напіввизначена*.

Може здаватися, що неможливо перевірити матрицю на визначеність, оскільки x вибирається довільно. Але ми вже маємо деякі результати для того, щоб зробити це. Повернемось до того, що симетрична матриця може бути декомпонована у

$$A = C\Lambda C'.$$

Тоді квадратична форма може записуватись як

$$x'Ax = x'C\Lambda C'x.$$

Нехай $y = C'x$. Тоді

$$x'Ax = y'\Lambda y = \sum_{i=1}^n \lambda_i y_i^2. \quad (7.80)$$

Якщо λ_i додатна для всіх i , тоді незалежно від y , тобто незалежно від x , q буде додатним. Це буде випадок додатно визначеної матриці. За аналогією, якщо всі характеристичні корені A від'ємні, то матриця є *від'ємно визначеною*. Якщо деякі корені рівні нулю, то матриця є *невід'ємно визначеною*, якщо решта коренів додатні, і *недодатно визначеною*, якщо вони від'ємні. Якщо A має і додатні, і від'ємні корені, то така матриця називається *невизначеною*.

7.6.1. Невід'ємно визначені матриці

Невід'ємно визначені матриці є досить цікавим випадком. З попередньої теореми випливає декілька важливих наслідків.

Наслідок 1. Якщо A невід'ємно визначена, тоді $|A| \geq 0$. (7.81)

Доведення випливає з того, що визначник є добутком коренів.

Але обернене твердження неправильне. Наприклад матриця розмірності $[2 \times 2]$ з двома від'ємними коренями очевидно не є додатно визначеною, але її визначник додатний.

Наслідок 2. Якщо A додатно визначена, то так само визначена і A^{-1} (7.82)

Доведення. Корені A^{-1} є оберненими величинами до коренів A , які, таким чином, є додатними.

Наслідок 3. Одиначна матриця I додатно визначена. (7.83)

Доведення: $x'Ix = x'x > 0$, якщо $x \neq 0$.

Дуже важливим наслідком для регресивної моделі є

Наслідок 4. Якщо A має розмірність $[n \times k]$ з повним рангом і $n > k$, то $A'A$ додатно визначена і AA' невід'ємно визначена.

Доведення. За припущенням, $Ax \neq 0$. Отже, $x'A'Ax = (Ax)'(Ax) = y'y = \sum_i y_i^2 > 0$.

Аналогічне доведення для невід'ємної визначеності AA' .

Наслідок 5. Якщо A додатно визначена, а B — невироджена матриця, то $B'AB$ додатно визначена. (7.84)

Доведення: $x'B'ABx = y'Ay > 0$, де $y = Bx$. Але y не може дорівнювати 0, бо B невироджена.

Нарешті зауважимо, що для того, щоб A була від'ємно визначена, всі її характеристичні корені мають бути від'ємними. Але у цьому випадку $|A|$ є додатним, якщо матриця A парного порядку, і від'ємним, якщо A непарного порядку.

7.6.2. Рівнопотужні квадратичні форми

Рівнопотужні квадратичні форми відіграють велику роль у розподілі багатьох статистичних результатів. Найважливішими є два наступних твердження.

Твердження 1. Будь-яка симетрична рівнопотужна матриця n від'ємно визначена. (7.85)

Доведення. Всі корені дорівнюють або 0, або 1; звідси випливає, що матриця невід'ємно визначена за означенням.

Твердження 2. Якщо A симетрична і рівнопотужна матриця розмірності $[n \times n]$ і має ранг J , то будь-яка квадратична форма може бути записана так:

$$x'Ax = \sum_{i=1}^J y_i^2. \quad (7.86)$$

Доведення. Здійснюється підстановкою в (7.80) $\lambda=1$ або 0.

7.6.3. Ранжування матриць

Як можна порівняти дві матриці — не зовсім очевидно. Спочатку, звичайно, необхідно, щоб вони були однакової розмірності. Корисне, хоч і не завжди вичерпне, порівняння засноване на:

$$d = x'Ax - x'Bx.$$

Якщо ця величина завжди додатна, незалежно від вибору x , тоді, принаймні за цим критерієм, можемо сказати, що

$$A > B. \quad (7.87)$$

Протилежну нерівність отримаємо, якщо d завжди від'ємне. Для деяких матриць ця величина може бути невизначеною. З означення випливає, що

якщо $d > 0$, то $A - B$ додатно визначена,

бо

$$d = x'(A - B)x.$$

Наприклад, визначити, чи $A > B$ за умов:

$$A = \begin{bmatrix} 3 & -1 \\ -1 & 49 \end{bmatrix}; B = \begin{bmatrix} 2 & 5 \\ 5 & 13 \end{bmatrix}; A - B = \begin{bmatrix} 1 & -6 \\ -6 & 36 \end{bmatrix}.$$

Характеристичні корені цієї матриці рангу 1 дорівнюють 37 та 0. Тому $A - B$ невід'ємно визначена. Для цих A та B квадратична форма з A буде принаймні така ж велика, як і квадратична форма з B , при однаковому x .

Особливий випадок загального результату такий: якщо A додатно визначена, а B від'ємно визначена, то

$$A + B > A. \quad (7.88)$$

7.7. ОБЧИСЛЕННЯ І МАТРИЧНА АЛГЕБРА

Багато задач економетрики включають розширення деяких відомих результатів з арифметики. Як приклади можна вказати формулювання наближених функцій, максимізацію або мінімізацію значення функції багатьох змінних, аналіз переходу від однієї множини величин до іншої.

7.7.1. Диференціали та ряди Тейлора

Змінна y є функцією від іншої змінної x і позначається:

$$y = f(x), y = g(x), y = y(x) \text{ і т. д.},$$

якщо кожному значенню змінної x відповідає єдине значення змінної y . У такому разі змінні x та y часто називають відповідно незалежною змінною та залежною змінною.

Припускаючи, що функція є неперервною та диференційованою, отримуємо:

$$f'(x) = \frac{dy}{dx},$$

$$f''(x) = \frac{d^2y}{dx^2}.$$

і т. д.

Похідні $f(x)$ часто використовуються у *наближенні за рядами Тейлора*. Ряди Тейлора є поліноміальним наближенням $f(x)$. Якщо $x^{(0)}$ є довільно вибраною точкою розвинення,

$$f'(x) \cong f(x^{(0)}) + \sum_{i=1}^p \frac{1}{i!} \left. \frac{d^i f(x)}{dx^i} \right|_{x=x^{(0)}} (x - x^{(0)})^i. \quad (7.89)$$

Вибір кількості членів є довільним: чим більше їх використовується, тим точнішим буде наближення. В економетриці найчастіше використовуються *лінійні наближення*

$$f(x) \cong [f(x^{(0)}) - f'(x^{(0)})x^{(0)}] + f'(x^{(0)})x = \beta_1 + \beta_2 x \quad (7.90)$$

і *квадратичні наближення*

$$f(x) \cong [f^{(0)} - f'^{(0)}x^{(0)} + 1/2 f''^{(0)}(x^{(0)})^2] + [f'^{(0)} - f''^{(0)}x^{(0)}]x + 1/2 f''^{(0)}x^2 = b_1 + b_2 x + b_3 x^2, \quad (7.91)$$

де верхні індекси вказують, що функція оцінюється у деякому околі точки $x^{(0)}$.

Ми можемо задати функцію $y = f(x_1, x_2, \dots, x_n)$ як функцію зі *скалярними значеннями*, тобто $y = f(x)$. Вектором часткових похідних, або *вектором-градієнтом*, або просто *градієнтом* є:

$$\frac{\partial f(x)}{\partial x} = \begin{bmatrix} \partial y / \partial x_1 \\ \partial y / \partial x_2 \\ \dots \\ \partial y / \partial x_n \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \dots \\ f_n \end{bmatrix}. \quad (7.92)$$

Вектор $g(x)$ або g використовується для позначення вектора-градієнта. Зауважимо, що це вектор-стовпець.

Матриця других похідних, або *матриця Гессе* розраховується так:

$$H = \begin{bmatrix} \partial^2 y / \partial x_1 \partial x_1 & \partial^2 y / \partial x_1 \partial x_2 & \dots & \partial^2 y / \partial x_1 \partial x_n \\ \partial^2 y / \partial x_2 \partial x_1 & \partial^2 y / \partial x_2 \partial x_2 & \dots & \partial^2 y / \partial x_2 \partial x_n \\ \dots & \dots & \dots & \dots \\ \partial^2 y / \partial x_n \partial x_1 & \partial^2 y / \partial x_n \partial x_2 & \dots & \partial^2 y / \partial x_n \partial x_n \end{bmatrix} = \left| f_{ij} \right|. \quad (7.93)$$

У загальному випадку матриця H — квадратна і симетрична. Кожний рядок та стовпець H є похідною вектора-градієнта за однією із змінних. Таким чином:

$$H = \begin{bmatrix} \frac{\partial[\partial y / \partial x]}{\partial x_1} & \frac{\partial[\partial y / \partial x]}{\partial x_2} & \dots & \frac{\partial[\partial y / \partial x]}{\partial x_n} \end{bmatrix} = \frac{\partial[\partial y / \partial x]}{\partial(x_1 \ x_2 \ \dots \ x_n)} = \frac{\partial[\partial y / \partial x]}{\partial x'}$$

$$= \frac{\partial^2 y}{\partial x \partial x'}.$$

Наближення першого порядку або лінійні ряди Тейлора такі:

$$y \cong f(x^0) + \sum_i f_i(x^0)(x_i - x_i^0)$$

$$= f(x^0) + \left[\partial f(x^0) / \partial x^0 \right] (x - x^0)$$

$$= f(x^0) + g(x^0)' (x - x^0)$$

$$= \left[f(x^0) - g(x^0)' x^0 \right] + g(x^0)' x \quad (7.94)$$

$$= \left[f^0 - g^0' x^0 \right] + g^0' x$$

$$= \beta_1 + \beta_2' x.$$

Наближення другого порядку, або квадратичне, додає у наближення члени другого порядку

$$1/2 \sum_i \sum_j f_{ij}^0 (x_i - x_i^0) (x_j - x_j^0) = 1/2 (x - x^0)' H^0 (x - x^0).$$

Згрупувавши члени таким же чином, як у (7.94), отримуємо:

$$y \cong \beta_1 + \beta_2' x + 1/2 x' \Gamma_3 x,$$

де

$$\beta_1 = f^0 - g^0' x^0 + 1/2 x^0' H^0 x^0,$$

$$\beta_2 = -g^0 + H^0 x^0$$

і

$$\Gamma_3 = H^0.$$

Лінійна функція може бути записана:

$$y = a'x = x'a = \sum_i a_i x_i.$$

Надалі

$$\frac{\partial(a'x)}{\partial x} = a. \quad (7.95)$$

Зауважимо, зокрема, що $\partial(a'x)/\partial x = a$, а не a' . У системі лінійних функцій

$$y = Ax$$

кожний елемент y_i з $y \in$

$$y_i = a^i x,$$

де $a^i \in i$ -им рядком A . Таким чином,

$$\frac{\partial y_i}{\partial x} = a^i = \text{транспонованому } i\text{-ому рядку } A$$

