

**КАБІНЕТ МІНІСТРІВ УКРАЇНИ**  
**НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ І**  
**ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ**

**Факультет технології виробництва і переробки продукції тваринництва**

**Кафедра розведення та генетики тварин**  
**ім. М.А. Кравченка**

**ОСНОВИ БІОМЕТРІЇ**

---

**для лабораторних і самостійних робіт студентів спеціальності “ТВППТ”**

УДК 57.087.1:636

Наведено методичні вказівки, завдання до вивчення розділу “Біометрія” дисципліни “Генетика з біометрією”.

Для студентів ОКР «Бакалавр» зі спеціальності 6.090.102 – «Технологія виробництва і переробки продукції тваринництва».

Рекомендовано навчально-методичною радою факультету технології виробництва і переробки продукції тваринництва, протокол №10 від 15 квітня 2010 року.

Укладачі: Л.М. Хмельничий, І.О. Супрун,

Рецензенти: Коваленко В.П., Касаткін Д.Ю.

Навчальне видання

## **ОСНОВИ БІОМЕТРІЇ**

Методичні вказівки до виконання лабораторних та самостійних робіт студентам освітньо-кваліфікаційного рівня «Бакалавр» спеціальність – 6.090102 “Технологія виробництва і переробки продукції тваринництва” та 6.110.100 “Ветеринарна медицина”

Укладачі: Л.М. Хмельничий, І.О. Супрун,

Відповідальна за випуск І.О. Супрун

Зав. видавничим центром НУБіП України А.П. Колесніков

Видання здійснено за авторським редагуванням

Підписано до друку 20.09 2010 Формат 60 x 84 1/16.

Ум. друк. арк. 5. Обл.-вид. арк. 5,5.

Наклад 50 пр. Зам. №

Видавничий центр НУБіП України.

м. Київ, вул. Героїв Оборони, 15, 03041

## ЗМІСТ

1.	Математичні основи біометрії.....	4
2.	Мінливість та її класифікація.....	9
3.	Репрезентативність та похибки показників вибірових сукупностей.....	19
	Приклади визначення біометричних показників.....	24
4.	Кореляційний аналіз.....	35
5.	Регресійний аналіз.....	46
	Приклади розв'язування задач.....	47
6.	Дисперсійний аналіз.....	49
	Приклади розв'язування задач.....	56
7.	Успадкування ознак та методи їх визначення.....	60
8.	Перевірка генетичних гіпотез методом Хі-квадрат.....	67
	Приклади розв'язування задач.....	68
	Задачі для самостійного розв'язування.....	70
	Контрольні запитання.....	72
	Бібліографічний список.....	72
	Додатки.....	73

# 1. МАТЕМАТИЧНІ ОСНОВИ БІОМЕТРІЇ

Без систематизації та належного опрацювання даних біологічних експериментів, без поглибленого та всебічного їхнього аналізу неможливо одержати достовірну інформацію, яка ґрунтується на фактичних даних цих досліджень. Вдале поєднання біології та математики, що визначає сутність біометрії, як науки, дозволяє ефективно вирішити цю проблему.

Біометрія виникла в процесі розвитку біології. Вперше математичні методи біометрії розробили у XIX столітті Френсіс Гальтон та Карл Пірсон. У XX столітті з'явилися праці В. Госета, який писав під псевдонімом Стьюдент, з розробкою теорії малих вибірок. Ф. Фішер розробив метод дисперсійного аналізу.

**Біометрія** (біологічна, варіаційна статистика, від *біо...* та грецького *μετρέω* – вимірюю) – наука про статистичний аналіз групових властивостей біологічних об'єктів. Біометрію визначають також як науку про рух групової інформації в популяціях. Підставою для використання біометрії в біології було встановлення кардинального факту, що багатьом біологічним процесам притаманні статистичні закономірності (ймовірність прояву, закон розподілення варіант за наявності їхньої великої кількості у вибірковій сукупності тощо). Предметом біометрії є група біологічних об'єктів, яка називається **сукупністю**. Розрізняють генеральну сукупність та вибірку з неї.

**Генеральна сукупність** – це весь масив об'єктів однієї категорії, подібних за однаковими ознаками і що різняться за іншими.

Наприклад, під генеральною сукупністю української червоно-рябої молочної породи слід розуміти всіх тварин, які відносяться до цієї породи, а також її зональні (центральный, південно-західний та прикарпатський) та заводські (вінницький, київський, прилуцький, харківський та черкаський) типи. Обсяг генеральної сукупності може бути дуже великим (від 5 – 100 тис. голів до кількох мільйонів), але обсяг буває і дуже малим (при вивченні нечисленних порід, генофондних стад). При вивченні фізіологічних параметрів до генеральної сукупності відносять загальну кількість еритроцитів, лейкоцитів у крові однієї тварини.