і

$$\begin{bmatrix} \partial y_1 / \partial x' \\ \partial y_2 / \partial x' \\ \vdots \\ \partial y_n / \partial x' \end{bmatrix} = \begin{bmatrix} a^1 \\ a^2 \\ \vdots \\ a^n \end{bmatrix}$$

Згрупувавши всі члени, знайдемо, що $\partial Ax / \partial x' = A$, тоді як більш знайомою формою буде:

$$\frac{\partial Ax}{\partial x} = A'. \quad (7.96)$$

Квадратична форма записується:

$$x'Ax = \sum_i \sum_j x_i x_j a_{ij}. \quad (7.97)$$

Наприклад,

$$A = \begin{bmatrix} 1 & 3 \\ 3 & 4 \end{bmatrix},$$

тоді

$$x'Ax = 1x_1^2 + 4x_2^2 + 6x_1x_2.$$

Тоді

$$\frac{\partial x'Ax}{\partial x} = \begin{bmatrix} 2x_1 + 6x_2 \\ 6x_1 + 8x_2 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 6 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2Ax, \quad (7.98)$$

що є загальним результатом, якщо A — симетрична матриця. Якщо A не є симетричною, то

$$\frac{\partial (x'Ax)}{\partial x} = (A + A')x. \quad (7.99)$$

Посилаючись на попереднє подвійне додавання, знаходимо, що для кожного члена коефіцієнтом при a_{ij} є $x_i x_j$. Таким чином,

$$\frac{\partial (x'Ax)}{\partial a_{ij}} = x_i x_j.$$

Квадратною матрицею, ij -им елементом якої є $x_i x_j$, є xx' , отже,

$$\frac{\partial (x'Ax)}{\partial A} = xx'. \quad (7.100)$$

Похідні, що містять визначники, з'являються у максимально наближених оцінках. З використанням алгебраїчних доповнень

$$\frac{\partial |A|}{\partial a_{ij}} = (-1)^{i+j} |C_{ji}|,$$

де C_{ij} є ij -им алгебраїчним доповненням A . Матриця, обернена до A , може бути обчислена з використанням

$$A_{ij}^{-1} = \frac{(-1)^{i+j} |C_{ij}|}{|A|}.$$

Зверніть увагу на зворотний порядок індексів, з чого виходить, що

$$\frac{\partial \ln |A|}{\partial a_{ij}} = \frac{(-1)^{i+j} |C_{ij}|}{|A|}$$

або, за умови групування членів:

$$\frac{\partial \ln |A|}{\partial A} = A^{-1}.$$

7.7.2. Оптимізація

Багато прикладних проблем вимагають знаходження такого значення x , при якому $f(x)$ є мінімальним або максимальним. Враховуючи, що $f'(x)$ є нахилом $f(x)$, маємо оптимум за умови $f'(x) = 0$. У інших випадках функція від x буде зростати або спадати. З цього випливає умова першого порядку, або *необхідна умова для знаходження екстремуму*:

$$\frac{\partial y}{\partial x} = 0. \quad (7.101)$$

Ця умова є необхідною як для максимуму, так і для мінімуму. Для максимуму функція має бути увігнутою; для мінімуму вона має бути опуклою.

Достатня умова оптимуму

$$\text{для максимуму: } \frac{\partial^2 y}{\partial x^2} < 0; \quad (7.102)$$

$$\text{для мінімуму: } \frac{\partial^2 y}{\partial x^2} > 0.$$

Деякі функції, такі як синус і косинус, мають багато *локальних оптимумів*, тобто багато мінімумів і максимумів. Така функція, як $\sin x / |x|$, котра є затухаючою синусоїдою, також має багато локальних оптимумів, але відрізняється тим, що має один максимум при $x = 0$, при якому $f(x)$ є більшою, ніж у будь-якій іншій точці. Таким чином, $x = 0$ — *глобальний*

максимум, тоді як інші максимуми — тільки **локальні максимуми**. Деякі функції, такі як квадратична, мають тільки один оптимум. Ці функції **глобально опуклі**, якщо оптимум є мінімумом, і **глобально увігнуті**, якщо він є максимумом.

Для максимізації чи мінімізації функції декількох змінних умовами першого порядку є:

$$\frac{\partial f(\bar{x})}{\partial x} = 0. \quad (7.103)$$

Це інтерпретується таким же чином, як і необхідна умова у разі однієї змінної. При оптимумі має виконуватись умова, що ніяка маленька зміна у будь-якій змінній не спричинить поліпшення значення функції. У випадку однієї змінної d^2y/dx^2 має бути додатним для мінімуму і від'ємним для максимуму. У випадку багатьох змінних можна записати аналогічну умову для матриці других похідних цільової функції. Зокрема, умовами другого порядку для оптимуму є такі, що при оптимальному значення матриці других похідних (Гессе)

$$H = \frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x}) \quad (7.104)$$

мають бути додатно визначеними для мінімуму і від'ємно визначеними для максимуму.

У задачі з однією змінною умова другого порядку звичайно може бути перевірена. У загальному випадку для багатьох змінних це не так. Як ми встановили раніше, перевірка визначеності матриці є, загалом, відносно складним завданням. Для багатьох проблем, з якими має справи економетрика, умова другого порядку може бути застосована залежно від структури задачі. Тобто матриця H буде звичайно мати таку форму, що завжди буде визначеною.

Для ілюстрації вищенаведеного розглянемо задачу:

$$\text{Maximize } R = a'x - x'Ax,$$

де

$$a' = (5 \ 4 \ 2)$$

і

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 3 & 2 \\ 3 & 2 & 5 \end{bmatrix}.$$

Використовуючи деякі вже відомі результати, отримуємо

$$\begin{aligned} \frac{\partial R}{\partial x} &= a - 2Ax \\ &= \begin{bmatrix} 5 \\ 4 \\ 2 \end{bmatrix} - \begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 4 \\ 6 & 4 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0. \end{aligned} \quad (7.105)$$

Розв'язком буде:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 4 \\ 6 & 4 & 10 \end{bmatrix}^{-1} \begin{bmatrix} 5 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 11.25 \\ 1.75 \\ -7.25 \end{bmatrix}.$$

Достатня умова полягає в тому, щоб

$$\frac{\partial^2 R(x)}{\partial x \partial x'} = -2A = \begin{bmatrix} -4 & -2 & -6 \\ -2 & -6 & -4 \\ -6 & -4 & -10 \end{bmatrix} \quad (7.106)$$

були від'ємно визначеними. Трьома характеристичними коренями цієї матриці є -15.776 , -4 і -0.25403 . Оскільки всі три корені від'ємні, матриця є від'ємно визначеною, як і вимагалось.

У вищевикладеному необхідно було обчислювати характеристичні корені гессіана для перевірки достатньої умови. Для загальної матриці, що має порядок, більший 2, для цього, звичайно, потрібен комп'ютер. Однак, припустимо, що A має форму

$$A = B'B,$$

де B є певною відомою матрицею. Тоді, як було показано раніше, A завжди буде додатно визначеною (за умови, що B має повний ранг). Немає потреби обчислювати характеристичні корені A для перевірки достатніх умов.

7.7.3. Умовна оптимізація

Часто буває необхідним розв'язати оптимізаційну задачу з певними обмеженнями розв'язків. Один метод полягає просто у "виведенні обмежень із задачі". Наприклад, у задачі максимізації, що розглядалася вище, припустимо, що на розв'язок накладається обмеження $x_1 = x_2 - x_3$. Для єдиного обмеження, такого як це, можна просто підставити праву частину

рівняння для x_1 у цільову функцію і розв'язати задачу як функцію двох змінних, що залишилися. Однак, для більш загальної форми обмежень, або якщо можливе більше, ніж одне обмеження, потужнішим методом розв'язку задачі є *метод множників Лагранжа*.

Максимізуємо по x $f(x)$ при обмеженнях

$$\begin{aligned} c_1(x) &= 0; \\ c_2(x) &= 0; \\ &\dots \\ c_j(x) &= 0. \end{aligned} \quad (7.107)$$

Формулювання лагранжиана полягає у

$$\text{Maximize } L^*(x, \lambda) = f(x) + \sum_{j=1}^J \lambda_j c_j(x), \quad (7.108)$$

де $c_1(), \dots, c_j()$ є J обмеженнями, а λ_j — множниками Лагранжа. Визначимо вектор множників Лагранжа λ і уявімо обмеження як вектор $c(x)$. Тоді формула лагранжиана запишеться так:

$$L^*(x, \lambda) = f(x) + \lambda c(x). \quad (7.109)$$

Умовами першого порядку є:

$$\begin{aligned} \frac{\partial L^*}{\partial x} &= \frac{\partial f(x)}{\partial x} + \frac{\partial \lambda' c(x)}{\partial x} = 0; \\ \frac{\partial L^*}{\partial \lambda} &= c(x) = 0. \end{aligned} \quad (7.110)$$

Другим членом у $\partial L^*/\partial x$ є

$$\begin{aligned} \frac{\partial \lambda' c(x)}{\partial x} &= \frac{\partial c(x)' \lambda}{\partial x} \\ &= \left[\frac{\partial c(x)'}{\partial x} \right] \lambda = C' \lambda, \end{aligned} \quad (7.111)$$

де C — похідні від обмежень за x і j -ий рядок матриці C , що має розмірність $[J \times n]$ (є вектором похідних j -го обмеження $c_j(x)$ за x'). Для всіх членів умовами першого порядку буде:

$$\begin{aligned} \frac{\partial L^*}{\partial x} &= \frac{\partial f(x)}{\partial x} + C' \lambda = 0; \\ \frac{\partial L^*}{\partial \lambda} &= c(x) = 0. \end{aligned} \quad (7.112)$$

Є дуже важливий аспект розв'язку системи обмежень, який варто розглянути. У безумовному розв'язку ми маємо $f(x)/(x) = 0$. З (7.112) отримуємо для розв'язку задачі з обмеженнями

$$\frac{\partial f(x)}{\partial x} = -C' \lambda,$$

що не дорівнює 0, якщо не виконується $\lambda=0$. Це має два важливих наслідки.

1. Умовний розв'язок має бути гіршим, ніж безумовний розв'язок. Це впливає із ненульового градієнта при умовному розв'язку.

2. Якщо множники Лагранжа нульові, тоді умовний розв'язок буде дорівнювати безумовному розв'язку.

Для продовження прикладу, розпочатого раніше, припустимо, що ми додаємо такі умови:

$$\begin{aligned} x_1 - x_2 + x_3 &= 0, \\ x_1 + x_2 + x_3 &= 0. \end{aligned}$$

Щоб привести це до способу уявлення загальної задачі, запишемо обмеження у вигляді

$$c(x) = Cx = 0,$$

де

$$C = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Запишемо лагранжиан:

$$R^*(x, \lambda) = a'x - x'Ax + \lambda'Cx.$$

Зверніть увагу на розмірність і групування різних частин. Зокрема, C є матрицею розмірності $[2 \times 3]$ з одним рядком для кожного обмеження і одним стовпцем для кожної змінної у цільовій функції. Таким чином, вектор множників Лагранжа має два елементи, один для кожного обмеження. Умовами першого порядку є:

$$a - 2Ax + C'\lambda = 0 \quad (\text{три рівняння})$$

і

$$Cx = 0 \quad (\text{два рівняння}). \quad (7.113)$$

Це можна об'єднати в одному рівнянні

$$\begin{bmatrix} -2A & C' \\ C & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} -a \\ 0 \end{bmatrix}.$$

Використовуючи комбіновану інверсію, отримуємо такі результати:

$$\lambda = -|CA^{-1}C'|CA^{-1}a \quad (7.114)$$

і

$$x = 1/2A^{-1}|I - C'(CA^{-1}C')^{-1}CA^{-1}|a. \quad (7.115)$$

Обидва результати (7.114) і (7.115) дають аналітичні розв'язки для λ і x . Для окремих матриць і векторів прикладу $\lambda = [-0.5 \ -7.5]'$ і вектор розв'язку $x^* = [1.5 \ 0 \ -1.5]'$. Зауважимо, що при розв'язанні задач такого типу немає потреби використовувати досить громіздку формулу (7.115). Коли вже λ отримано з (7.114), розв'язок може бути підставлено у (7.113) для набагато простіших обчислень. Розв'язок

$$x = 1/2A^{-1}a + 1/2A^{-1}C'\lambda$$

дає корисний результат для умовного оптимуму:

$$\text{умовний розв'язок} = \text{безумовному розв'язку} + 1/[2A]^{-1}C'\lambda. \quad (7.116)$$

Нарешті, підставляючи два розв'язки у первинну функцію, знаходимо, що $R = 24.375$ і $R^* = 2.25$, а це знову ілюструє, що умовний розв'язок (у цій задачі максимізації) є гіршим за безумовний.

7.7.4. Перетворення

Якщо функція є строго монотонною, то вона є взаємно однозначною. Кожному y відповідає точно одне значення x , і навпаки. В такому разі є обернена функція, що визначає x як функцію y , яка записується так:

$$y = f(x) \text{ і } x = f^{-1}(y). \quad (7.117)$$

Наприклад, обернена залежність між логарифмічною та степеневою функціями. Нахил оберненої функції:

$$\frac{dx}{dy} = \frac{df^{-1}(y)}{dy} = f^{-1\prime}(y) \quad (7.118)$$

і називається матрицею Якобі перетворення з x в y . Наприклад, якщо

$$y = ax + b, \quad (7.119)$$

тоді

$$x = -\frac{a}{b} + \left(\frac{1}{b}\right)y \quad (7.120)$$

буде оберненим перетворенням і

$$J = \frac{dx}{dy} = \frac{1}{b}. \quad (7.121)$$

Передбачаючи потребу статистичного застосування цього матеріалу, зазначимо, що якщо $y=f(x)$ є вертикальною лінією, то вона не буде функціональною залежністю. Одному x буде відповідати більше ніж одне значення y . У цьому разі при даному значенні x $J=0$, що вказує на виродженість функції.

Якщо y — вектор-стовпець, $y=f(x)$, то:

$$J = \frac{\partial x}{\partial y'} = \begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \dots & \frac{\partial x_1}{\partial y_n} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \dots & \frac{\partial x_2}{\partial y_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \dots & \frac{\partial x_n}{\partial y_n} \end{bmatrix}. \quad (7.122)$$

Розглянемо множину лінійних функцій $y = Ax = f(x)$. Обернене перетворення буде $x = f^{-1}(y)$, і матиме вигляд:

$$x = A^{-1}y,$$

якщо A не вироджена,
або

$$x = 0,$$

якщо A вироджена.

Нехай J — матриця часткових похідних оберненої функції.

$$J = \left[\frac{\partial x_i}{\partial y_j} \right] \quad (7.123)$$

Абсолютна величина визначника J

$$\text{abs}(J) = \left| \frac{\partial x}{\partial y'} \right| \quad (7.124)$$

є матрицею Якобі трансформації з x в y . У разі не виродженої матриці

$$\text{abs}(J) = \text{abs}(A^{-1}) = \frac{1}{\text{abs}(A)}. \quad (7.125)$$

У разі виродженої матриці всі часткові похідні будуть дорівнювати 0,

як і матриця Якобі. У даній ситуації з того, що матриця Якобі є нульовою, випливає, що A — вироджена, і трансформація з x в y функціонально залежна. Випадок виродження аналогічний випадку однієї змінної.

Очевидно, що якщо вектор x заданий, то $y = Ax$ можна обчислити. Чи може x бути вираженим через y — це вже інше питання. Взагалі, це залежить від матриці Якобі. Якщо матриця Якобі ненульова, то обернена залежність можлива, і ми можемо виразити x через y . Якщо ні, то ми не можемо отримати x .

Вправи

Вправа 1. Для матриць

$$A = \begin{bmatrix} 1 & 3 & 3 \\ 2 & 4 & 1 \end{bmatrix},$$

$$B = \begin{bmatrix} 2 & 4 \\ 1 & 5 \\ 6 & 2 \end{bmatrix}$$

обчислити AB , $A'B'$, і BA .

Вправа 2. Довести, що $\text{tr}(AB) = \text{tr}(BA)$, де A і B — дві довільні матриці, зручні для множення з обох сторін, причому вони не обов'язково квадратні.

Вправа 3. Розвинути добуток матриць:

$$X = ([AB + (CD)'][(EF)^{-1} + GH])'.$$

Припустіть, що всі матриці квадратні, а E і F невироджені.

Вправа 4. Доведіть, що для n [$k \times 1$] векторів-стовпців x_i , $i=1, \dots, n$ і деякого ненульового вектора a ,

$$\sum_i (x_i - a)(x_i - a)' = X' M^0 X + n(x - a)(x - a)',$$

де i -ий рядок X позначено x_i' , а M^0 визначена в (7.28).

Вправа 5. Нехай A довільна квадратна матриця, стовпцями якої є $[a_1, a_2, \dots, a_n]$, і нехай B — будь-яка перестановка стовпців одиничної матриці розмірності $[M \times M]$. Яка операція виконується при множенні AB ? При множенні BA ?

Вправа 6. Розглянемо матрицю B із вправи 5 розмірністю $[3 \times 3]$. Наприклад:

$$B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

Обчисліть B^2 і B^3 . Повторіть обчислення для $[4 \times 4]$. Чи можете ви узагальнити результати?

Вправа 7. Обчисліть визначник матриці A , $\text{tr}(A)$ і A^{-1} для

$$A = \begin{bmatrix} 1 & 4 & 7 \\ 3 & 2 & 5 \\ 5 & 2 & 8 \end{bmatrix}.$$

Вправа 8. Отримайте декомпозицію Чолескі для матриці

$$A = \begin{bmatrix} 25 & 7 \\ 7 & 13 \end{bmatrix}.$$

Скористайтесь тим, що декомпозицією Чолескі матриці A є добуток $LU = A$, де L — нижньотрикутна матриця і $U = L'$.

Вправа 9. Симетрична додатно визначена матриця A може бути записана як $A = UL$, де U — верхньотрикутна матриця і $L = U'$. Але це не є декомпозицією Чолескі. Зробіть цю декомпозицію для матриці з вправи 8.

Вправа 10. Які дії виконуються при множенні зліва матриці на діагональну матрицю?

Що можна сказати з приводу множення справа?

Вправа 11. Чи є такі квадратичні форми додатними при всіх значеннях x ?

a) $y = x_1^2 - 28x_1x_2 + (11x_2)^2$;

b) $y = 5x_1^2 + x_2^2 + 7x_3^2 + 4x_1x_2 + 6x_1x_3 + 8x_2x_3$.

Вправа 12. Доведіть, що $\text{tr}(A \otimes B) = \text{tr}(A)\text{tr}(B)$.

Вправа 13. Обчисліть характеристичні корені матриці

$$A = \begin{bmatrix} 2 & 4 & 3 \\ 4 & 8 & 6 \\ 3 & 6 & 5 \end{bmatrix}.$$

Вправа 14. Припустимо, що $A = A(z)$, де z — скалярна величина. Що являє собою $\partial (x'Ax) / \partial z$? Припустимо, що кожний x є функцією від z . Що являє собою $\partial (x'Ax) / \partial z$?

Вправа 15. Покажіть, що розв'язки визначникових рівнянь $|B - \lambda A| = 0$ і $|A^{-1}B - \lambda I| = 0$ однакові. Як розв'язки цих рівнянь співвідносяться з розв'язками рівняння $B^{-1}A - \mu I = 0$?

Вправа 16. Використовуючи матрицю A з вправи 8, знайдіть такий вектор x , що мінімізує $y = x'Ax + 2x_1 + 3x_2 - 10$. Яким є значення y в цьому мінімумі? Мінімізуйте y , що задовольняє обмеження $x_1 + x_2 = 1$. Порівняйте обидва розв'язки.

Вправа 17. Яким буде якобіан таких перетворень:

$$\begin{aligned} y_1 &= \frac{x_1}{x_2}, \\ \ln y_1 &= \ln x_1 - \ln x_2 + \ln x_3, \\ y_3 &= x_1 x_2 x_3? \end{aligned}$$

Вправа 18. Доведіть, що перестановка двох стовпців квадратної матриці змінює на протилежний знак її визначник.

Вправа 19. Для матриці

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 4 & -2 & 3 & -5 \end{bmatrix}$$

обчисліть $P = X(X'X)^{-1}$ і $M = (I - P)$.

Вправа 20. Знайдіть обернену матрицю до

$$P = \begin{bmatrix} \cos(x) & \sin(x) \\ -\sin(x) & \cos(x) \end{bmatrix}.$$

Знайдіть характеристичні корені P .

7.8. ОСНОВНІ ПОНЯТТЯ ТЕОРІЇ ЙМОВІРНОСТЕЙ

7.8.1. Інтегральна функція розподілу ймовірностей випадкової величини

Інтегральною функцією розподілу називають функцію $F(x)$, яка визначає для кожного значення x ймовірність того, що випадкова величина X набере значення, менше ніж x , тобто

$$F(x) = P(X < x).$$

Геометрично цю рівність можна уявити як x -ймовірність того, що випадкова величина X набере значення, менше за те, яке відображається на числовій осі точкою, котра знаходиться ліворуч від точки x .

Випадкова величина буде називатися неперервною, якщо її функція розподілу неперервно диференційована.

Розглянемо властивості інтегральної функції розподілу.

1. Значення функції розподілу належать відрізку $[0,1]$:

$$0 \leq F(x) \leq 1.$$

2. $F(x)$ — неспадна функція, тобто

$$F(x_2) > F(x_1), \text{ якщо } x_2 > x_1.$$

Наслідок 1. Імовірність того, що випадкова величина набере значення, яке належить інтервалу (a,b) , дорівнює приросту інтегральної функції на цьому інтервалі:

$$P(a \leq X \leq b) = F(b) - F(a).$$

Наслідок 2. Імовірність того, що випадкова величина X набере одного визначеного значення, дорівнює нулеві. $P(x = x_1) = 0$.

$$P(a \leq X < b) = P(a < X < b) = P(a < X \leq b) = P(a \leq X \leq b).$$

Наслідок 3. Якщо можливі значення випадкової величини належать інтервалу (a,b) , то

$$\begin{aligned} F(x) &= 0 \text{ при } x \leq a; \\ F(x) &= 1 \text{ при } x \geq b. \end{aligned}$$

Наслідок 4. Якщо можливі значення неперервної випадкової величини належать усій числовій осі x , то правильне таке граничне відношення:

$$\lim_{x \rightarrow -\infty} F(x) = 0; \quad \lim_{x \rightarrow \infty} F(x) = 1.$$

7.8.2. Диференційна функція (функція густини) розподілу ймовірностей випадкової величини

Перша похідна від функції розподілу називається диференційною функцією: $f(x) = F'(x)$.