Дані, одержані при вивченні генеральної сукупності, найбільш чітко відображають характеристику тварин, яких вивчають, але вивчення будь-яких ознак усіх особин популяції надзвичайно важкі у практичному виконанні та досить часто вони дорого коштують. Тому охарактеризувати всю генеральну сукупність за ознаками, що вивчають, в більшості випадків неможливо. А при

вивченні показників, які визначаються тільки після забою (якість м'яса, розвиток внутрішніх органів), вивчення генеральної сукупності взагалі неможливе, оскільки це спричинило б до її знищення. Тому з генеральної сукупності відокремлюють вибірккову сукупність і вивчають її.

**Вибірка** – це частина типових представників генеральної сукупності, яка її відображає, тобто, це частина від цілого. Відбір даних до неї за особинами, членами вибірки, проводиться за принципом рендемічності, випадковості. Наприклад в свинарстві, враховується кожний 5-й, 10-й, 15-й, 20-й підсвинок з кожної породи, які знаходяться на контрольній відгодівлі. Наступний варіант полягає у послідовності включення даних у вибірку з таблиці випадкових чисел.

Вихідні дані беруться з матеріалів первинного зоотехнічного обліку або з результатів досліджень. Вибірка складається з окремих варіантів, під якими розуміють величину ознаки для одиниці сукупності (величина ознаки – надій, одиниця сукупності – група корів певної породи).

**Ознака** – елементарна особливість чи властивість організму, яка може бути визначена і відокремлена від багатьох інших.

Варіанти позначають символами  $V_1, V_2 \dots V_n$  (або  $X_1, X_2 \dots X_n$ ).

Розмір генеральної сукупності позначається великою літерою  $N$ . Число варіант у вибірці називають її обсягом і позначають символом  $n$ .

Розрізняють вибірки малі (при  $n < 30$ ) і великі ( $n > 30$ ). Це пов'язано з технікою (формулами) обчислень, але при використанні ЕВМ всі вибірки можна вважати малими.

До варіант можуть бути віднесені *кількісні* й *якісні* ознаки. До *кількісних* ознак відносять такі, що можна безпосередньо виміряти або підрахувати (наприклад, на вагах визначити живу масу кожного теляти в групі, при контрольному доїнні визначити надій за добу індивідуально для кожної корови, тощо). Ці ознаки характеризуються безперервною мінливістю, їх величини залежать від дії багатьох генів.

**Якісні** ознаки характеризуються окремими описовими визначеннями – забарвлення, масть, здорова чи хвора тварина, самець або самка. У випадку, коли існують два взаємовиключаючі варіанти, ознаки називаються альтернативними (чорний-білий, рогатий-комолий).

Розподіл ознак на якісні і кількісні умовний, оскільки при сучасній техніці вимірювання можна більшість якісних ознак описати кількісно (забарвлення – за кількістю пігменту, вмісту каратиноїдів).

### 1.1. Середні величини прояву ознак

Експериментальні дослідження у біології відрізняються цифровими значеннями різної варіативності, усвідомлення яких потребує використання відповідних математичних обробок та систематизації.

Однією із найважливіших характеристик варіюючих ознак експериментів є їхня середня величина. Середній показник, нівелюючи значення кожної окремо взятої ознаки, дозволяє визначити середню величину вибіркової сукупності, за якою можна порівнювати окремі піддослідні групи тварин, стада та популяції.

Аналіз вибіркової сукупності починається з групування даних, під яким розуміють процес систематизації, або розташування виражених кількостями даних з метою одержання вміщеної в них інформації, встановлення закономірності, якій підпорядковується ознака, що вивчають, або явище.

При роботі з кількісними ознаками і при великому числі варіант групування виконують за класами варіаційного ряду.

**Варіаційний** (ранжирований) ряд – це упорядковане розміщення варіант у сукупності відповідно до наростання або спадання їх чисельних значень.

**Середні величини** – це абстрактні, цілі або дробові кількості, які кількісно характеризують ознаки тієї чи іншої сукупності.

Основна мета обчислення середніх значень у селекційній роботі – це характеристика селекційних груп, стад, ліній, родин за рівнем продуктивності або іншим біологічним показником. Порівнюючи середні двох послідовних поколінь, селекціонер може визначити фактично одержаний генетичний прогрес. Важлива роль відводиться середнім величинам при оцінці плідників за якістю потомства.

Існує декілька видів середніх показників. До них належать: середня арифметична, середня зважена, середня геометрична, середня квадратична, середня гармонічна, середня кубічна, мода і медіана.