7.8.2.1. Імовірність влучення неперервної випадкової величини у визначений інтервал

Теорема. Імовірність того, що неперервна випадкова величина X набере значення, яке належить інтервалу (a,b) , дорівнює визначеному інтегралу від диференційної функції у межах від a до b .

$$P(a < X < b) = \int_a^b f(x) dx.$$

Якщо $f(x)$ — функція парна і кінці інтервалу симетричні відносно початку координат, то

$$P(-a < X < a) = P(|x| < a) = 2 \int_0^a f(x) dx.$$

7.8.2.2. Властивості диференційної функції

1. Диференційна функція невід'ємна: $f(x) \geq 0$.
2. Невласний інтеграл від диференційної функції у межах від $-\infty$ до ∞ дорівнює 1.

7.8.3. Імовірнісний зміст диференційної функції

$$f(x) = F'(x);$$

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x}.$$

Чисельник — це ймовірність того, що X прийме значення, яке належить інтервалу $(x, x + \Delta x)$.

Таким чином, границя відношення ймовірності того, що неперервна випадкова величина набере значення, яке належить інтервалу $(x, x + \Delta x)$, до довжини цього інтервалу (при $\Delta x \rightarrow 0$), дорівнює значенню диференційної функції в точці x .

$f(x)$ в точці x — це густина ймовірності в цій точці.

З теорії диференційних обчислень відомо:

$$F(x + \Delta x) - F(x) \cong dF(x),$$

$$F(x + \Delta x) - F(x) \cong F'(x) dx \Rightarrow F'(x) = f(x),$$

$$dx = \Delta x.$$

$$F(x + \Delta x) - F(x) \cong f(x) \Delta x.$$

Імовірнісний зміст цієї рівності: ймовірність того, що випадкова величина набере значення, яке належить інтервалу $(x, x + \Delta x)$, приблизно дорівнює (з точністю до нескінченно малих вищого порядку відносно Δx) добутку густини ймовірності в точці X на довжину інтервалу Δx .

7.8.4. Закон рівномірного розподілу ймовірностей

Диференційні функції розподілу також називають законами розподілу

Розподіл ймовірностей називають *рівномірним*, якщо на інтервалі, в якому знаходяться всі можливі значення випадкової величини, диференційна функція має сталі значення.

Приклад. Диференційна функція рівномірного розподілу, всі можливі значення якої знаходяться в проміжку (a, b) , на якому вона зберігає постійне значення c .

При цьому: $f(x) = 0$ при $x < a$; $x > b$.

Оскільки всі можливі значення знаходяться в інтервалі $[a, b]$, то справедлива рівність:

$$\int_a^b f(x) dx = 1; \int_a^b c dx = 1; c = \frac{1}{\int_a^b dx} = \frac{1}{b - a};$$

$$f(x) = \begin{cases} 0 & \text{при } x < a \\ \frac{1}{b - a} & a \leq x \leq b \\ 0 & \text{при } x > b \end{cases}.$$

7.8.5. Нормальний закон розподілу

Числові характеристики неперервних випадкових величин

Нехай неперервна випадкова величина X задана диференційною функцією $f(x)$. Нехай усі можливі значення X належать проміжку $[a, b]$. Розіб'ємо відрізок на n часткових відрізків довжиною $\Delta x_1, \Delta x_2, \dots, \Delta x_n$ і виберемо у кожному з них довільну точку $X_i (i = 1, n)$.

Математичні сподівання:

$$\sum x_i f(x_i) \Delta x_i \Rightarrow \int_a^b x f(x) dx.$$

$f(x_i) \Delta x_i$ наближено дорівнює ймовірності влучення x в інтервалі Δx_i .

Математичним сподіванням неперервної випадкової величини X , можливі значення якої розташовані на відрізку $[a, b]$, називають визначений інтеграл:

$$M(x) = \int_a^b x f(x) dx.$$

Якщо можливі значення належать усій осі, то $M(x) = \int_{-\infty}^{\infty} x f(x) dx$.

Дисперсія неперервної випадкової величини відповідно визначається за формулами:

$$D(x) = \int_a^b [x - M(x)]^2 f(x) dx, x \in [a, b]$$

$$D(x) = \int_a^b [x - M(x)]^2 f(x) dx, x \in [-\infty, \infty]$$

$$D(x) = \int_a^b x^2 f(x) dx [M(x)]^2,$$

$$D(x) = \int_{-\infty}^{\infty} x^2 f(x) dx [M(x)]^2.$$

Нормальним називають розподіл імовірностей, який описується диференційною функцією виду:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \text{ де } a \text{ — математичне сподівання, } \sigma = \sqrt{D(x)} \text{ —}$$

середнє квадратичне відхилення випадкової величини.

Нормованим називають розподіл з $a = 0$ і $\sigma = 1$,

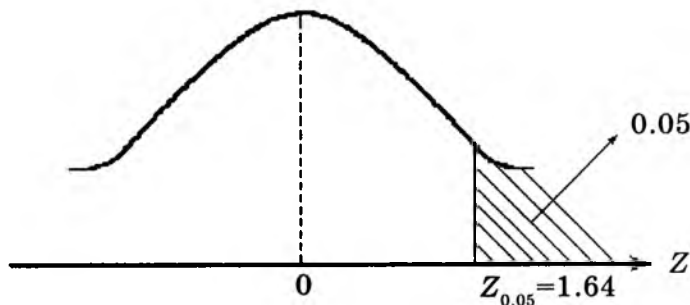
$$t = \frac{x-a}{\sigma} \text{ — нормована випадкова величина з } M(t) = 0, \sigma(t) = 1.$$

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Властивості функції густини нормального закону розподілу.

1. При $x = a$ функція має максимум $= \frac{1}{\sigma\sqrt{2\pi}}$.
2. Графік функції є симетричним відносно прямої $x = a$.
3. Точки графіка

$$\left(a - \sigma, \frac{1}{\sigma\sqrt{2\pi e}}\right) \text{ та } \left(a + \sigma, \frac{1}{\sigma\sqrt{2\pi e}}\right) \text{ — точки перегину.}$$



Малюнок 7.4. Функція густини нормального закону розподілу нормованої випадкової величини

7.8.5.1. Вплив параметрів нормального розподілу на форму нормальної кривої

Як бачимо з малюнка, функція густини нормального закону розподілу симетрична відносно свого математичного сподівання.

Із зростанням математичного сподівання a (на мал.7.4. математичне сподівання дорівнює нулю) **функція густини зсувається праворуч відносно осі x** . **Із спаданням a вона зсувається ліворуч.**

Якщо σ зростає, крива стає більш похилою, тобто опускається до осі x ; **при зменшенні значення σ нормальна крива стає гостроверхою.**

Будь-яку випадкову величину x , розподілену нормально з математичним сподіванням a та дисперсією σ^2 , можна пронормувати, тобто звести до випадкової величини з математичним сподіванням нуль та дисперсією одиниця, шляхом такого перетворення:

$$T = \frac{x-a}{\sigma}, \text{ де } x \text{ — випадкова величина;}$$

a — математичне сподівання цієї випадкової величини;

σ — її середнє квадратичне відхилення $\sigma = \sqrt{\sigma^2}$.

Для нормального закону розподілу 50% значень випадкової величини

лежать у межах $\left(a \pm \frac{2}{3}\sigma\right)$, 68% — у межах $(a \pm \sigma)$ і 95% — у межах $(a \pm 2\sigma)$ та 99.7% — у межах $(a \pm 3\sigma)$.

Щоб проілюструвати нормальний закон розподілу, розглянемо декілька прикладів.

Приклад 1. Показати, що при досить великій кількості класів (B) та при частоті n_i кожного окремого значення x_i змінна $T = \frac{x-a}{\sigma}$ має нульове

математичне сподівання ($t = 0$) і дисперсію $\sigma^2 = 1$.

Розв'язок. За означенням математичного сподівання для дискретних випадкових величин маємо:

$$\bar{t} = \frac{1}{n} \sum_{i=1}^k n_i \left(\frac{x_i - a}{\sigma}\right) = \frac{1}{n} \cdot \frac{1}{\sigma} \sum_{i=1}^k n_i (x_i - a) = \frac{1}{n} \cdot \frac{1}{\sigma} \left(\sum_{i=1}^k n_i x_i - n \cdot a\right) = \frac{1}{n} \cdot \frac{1}{\sigma} \cdot 0 = 0,$$

враховуючи, що $a = \bar{x} = \frac{1}{n} \sum_{i=1}^k n_i x_i$; $\sum_{i=1}^k n_i = n$.

Дисперсія величини T обчислюється за формулою:

$$\text{var}(T) = \frac{1}{n} \sum_{i=1}^k \left(\frac{x_i - a}{\sigma} - 0\right)^2 = \frac{1}{n} \cdot \frac{1}{\sigma^2} \sum_{i=1}^k (x_i - a)^2 = \frac{1}{n} \cdot \frac{1}{\sigma^2} (n\sigma^2) = 1.$$

Приклад 2. Нехай нам відомо, що надходження чеків до банку відповідає нормальному закону розподілу з математичним сподіванням $x=5232$ (грн.) та середнім квадратичним віхиленням $\sigma = 1972$ (грн.). Нам потрібно знайти ймовірність надходження чека, вартість якого перевищує 6000 (грн.). Формалізовано це можна записати таким чином:

X — випадкова величина (вартість чеків);

$X \sim N(5232, 1972)$.

Треба знайти: $P(X > 6000) = ?$

Розв'язок. Пронормуємо випадкову величину x_i та знайдемо значення

$$t_i, \text{ яке відповідає } x_i = 6000: t_i = \frac{6000 - 5232}{1972} = 0.39.$$

Далі використаємо таблицю нормального розподілу. Враховуючи симетричність функції густини нормального закону розподілу, отримаємо ймовірність того, що $P(T > t_i)$ буде дорівнювати $P(T > 0.39) = 0.5 - P(T < 0.39) = 0.5 - 0.1517 = 0.3483$. Отже, ймовірність того, що до банку надійде чек вартістю більше ніж 6000, дорівнює 0.3483. Тепер нас цікавить ймовірність того, що до банку надійде чек на суму меншу ніж 6000 (гривень). Оскільки ми вже маємо $P(x > 6000) = 0.3483$, то знайти $P(x < 6000)$ дуже просто, знаючи, що $P(x > 6000) + P(x < 6000) = 1$. Отже, $P(x < 6000) = 0.6517$.

Але якби нам був невідомий попередній результат, ми б шукали цю ймовірність аналогічно попередньому розрахунку. Справді, для нашого прикладу: $P(T < 0.39) = 0.6517$. З властивості того, що функція густини нормального закону розподілу симетрична відносно свого математичного сподівання, впливає:

$$P(T < -t) = P(T > t) \text{ та } P(T > -t) = P(T < t).$$

Приклад 3. За умов прикладу 2 необхідно знайти ймовірність того, що до банку потрапить чек вартістю нижче ніж 4000 (гривень).

Розв'язок. Нормуючи випадкову величину x , отримаємо значення t :

$$t = \frac{4000 - 5232}{1972} = -0.62,$$

звідки $P(x < 4000) = P(T < -0.62) = 0.2676$.

Приклад 4. За умов прикладу 2 визначимо ймовірність того, що значення випадкової величини потрапить до інтервалу $(x_1; x_2)$. Наприклад, нас цікавить ймовірність того, що вартість чека, який надходить, знаходиться у межах 4000 та 7000 (грн.). Нормування значення 4000 (грн.) та 7000 (грн.) дасть нам:

$$t_1 = \frac{4000 - 5232}{1972} = -0.62; t_2 = \frac{7000 - 5232}{1972} = 0.90.$$

Враховуючи ці межі, можна розписати:

$$\begin{aligned} P(4000 < x < 7000) &= P(t_1 < T < t_2) = P(T < t_2) - P(T < t_1) = \\ &= P(T < t_2) - P(T > -t_1). \end{aligned}$$

За таблицею нормального закону розподілу, зробивши додаткові розрахунки, знаходимо: $P(T > 0.62) = 0.267$ та $P(T < 0.90) = 0.8159$.

Звідси маємо розв'язок: $P(4000 < x < 7000) = 0.8159 - 0.2676 = 0.5483$.

Звичайно, на практиці нас цікавить і протилежна проблема: як, знаючи ймовірність того, чи відбулася подія, встановити її значення. Це можна зробити, використавши таблицю нормального закону розподілу та провівши зворотні перетворення, що можна зобразити за допомогою схеми 1:

$$P(X < x_i) = p = ? \rightarrow t_i = \frac{x_i - \bar{x}}{\sigma} \rightarrow P(T < t_i) = p = ? \rightarrow \text{табл.} \rightarrow \text{відповідь: } p;$$

$$P(X < x_i) = p; x_i = ? \rightarrow \text{табл.} \rightarrow P(T < t_i) = p \rightarrow t_i = \frac{x_i - \bar{x}}{\sigma} \rightarrow \text{відповідь: } x_i.$$

7.8.6. Розподіл χ^2 (x_i — квадрат)

Нехай ми маємо ряд нормально розподілених та незалежних змінних X_1, X_2, \dots, X_v . Пронормуємо їх та отримаємо стандартні нормалізовані значення:

$$Z_1 = \frac{x_1 - \mu_1}{\sigma} \sim N(0,1); Z_2 = \frac{x_2 - \mu_2}{\sigma_2} \sim N(0,1) \dots, Z_v = \frac{x_v - \mu_v}{\sigma_v}.$$

Сума квадратів нормалізованих величин розподіляється за розподілом χ^2 (x_i — квадрат) із ступенями вільності, які дорівнюють v :

$$\chi^2 = \sum_{i=1}^v Z_i^2 = \sum_{i=1}^v \left(\frac{x_i - \mu_i}{\sigma_i} \right)^2,$$

де v — кількість змінних, які ми сумуємо.

Функція густини розподілу χ^2 зміщена праворуч відносно осі ординат, починається на початку координат та прямує до нескінченності. Із зростанням n функція густини стає дедалі симетричнішою. Є спеціальна таблиця розподілу χ^2 , яка є кумульованою: вона дає ймовірність того, що χ^2 більше, ніж якесь значення при певних ступенях вільності та заданому рівні значимості.

Ми символічно можемо записати $\chi^2_{0.025}$ для оцінки χ^2 , для якої 2,5% спостережень матимуть значення більше, ніж χ^2 , тобто будуть знаходитися праворуч від χ^2 на осі абсцис функції густини.

Так, для ступенів вільності $\nu = 10$ та рівня значимості 5% за таблицею χ^2 знаходимо критичне значення 18.3, що можна записати:

$$P\{18.3 < \chi^2 < \infty\} = 0.05 \text{ (для } \nu = 10\text{)}.$$

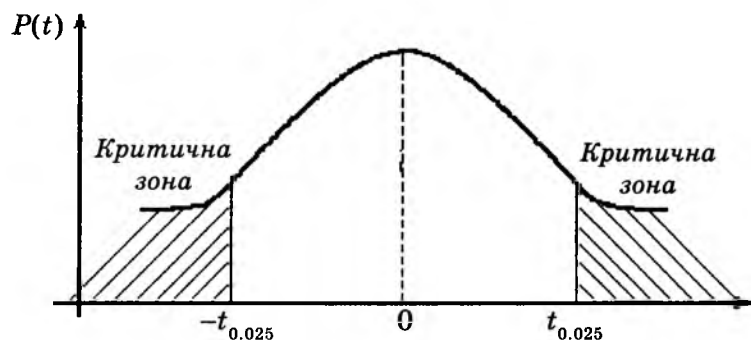
Тобто ймовірність того, що $\chi^2 > 18.3$, дорівнюватиме 0.05 для 10 ступенів вільності.

7.8.7. T-розподіл Ст'юдента

Якщо змінна Z має стандартний нормальний розподіл з нульовим математичним сподіванням та дисперсією 1; $Z_i \sim N(0,1)$, та інша незалежна змінна V^2 розподілена за χ^2 розподілом з ν ступенями вільності, то число $t = Z\sqrt{\nu}/V$ має t -розподіл з ν ступенями вільності.

Характеристики t -розподілу такі:

- 1) t -розподіл є розподілом, який має форму дзвона, симетричного відносно нульового значення (див. мал. 7.5);
- 2) значення t належать проміжку $-\infty < t < +\infty$;
- 3) функція густини t -розподілу більш розтягнута, ніж функція нормального закону розподілу. Це означає, що площа на кінцях більша у t -розподілі, ніж у стандартному нормальному розподілі;
- 4) із збільшенням числа спостережень (n) t -розподіл наближається до стандартного нормального розподілу. Відмітимо, якщо $n > 30$, то t -розподіл Ст'юдента в нормальному законі розподілу;



Малюнок 7.5. Функція густини t -розподілу Ст'юдента

- 5) t -розподіл залежить від ступеня вільності, тобто нам потрібно знайти ступені вільності, щоб отримати критичне значення з t -таблиці. Вона

відрізняється від таблиці стандартного нормального розподілу тим, що включає ступені вільності. Таблиця перераховує значення t , з правої сторони якого ми знаходимо 10, 5, 2.5 та 0.5 відсотка площі під кривою. Символічно ми будемо записувати $t_{0.025}$ для оцінки t , праворуч — площа під кривою, яка становить 2.5%; $t_{0.01}$ для оцінки t , праворуч площа під кривою становить 1% від загальної суми і так далі. Наприклад, припустимо, ступені вільності дорівнюють 15 ($df = 15$). Ми можемо знайти з t -таблиці критичні значення t , які відповідають різним рівням значимості. Для рівня значимості 5% та $df = 15$ маємо: $t_{0.025} = \pm 2.131$.

Це можна розписати таким чином:

$$P\{-2.131 < t < 2.131\} = 0.95 \text{ з 15 ступенями вільності.}$$

Перетворення нормально розподіленої величини x_i і її вибіркового середнього в t -величини виконується за допомогою формул:

$$t = \frac{x_i - \mu}{\hat{\sigma}_x} \quad \text{та} \quad t = \frac{\bar{x}_i - \mu}{\hat{\sigma}_x / \sqrt{n}},$$

де $\hat{\sigma}_x$ — оцінка середнього квадратичного відхилення для малої вибірки ($n < 30$).

$$\hat{\sigma}_x = \sqrt{\hat{\sigma}_x^2},$$

де $\hat{\sigma}_x^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$ — оцінка дисперсії σ_x^2 генеральної сукупності.

7.8.8. F-розподіл Фішера

Якщо дві змінні мають незалежні χ^2 -квадрат розподіли χ_1^2 і χ_2^2 , відповідно з ν_1 і ν_2 ступенями вільності, тоді статистика

$$F = \frac{\chi_1^2 / \nu_1}{\chi_2^2 / \nu_2}$$

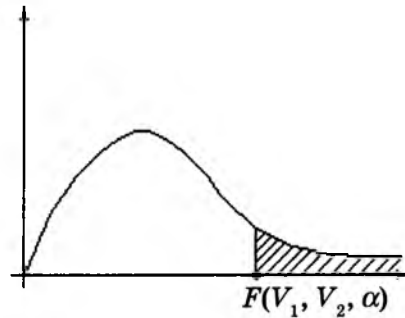
має F -розподіл Фішера з ν_1 і ν_2 ступенями вільності.

F -статистику можна розглядати також як відношення двох незалежних оцінок дисперсій, і тому на практиці F -розподіл найчастіше використовується для тестування рівності цих оцінок дисперсій. З цієї причини F -статистику інколи називають відношенням дисперсій:

$$F = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} = \text{відношення дисперсій}$$

з ν_1 і ν_2 ступенями вільності.

Функція F -розподілу асиметрична (див. мал. 7.6).



Малюнок 7.6. Функція густини F -розподілу Фішера

Слід зазначити, що значення F -відношення завжди додатне і знаходиться в межах від нуля до плюс нескінченності ($0 \leq F \leq +\infty$). Значення F -відношення з різними ступенями вільності (і різними рівнями значимості) зведені в таблицю F -розподілу Фішера. Ступені вільності ν_1 і ν_2 залежать від оцінок дисперсій чисельника та знаменника F -відношення. F -таблиця дає критичні значення праворуч спрямованого хвоста. Якщо дві оцінки дисперсії близькі одна до одної, тоді відношення їх наближається до одиниці. Чим більша різниця між двома дисперсіями, тим більше значення F -відношення. Таким чином, загалом, великі значення F допускають, що різниця між двома оцінками є значною.

Розглянемо приклад перевірки на адекватність багатофакторної регресійної моделі з використанням F -критерію Фішера.

Нагадаємо, що при цьому ми формуємо нуль-гіпотезу рівності всіх параметрів багатофакторної регресії нулеві на протипагу альтернативній гіпотезі, що хоча б один параметр не дорівнює нулю, тобто має місце регресійний зв'язок. При цьому нуль-гіпотеза має вигляд:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$$

на протипагу альтернативній гіпотезі, що

$$H_1: \text{хоча б одне значення } \beta_i \text{ відмінне від нуля.}$$

Якщо нуль-гіпотеза H_0 неправильна, то тоді правильна гіпотеза H_1 , тобто не всі параметри незначно відрізняються від нуля, що дає підставу вважати, що побудована регресійна модель відповідає дійсності, тобто адекватна. Для перевірки H_0 -гіпотези розраховується F -відношення Фішера з p та $(n - p - 1)$ ступенями вільності:

$$F_{p, n-p-1} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 / p}{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n - p - 1)},$$

де p — кількість факторів, які увійшли в модель; n — загальна кількість спостережень.

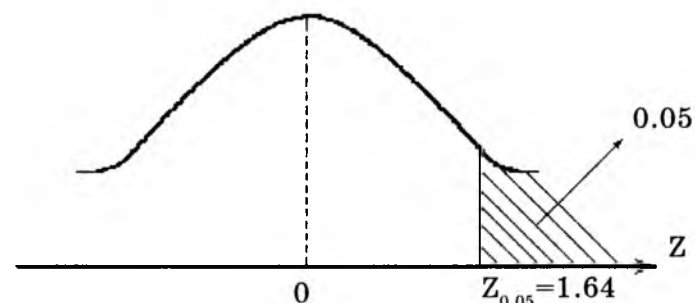
За F -таблицями Фішера, як і у випадку простої регресії, знаходимо критичне значення $F_{кр}$ з p та $(n - p - 1)$ ступенями вільності, задавши попередньо рівень значимості $\alpha \cdot 100\%$ (або рівень довіри $(1 - \alpha) \cdot 100\%$).

Якщо $F > F_{кр}$, тоді нуль-гіпотеза відкидається, що свідчить про адекватність побудованої моделі. У протилежному випадку вона приймається і модель вважається неадекватною.

ДОДАТКИ

ДОДАТОК 1

Графік і таблиця нормального закону розподілу



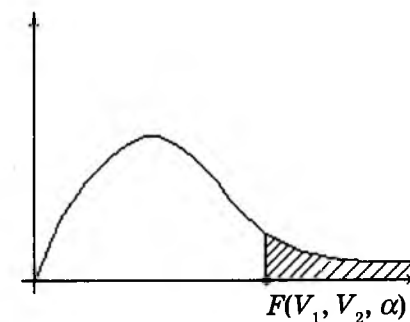
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

ДОДАТОК 2

Функція щільності нормального закону розподілу симетрична відносно осі $z = 0$, тому вся площа під кривою зліва від нуля також дорівнює 0,5. Це дозволяє шляхом досить легких перетворень отримувати ймовірності, які не передбачені в наведеній вище таблиці.

Наприклад, $P(z \leq 1.84) = 0.5 + P(0 \leq z \leq 1.84) = 0.5 + 0.4671 = 0.9671$.

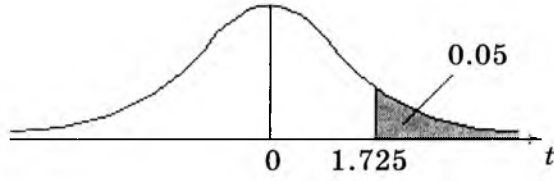
Графік і таблиця F -розподілу Фішера



v_2/v_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,9	245,9	248,0	249,1	250,1	251,1	252,2	253,3	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,41	19,43	19,45	19,45	19,46	19,47	19,48	19,49	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,99	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,64	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,19	2,15	2,11	2,06	2,02	1,97
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,15	2,11	2,07	2,03	1,98	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,07	1,99	1,95	1,90	1,85	1,80	1,75	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,13	2,06	1,97	1,93	1,88	1,84	1,79	1,73	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,04	1,96	1,91	1,87	1,82	1,77	1,71	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,10	2,03	1,94	1,90	1,85	1,81	1,75	1,70	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35	1,25
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00

ДОДАТОК 3

Графік і таблиця t -розподілу Ст'юдента



Df/Pr	0,25 0,50		0,10 0,20		0,05 0,10		0,025 0,05		0,01 0,02		0,005 0,010		0,001 0,002	
	1	1,000	3,078	6,314	12,706	31,821	63,657	318,31						
2	0,861	1,886	2,920	4,303	6,965	9,925	22,327							
3	0,765	1,638	2,353	3,182	4,541	5,841	10,214							
4	0,741	1,533	2,132	2,776	3,747	4,604	7,173							
5	0,727	1,476	2,015	2,571	3,365	4,032	5,893							
6	0,718	1,440	1,943	2,227	3,143	3,707	3,208							
7	0,711	1,415	1,895	2,365	2,998	3,499	4,785							
8	0,706	1,397	1,860	2,306	2,896	3,355	4,501							
9	0,703	1,383	1,833	2,262	2,821	3,250	4,297							
10	0,700	1,372	1,812	2,228	2,764	3,169	4,144							
11	0,697	1,363	1,796	2,201	2,718	3,106	4,025							
12	0,695	1,356	1,782	2,179	2,681	3,055	3,930							
13	0,694	1,350	1,771	2,160	2,650	3,012	3,852							
14	0,692	1,345	1,761	2,145	2,624	2,977	3,787							
15	0,691	1,341	1,753	2,131	2,602	2,947	3,733							
16	0,690	1,337	1,746	2,120	2,583	2,921	3,686							
17	0,689	1,333	1,740	2,110	2,567	2,898	3,646							
18	0,688	1,330	1,734	2,101	2,552	2,878	3,610							
19	0,688	1,328	1,729	2,093	2,539	2,861	3,579							
20	0,687	1,325	1,725	2,086	2,528	2,845	3,552							
21	0,686	1,323	1,721	2,080	2,518	2,831	3,527							
22	0,686	1,321	1,717	2,074	2,508	2,819	3,505							
23	0,685	1,319	1,714	2,069	2,500	2,807	3,485							
24	0,685	1,318	1,711	2,064	2,492	2,797	3,467							
25	0,684	1,316	1,708	2,060	2,485	2,787	3,450							
26	0,684	1,315	1,706	2,056	2,479	2,779	3,435							
27	0,684	1,314	1,703	2,052	2,473	2,771	3,421							
28	0,683	1,313	1,701	2,048	2,467	2,763	3,408							
29	0,683	1,311	1,699	2,045	2,462	2,756	3,396							
30	0,683	1,310	1,697	2,042	2,457	2,750	3,385							
40	0,681	1,303	1,684	2,021	2,423	2,704	3,307							
60	0,679	1,296	1,671	2,000	2,390	2,660	3,232							
120	0,677	1,289	1,658	1,980	2,358	2,617	3,160							
∞	0,674	1,282	1,645	1,960	2,326	2,576	3,090							

Приклад. Для df = 20
 $Pr(t > 2,089) = 0,025$;
 $Pr(t > 1,725) = 0,05$;
 $Pr(|t| > 1,725) = 0,10$.

ДОДАТОК 4

DW-статистика Дарбіна — Уотсона.
 Критичні точки d_L та d_U при рівні значимості $d = 0.05$

n	k=1		k=2		k=3		k=4		k=5		k=6		k=7		k=8		k=9		k=10	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
6	0.610	1.400	0.467	1.896	0.368	2.287	0.296	2.588	0.243	2.822	0.203	3.005	0.171	3.149	0.147	3.266	0.127	3.360	0.111	3.438
7	0.700	1.356	0.467	1.896	0.368	2.287	0.296	2.588	0.243	2.822	0.203	3.005	0.171	3.149	0.147	3.266	0.127	3.360	0.111	3.438
8	0.763	1.332	0.459	1.777	0.368	2.287	0.296	2.588	0.243	2.822	0.203	3.005	0.171	3.149	0.147	3.266	0.127	3.360	0.111	3.438
9	0.824	1.320	0.455	1.628	0.455	2.016	0.376	2.414	0.343	2.645	0.288	2.892	0.230	2.985	0.188	3.111	0.151	3.216	0.111	3.304
10	0.879	1.320	0.455	1.628	0.455	2.016	0.376	2.414	0.343	2.645	0.288	2.892	0.230	2.985	0.188	3.111	0.151	3.216	0.111	3.304
11	0.927	1.324	0.458	1.604	0.595	1.928	0.444	2.283	0.316	2.822	0.266	3.005	0.218	3.062	0.169	3.266	0.134	3.360	0.111	3.438
12	0.971	1.311	0.461	1.579	0.658	1.864	0.512	2.177	0.379	2.506	0.306	2.992	0.247	3.033	0.192	3.266	0.151	3.360	0.111	3.438
13	1.010	1.340	0.461	1.562	0.715	1.816	0.574	2.094	0.445	2.390	0.328	2.892	0.266	3.005	0.218	3.266	0.169	3.360	0.111	3.438
14	1.045	1.350	0.465	1.551	0.767	1.779	0.632	2.030	0.505	2.296	0.349	2.992	0.285	3.062	0.247	3.266	0.188	3.360	0.111	3.438
15	1.077	1.361	0.468	1.543	0.814	1.750	0.685	1.977	0.562	2.220	0.368	2.992	0.306	3.033	0.266	3.266	0.207	3.360	0.134	3.438
16	1.106	1.371	0.468	1.539	0.857	1.728	0.734	1.935	0.615	2.157	0.387	2.992	0.325	3.062	0.285	3.266	0.226	3.360	0.151	3.438
17	1.133	1.381	0.468	1.536	0.897	1.710	0.779	1.900	0.664	2.104	0.406	2.992	0.344	3.091	0.304	3.266	0.245	3.360	0.169	3.438
18	1.158	1.391	0.468	1.535	0.935	1.695	0.820	1.872	0.710	2.060	0.425	2.992	0.363	3.120	0.323	3.266	0.264	3.360	0.188	3.438
19	1.180	1.401	0.474	1.536	0.967	1.685	0.859	1.848	0.752	2.023	0.444	2.992	0.382	3.149	0.342	3.266	0.283	3.360	0.207	3.438
20	1.201	1.411	0.474	1.536	0.998	1.676	0.894	1.828	0.792	1.991	0.463	2.992	0.401	3.178	0.361	3.266	0.302	3.360	0.226	3.438
21	1.221	1.420	0.474	1.538	1.026	1.669	0.927	1.812	0.829	1.964	0.482	2.992	0.420	3.207	0.380	3.266	0.321	3.360	0.245	3.438
22	1.239	1.429	0.474	1.541	1.053	1.664	0.958	1.797	0.863	1.940	0.501	2.992	0.440	3.236	0.400	3.266	0.340	3.360	0.264	3.438
23	1.257	1.437	0.474	1.543	1.078	1.660	0.986	1.785	0.895	1.920	0.520	2.992	0.460	3.265	0.380	3.266	0.360	3.360	0.283	3.438
24	1.273	1.446	0.474	1.546	1.101	1.656	1.013	1.775	0.925	1.902	0.539	2.992	0.480	3.294	0.360	3.266	0.380	3.360	0.302	3.438
25	1.288	1.454	0.474	1.549	1.123	1.654	1.038	1.767	0.953	1.886	0.558	2.992	0.500	3.322	0.340	3.266	0.400	3.360	0.321	3.438
26	1.302	1.461	0.474	1.553	1.148	1.652	1.062	1.759	0.979	1.873	0.577	2.992	0.520	3.351	0.320	3.266	0.420	3.360	0.340	3.438
27	1.316	1.469	0.474	1.556	1.162	1.651	1.084	1.753	0.994	1.861	0.596	2.992	0.540	3.380	0.300	3.266	0.440	3.360	0.360	3.438
28	1.328	1.476	0.474	1.560	1.181	1.650	1.104	1.747	0.928	1.850	0.615	2.992	0.560	3.409	0.280	3.266	0.460	3.360	0.380	3.438
29	1.341	1.483	0.474	1.563	1.198	1.650	1.124	1.743	0.930	1.841	0.634	2.992	0.580	3.438	0.260	3.266	0.480	3.360	0.400	3.438
30	1.352	1.489	0.474	1.567	1.214	1.650	1.143	1.739	0.933	1.833	0.653	2.992	0.600	3.467	0.240	3.266	0.500	3.360	0.420	3.438
31	1.363	1.496	0.474	1.570	1.229	1.650	1.160	1.735	0.936	1.825	0.672	2.992	0.620	3.496	0.220	3.266	0.520	3.360	0.440	3.438
32	1.373	1.502	0.474	1.574	1.244	1.650	1.177	1.732	0.939	1.819	0.691	2.992	0.640	3.525	0.200	3.266	0.540	3.360	0.460	3.438
33	1.383	1.508	0.474	1.577	1.258	1.651	1.193	1.730	0.942	1.813	0.710	2.992	0.660	3.554	0.180	3.266	0.560	3.360	0.480	3.438
34	1.393	1.514	0.474	1.580	1.271	1.652	1.208	1.728	0.945	1.808	0.729	2.992	0.680	3.583	0.160	3.266	0.580	3.360	0.500	3.438
35	1.402	1.519	0.474	1.584	1.283	1.653	1.222	1.726	0.948	1.803	0.748	2.992	0.700	3.612	0.140	3.266	0.600	3.360	0.520	3.438
36	1.411	1.525	0.474	1.587	1.296	1.654	1.236	1.724	0.951	1.800	0.767	2.992	0.720	3.641	0.120	3.266	0.620	3.360	0.540	3.438
37	1.419	1.530	0.474	1.590	1.307	1.655	1.249	1.723	0.954	1.795	0.786	2.992	0.740	3.670	0.100	3.266	0.640	3.360	0.560	3.438
38	1.427	1.535	0.474	1.594	1.318	1.656	1.261	1.722	0.957	1.792	0.805	2.992	0.760	3.700	0.080	3.266	0.660	3.360	0.580	3.438
39	1.435	1.540	0.474	1.597	1.329	1.658	1.273	1.722	0.960	1.789	0.824	2.992	0.780	3.729	0.060	3.266	0.680	3.360	0.600	3.438
40	1.442	1.544	0.474	1.600	1.338	1.659	1.285	1.721	0.963	1.786	0.843	2.992	0.800	3.758	0.040	3.266	0.700	3		

ДОДАТОК 5

DW-статистика Дарбіна — Уотсона. Критичні точки d_L та d_U при рівні значимості $d=0.01$

Table with columns for n (16 to 200) and k (11 to 20). Rows contain critical values d_L and d_U for each k.

Table with columns for n (60 to 200) and k (1 to 10). Rows contain critical values d_L and d_U for each k.

СПИСОК ЛІТЕРАТУРИ

1. Aigner D. J. Basic Econometrics. — New York: Prentice-Hall, Inc., Englewood Cliffs, 1971.
2. Ashley R. On the Usefulness of Macroeconomic Forecasts as Inputs to Forecasting Models //Journal of Forecasting. — 1983. — 2. — P. 211—223.
3. Ashley R. A Simple Test for Regression Parameter Stability //Economic Inquiry. — 1984. — 22. — P. 253—268.
4. Banerjee A., Galbraith J., Dolado J. Dynamic Specification and Linear Transformations of the Autoregressive Distributed Lag Model //Oxford Bulletin of Economics and Statistics. — 1990. — 52. — P. 95—104.
5. Belsley David A., Kuh Edwin, Welsh Roy E. Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. — New York: John Wiley & Sons, Inc., 1980.
6. Berndt Ernst R. The Practice of Econometrics: Classic and Contemporary. — Addison-Wesley, 1991.
7. Bincley J. K., Abbott P. C. The Fixed X Assumption in Econometrics: Can the Textbooks Be Trusted? //American Statistician. — 1987. — 41. — P. 206—214.
8. Bodkin R. G., Klein L. R., Marwah K. A History of Macroeconometric Model-Building. — Brookfield, VT: Edward Elgar, 1991.
9. Bollerslev T. Generalized Autoregressive Conditional Heteroscedasticity //Journal of Econometrics. — 1986. — 31. — P. 307—327.
10. Box G. E. P., Jenkins G. M. Time Series Analysis: Forecasting and Control. — San Francisco: Holden Day, 1970 (revised edition 1976).
11. Breusch T. S. Testing for Autocorrelation in Dynamic Linear Models //Australian Economic Papers. — 1978. — 17. — P. 334—355.
12. Breusch T. S., Pagan A. R. A Simple Test for Heteroscedasticity and Random Coefficient Variation //Econometrica. — 1979. — 47. — P. 187—194.
13. Bridge J. I. Applied Econometrics. — Amsterdam: North-Holland Publishing Company, 1971.
14. Brown R., Durbin J., Evans J. Techniques for Testing the Constancy of Regression Relationships over Time //Journal of the Royal Statistical Society. — 1975. — B37. — P. 149—163.
15. Cameron A., Trivedi P. Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests //Journal of Applied Econometrics. — 1986. — 1. — P. 29—54.
16. Casella G. Leverage and Regression through the Origin //American Statistician. — 1983. — 37. — P. 147—152.
17. Chatfield C. The Analysis of Time Series: an Introduction. London: Chapman and Hall, 1984.
18. Frank C. R., Jr. Statistics and Econometrics. — New York: Holt, Rinehart and Winston, Ink., 1971.
19. Godfrey L. G., McAleer M., McKenzie C. R. Variable Addition and

	k=11	k=12	k=13	k=14	k=15	k=16	k=17	k=18	k=19	k=20
n	16	17	18	19	20	21	22	23	24	25
d _L	0.060	3.446	0.053	3.506	0.053	3.562	0.053	3.618	0.053	3.674
d _T	0.094	3.286	0.075	3.358	0.075	3.420	0.075	3.482	0.075	3.544
d _L	0.113	3.146	0.087	3.218	0.087	3.280	0.087	3.342	0.087	3.404
d _T	0.145	3.023	0.102	3.095	0.102	3.157	0.102	3.219	0.102	3.281
d _L	0.178	2.914	0.131	3.109	0.131	3.171	0.131	3.233	0.131	3.295
d _T	0.212	2.817	0.162	3.004	0.162	3.066	0.162	3.128	0.162	3.190
d _L	0.246	2.729	0.194	2.909	0.194	2.971	0.194	3.033	0.194	3.095
d _T	0.281	2.651	0.227	2.822	0.227	2.884	0.227	2.946	0.227	3.008
d _L	0.315	2.580	0.260	2.744	0.260	2.806	0.260	2.868	0.260	2.930
d _T	0.348	2.517	0.292	2.674	0.292	2.736	0.292	2.798	0.292	2.860
d _L	0.381	2.460	0.324	2.610	0.324	2.672	0.324	2.734	0.324	2.796
d _T	0.413	2.409	0.356	2.552	0.356	2.614	0.356	2.676	0.356	2.738
d _L	0.444	2.363	0.387	2.499	0.387	2.561	0.387	2.623	0.387	2.685
d _T	0.474	2.321	0.417	2.451	0.417	2.513	0.417	2.575	0.417	2.637
d _L	0.503	2.283	0.447	2.407	0.447	2.469	0.447	2.531	0.447	2.593
d _T	0.531	2.248	0.475	2.367	0.475	2.429	0.475	2.491	0.475	2.553
d _L	0.558	2.216	0.503	2.330	0.503	2.392	0.503	2.454	0.503	2.516
d _T	0.585	2.187	0.530	2.296	0.530	2.358	0.530	2.420	0.530	2.482
d _L	0.610	2.160	0.556	2.266	0.556	2.328	0.556	2.390	0.556	2.452
d _T	0.634	2.136	0.581	2.237	0.581	2.299	0.581	2.361	0.581	2.423
d _L	0.658	2.113	0.605	2.210	0.605	2.272	0.605	2.334	0.605	2.396
d _T	0.680	2.092	0.628	2.186	0.628	2.248	0.628	2.310	0.628	2.372
d _L	0.702	2.073	0.651	2.164	0.651	2.226	0.651	2.292	0.651	2.354
d _T	0.723	2.055	0.673	2.143	0.673	2.205	0.673	2.271	0.673	2.333
d _L	0.744	2.039	0.694	2.123	0.694	2.185	0.694	2.251	0.694	2.313
d _T	0.835	1.972	0.790	2.044	0.790	2.106	0.790	2.172	0.790	2.248
d _L	0.913	1.925	0.871	1.987	0.871	2.048	0.871	2.114	0.871	2.189
d _T	0.979	1.891	0.940	1.945	0.940	2.010	0.940	2.076	0.940	2.151
d _L	1.037	1.865	1.001	1.914	1.001	1.972	1.001	2.038	1.001	2.113
d _T	1.087	1.845	1.053	1.889	1.053	1.947	1.053	2.012	1.053	2.087
d _L	1.131	1.831	1.099	1.870	1.099	1.932	1.099	1.998	1.099	2.071
d _T	1.170	1.819	1.141	1.856	1.141	1.918	1.141	1.984	1.141	2.055
d _L	1.205	1.810	1.177	1.844	1.177	1.906	1.177	1.972	1.177	2.049
d _T	1.236	1.803	1.210	1.834	1.210	1.894	1.210	1.960	1.210	2.043
d _L	1.264	1.798	1.240	1.827	1.240	1.887	1.240	1.953	1.240	2.037
d _T	1.290	1.793	1.267	1.821	1.267	1.881	1.267	1.947	1.267	2.031
d _L	1.314	1.790	1.292	1.816	1.292	1.876	1.292	1.942	1.292	2.025
d _T	1.473	1.783	1.458	1.799	1.458	1.863	1.458	1.929	1.458	1.919
d _L	1.561	1.791	1.550	1.801	1.550	1.865	1.550	1.931	1.550	1.925

Lagrange Multiplier Tests for Linear and Logarithmic Regression Models // Review of Economics and Statistics. — 1988. — 70. — P. 492—503.

20. Goldberger A. S. A Course in Econometrics. — Cambridge, Mass.: Harvard University Press, 1991.

21. Goldberger A. S. Best Linear Unbiased Prediction in the Generalized Linear Regression Model // Journal of the American Statistical Association. — 1962. — 57. — P. 369—375.

22. Goldberger A. S. Topics in Regression Analysis. — New York: The Macmillan Company, 1968.

23. Goldfeld S. M., Quandt R. E. Nonlinear Methods of Econometrics. — Amsterdam: North-Holland Publishing Company, 1972.

24. Goldfeld S., Quandt R. Studies in Nonlinear Estimation. — Cambridge, MA: Ballinger, 1976.

25. Granger C. W. J., Newbold P. Forecasting Economic Time Series. — 2nd ed. — London: Academic Press, 1986.

26. Granger C., Uhlig H. Reasonable Extreme-Bounds Analysis // Journal of Econometrics. — 1990. — 44. — P. 159—170.

27. Greene William H. Econometric Analysis. — New York: Macmillan Publishing Company, 1990.

28. Griliches Z. Economic Data Issues // In Griliches Z., Intriligator M. D. (eds). Handbook of Econometrics. — Amsterdam: North Holland, 1986. — Vol. III. — Ch. 25.

29. Griliches Z., Intriligator M. D. (eds). Handbook of Econometrics. — Amsterdam: North Holland, 1983. — Vol. I.

30. Griliches Z., Intriligator M. D. (eds). Handbook of Econometrics. — Amsterdam: North Holland, 1984. — Vol. II.

31. Griliches, Z., Intriligator M. D. (eds). Handbook of Econometrics. — Amsterdam: North Holland, 1986. — Vol. III.

32. Gujarati D. Basic Econometrics. — New York: McGraw-Hill Book Company, 1978.

33. Hackl P., Westlund A. H. Statistical Analysis of Structural Change: an Annotated Bibliography // Empirical Economics. — 1989. — 14. — P. 167—192.

34. Halvorsen R., Palmquist R. The Interpretation of Dummy Variables in Semilogarithmic Equations // American Economic Review. — 1980. — 70. — P. 474—475.

35. Harvey A. C. The Econometric Analysis of Time Series. — 2d ed. — Cambridge, Mass: MIT, Oxford: Philip Allan, 1990.

36. Hendry D. F. Econometrics — Alchemy or Science? // Economica. — 1980. — 47. — P. 387—406.

37. Hendry D. F., Ericsson N. R. An Econometric Analysis of UK Money Demand in Monetary Trends in the United States and the United Kingdom by Milton Friedman and Anna J. Schwartz // American Economic Review. — 1991. — 81. — P. 8—38.

38. Hendry D. F., Mizon G. E. Serial Correlation as a Convenient Simplification, Not a Nuisance: a Comment on a Study of the Demand for Money by the Bank of England // Economic Journal. — 1978. — 88. — P. 549—563.

39. Hendry D., Pagan A., Sargan D. Dynamic Specification // In Griliches Z., Intriligator M. D. (eds). Handbook of Econometrics. — Amsterdam: North Holland, 1984. — Vol. II. — Ch. 18.

40. Hendry D. F., Wallis K. F. (eds). Econometrics and Quantitative Economics. — Oxford: Basil Blackwell, 1984.

41. Hoff J. C. A Practical Guide to Box-Jenkins Forecasting. — Belmont, CA: Lifetime Learning, 1983.

42. Hoffman D., Schmidt P. Testing the Restrictions Implied by the Rational Expectations Hypothesis // Journal of Econometrics. — 1981. — 15. — P. 265—288.

43. Hu Teh-Wei. Econometrics: An Introductory Analysis. — Baltimore: University Park Press, 1973.

44. Johnston J. Econometric Methods. — 3d ed. — New York: McGraw-Hill Book Company, 1984.

45. Judge George G., Hill Carter R., Griffiths William E., Lutkepohl Helmut, Lee Tsoung-Chao. Introduction to the Theory and Practice of Econometrics. — New York: John Wiley & Sons, Inc., 1982.

46. Judge George G., Hill Carter R., Griffiths William E., Lutkepohl Helmut, Lee Tsoung-Chao. Theory and Practice of Econometrics. — New York: John Wiley & Sons, Inc., 1980.

47. Kelejian H. A., Oates W. E. Introduction to Econometrics: Principles and Applications. — 2d ed. — New York: Harper & Row, Publishers, Incorporated, 1981.

48. Kennedy Peter. A Guide to Econometrics. — 2d ed. — Cambridge, Mass.: MIT Press, 1985.

49. Koutsoyiannis A. Theory of Econometrics. — New York: Harper & Row, Publishers, Incorporated, 1973.

50. McNees S. Forecasting Accuracy of Alternative Techniques: a Comparison of U. S. Macroeconomic Forecasts // Journal of Business and Economic Statistics. — 1986. — 4. — P. 5—15.

51. Maddala G. S. Econometrics. — New York: McGraw-Hill Book Company, 1977.

52. Maddala G. S. Introduction to Econometrics. — New York: Macmillan, 1988.

53. Malinvaud E. Statistical Methods of Econometrics. — 2d ed. — Amsterdam: North-Holland Publishing Company, 1976.

54. Malley J. R. Dynamic Specification in Econometric Estimation // Journal of Agricultural Economics Research. — 1990. — 42(2). — P. 52—55.

55. Manski C. F. Regression // Journal of Economic Literature. — 1991. — 29. — P. 34—50.

56. Mark Stewart B., Wallis Kenneth F. *Introductory Econometrics*. — 2d ed. — New York: John Wiley & Sons, Inc., 1981. (A Halsted Press Book.)
57. Mayer L. S. *The Use of Exploratory Methods in Economic Analysis: Analyzing Residential Energy Demand* //In Kmenta J., Ramsey J. B. (eds). *Evaluation and Econometric Models*. — New York: Academic Press, 1980. — P. 15—45 (commentary by V. K. Smith, p. 123—128).
58. Mills T. C. *Time Series Techniques for Economists*. — Cambridge: Cambridge University Press, 1990.
59. Mills T. C. *Nonlinear Time Series Models in Economics* //Journal of Economic Surveys, forthcoming. — 1991.
60. Moon C. G. *A Monte Carlo Comparison of Semi parametric Tobit Estimators* //Journal of Applied Econometrics. — 1989. — 4. — P. 1361—1382.
61. Morgan M. *The History of Econometric Ideas*. — Cambridge: Cambridge University Press, 1990.
62. Morgan M. *Perspectives in the History of Econometrics: a Review Essay of R. J. Epstein: A History of Econometrics* //Econometric Theory. — 1990. — 6. — P. 151—164.
63. Mullahy J. *Economic Estimation of Microeconomic Recreation Participation Models* //In Pearce D. W., Rau N. J. (eds). *Economic Perspectives*. — Caur, Switzerland: Harwood, 1986. — Vol. 4. — P. 91—150.
64. Murphy James L. *Introductory Econometrics*. — Homewood, Ill.: Richard D. Irwin, Inc., 1973.
65. Netter J., Wasserman W. *Applied Linear Statistical Models*. — Homewood, Ill.: Richard D. Irwin, Inc., 1974.
66. Pindyck R. S., Rubinfeld D. L. *Econometric Models and Econometric Forecasts*. — 3d ed. — New York: McGraw-Hill Book Company, 1990.
67. Press S. J. *Bayesian Computer Programs* //In Zellner A. (ed.). *Bayesian Analysis in Econometrics and Statistics: Essays in Honor of Harold Jeffreys*. — Amsterdam: North Holland, 1980. — Ch. 27.
68. Rao C. R. *Linear Statistical Inference and Its Applications*. — 2d ed. — New York: John Wiley & Sons, 1975.
69. Rao Potluri, Miller Roger LeRoy. *Applied Econometrics*. — Belmont, Calif: Wadsworth Publishing Company, Inc., 1971.
70. Robinson P. M. *Non-parametric Methods in Specification* //Economic Journal. — 1986. — 96, supplement. — P. 134—141.
71. Robinson P. M. *Semi-Parametric Econometrics: a Survey* //Journal of Applied Econometrics. — 1988. — 3. — P. 35—51.
72. Seaks T. J., Vines D. P. *A Monte Carlo Evaluation of the Box-Cox Difference Transformation* //Review of Economics and Statistics. — 1990. — 72. — P. 506—510.
73. Sims C. A. *Macroeconomics and Reality* //Econometrica. — 1980. — 48. — P. 1—47.
74. Smith M. A., Smith D. J. *Choosing among Multiple Nonlinear Nonnested Regression Models with Different Dependent Variables* //Economics

Letters. — 1990. — 34. — P. 147—150.

75. Watson M. W., Engle R. F. *Testing for Regression Coefficient Stability with a Stationary AR(1) Alternative* //Review of Economics and Statistics. — 1985. — 67. — P. 341—346.

76. Welsch R. E. *Regression Sensitivity Analysis and Bounded-influence Estimation* //In Kmenta J., Ramsey J. (eds). *Evaluation of Econometric Models*. — New York: Academic Press, 1980. — P. 153—167.

77. White H. *A Heteroscedasticity-consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity* //Econometrica. — 1980. — 48. — P. 817—838.

78. Wickens M., Breusch T. *Dynamic Specification, the Long-run and the Estimation of Transformed Regression Models* //Economic Journal. — 1988. — 98, supplement. — P. 189—205.

79. Wonnacott R. J., Wonnacott T. H. *Econometrics*. — 2d ed. — New York: John Wiley & Sons, Inc., 1979.

80. Wooldridge J. M. *A Note on the Lagrange Multiplier and F-statistics for Two Stage Least Squares Regressions* //Economics Letters. — 1990. — 34. — P. 151—155.

81. Zellner A. *An Introduction to Bayesian Inference in Econometrics*. — New York: John Wiley & Sons, Inc., 1971.

82. Zellner A. *Basic Issues in Econometrics*. — Chicago: University of Chicago Press, 1984.

83. Zellner A. *The ET Interview* //Econometric Theory. — 1989. — 5. — P. 287—317.

84. Zellner A., Richard J. F. *Use of Prior Information in the Analysis and Estimation of Cobb-Douglas Production Function Models* //International Economic Review. — 1973. — 14. — P. 107—119.

85. Винн Р., Холден К. *Введение в прикладной эконометрический анализ*. — М., 1981. — 294 с.

86. Грубер Й. *Эконометрия. Учебное пособие для студентов экономических специальностей*. — К., 1996. — Т. 1. Введение в эконометрию. — 400 с.

87. Джонстон Дж. *Эконометрические методы*. — М., 1980. — 444 с.

88. Дрейпер Смит. *Прикладной регрессионный анализ*. — Москва: Мир, 1988. — Т. 1—2.

89. Фишер Ф. *Проблема идентификации в эконометрии*. — М., 1978. — 223 с.

90. Клас А., Герики К., Колен Ю., Шуян И. *Введение в эконометрическое моделирование*. — М., 1978. — 152 с.

91. Маленбо Э. *Статистические методы в эконометрии*. — Вып. 1: М., 1975. — 423 с.; Вып. 2: М., 1976. — 325 с.

92. Тинтнер Г. *Введение в эконометрию*. — М., 1965. — 361 с.

Навчальне видання

Ірина Григорівна Лук'яненко
Лариса Іванівна Краснікова

ЕКОНОМЕТРИКА

Підручник

Керівник видавничих проектів В. І. Карасьов
Відповідальний редактор М. М. Садовий
Редактори М. Г. Овдієнко-Обухова, В. П. Розумний, Л. В. Кирпич
Коректор В. Д. Мозолєвська
Комп'ютерна верстка С. Ф. Кужанов

Книгу можна придбати за адресами:

*м. Київ, вул. Хрещатик, 44 (книгарня "Знання"),
тел. 224-82-19;*

*м. Київ, пл. Лесі Українки, 1,
тел. 296-80-20;*

*м. Донецьк, вул. Артема, 147-а ("Дом книги"),
тел. (0622) 57-73-94.*

**Книготорговельним організаціям звертатись за тел.:
(044) 224-23-36, 224-80-43.**

Підписано до друку 16.10.1997. Формат 70x100 1/16.
Друк офсетний. Папір офсетний № 1. Гарнітура шкільна.
Авт. друк. арк. 35. Обл.-вид. арк. 31.
Замовлення № 8-290.

Київська обласна організація товариства "Знання" України
252034, Київ, вул. Стрілецька, 28
Тел. (044) 224-80-43, 224-23-36

ВАТ "Книжкова друкарська фабрика ім. Шевченка"
м. Київ, вул. Боггову

НБ ПНУС



611273