**Середня арифметична**, позначається буквою  $M$  або  $\bar{X}$  – це одна з основних характеристик вибіркової сукупності, яку найчастіше визначають у біологічних дослідженнях. Вона має кілька властивостей. Перша – займає середнє місце і знаходиться у центрі розподілу варіантів. Друга властивість полягає у тому, що середній показник може біти величиною абстрактною, тобто приймати значення, яке

не існує у природі, наприклад, багатоплідність – 8,2 поросяти. Головна властивість стверджує, що сума відхилення всіх варіант від середньої арифметичної завжди дорівнює нулю, тобто можна скласти формулу:

$$\Sigma (Vi - M) = 0 \quad (1)$$

де:  $\Sigma$  – знак суми,  $Vi$  – індивідуальні значення варіант.

Середня арифметична є величина поіменована і виражається у тих самих одиницях, що й варіанти.

Для вирахування середньої арифметичної для малих вибірок використовують наступну формулу:

$$M = \frac{V_1 + V_2 + V_3 + \dots + V_n}{n} = \frac{\Sigma V}{n} \quad (2)$$

**Середня зважена** – ( $M_{зг}$ ), розраховують за даними середніх арифметичних кількох вибірок, тобто коли виникає необхідність одержання узагальнених даних про продуктивність, наприклад, ліній за даними окремих родин, продуктивність лінійного та гібридного потомства. Коли необхідно вирахувати середню величину варіант, які знаходяться в залежності одна від одної, наприклад середній відсоток жиру в молоці за лактацію за спостереженнями певної кількості місяців. Крім того, використовується, коли необхідно знайти сумарну середню арифметичну з декількох уже відомих середніх різних за обсягом вибірок. Загальна формула середньої зваженої така:

$$M_{зг} = \frac{V_1v_1 + V_2v_2 + \dots + V_iv_i}{V_1 + V_2 + \dots + V_i} = \frac{\Sigma V_iv_i}{\Sigma V_i} \quad (3)$$

де:  $V_i$  – величина першого показника,  $v_i$  – величина другого показника, зв'язаного з першим.

**Середня геометрична** ( $G$ ) розраховується у тих випадках, коли необхідно визначити на скільки збільшується чи зменшується середній показник сукупності через відповідний проміжок часу. Наприклад при вивченні темпів приросту живої маси молодняку, росту популяцій тварин за певний період. Середня геометрична дає більш точну характеристику при визначенні середньодобових приростів, збільшення лінійних промірів.

**Середня квадратична** ( $S$ ). Існує серед досліджень багато показників, які характеризуються площиною, діаметром, радіусом. Наприклад, площа колоній мікроорганізмів, площа в'язевого вічка тощо. Для вирахування середньої квадратичної використовують наступну формулу:

$$S = \sqrt{\frac{\sum V^2}{n}} \quad (4)$$

де:  $V^2$  – значення конкретної варіанти взяте у квадраті,  $n$  – кількість варіант вибіркової сукупності.

**Середня гармонічна** ( $M_h$ ). Використовується для обчислення середніх ознак, які характеризують показники мінливих швидкостей.

**Середня кубічна** ( $M_q$ ). Використовується коли необхідно визначити середній розмір об'ємних ознак, наприклад, діаметр яєць.

**Медіаною** ( $Me$ ) називається значення ознаки, яка займає середнє положення і ділить весь розподіл на дві рівні за чисельністю частини. Тобто її значення практично не відрізняється від середньої арифметичної.

**Мода** ( $Mo$ ) – це значення ознаки, яке найчастіше зустрічається в даному варіаційному ряді. Ця величина знаходиться в межах модального класу і за своїм значенням є близькою до середньої арифметичної чи середньої зваженої.



## 2. МІНЛИВІСТЬ ТА ЇЇ КЛАСИФІКАЦІЯ

Властивість організмів змінюватись під впливом спадкових і не спадкових факторів називається *мінливістю*. Вона визначає різницю між спорідненими особинами одного чи кількох поколінь, між батьками і нащадками. Мінливість – це процес взаємозв'язку організму з середовищем, що є основним джерелом для природного і штучного добору та одним із головних факторів еволюції.

**Види мінливості.** Відмінності між окремими особинами і групами особин можуть бути двоякі. Одні передаються від покоління до покоління – це спадкова мінливість. Інші можуть існувати лише упродовж життя одного покоління – це не спадкова мінливість (*модифікаційна*). В свою чергу, спадкова мінливість поділяється на комбінаційну, мутаційну, онтогенетичну і кореляційну.

**Комбінаційна мінливість.** У даному випадку нові спадкові поєднання ознак у потомстві виникають у результаті перекомбінацій (рекомбінацій) батьківської і материнської форм. **Рекомбінація** – це перерозподіл генетичної інформації у нащадків. Цей вид мінливості широко використовується у сільському господарстві при створенні нових порід тварин і сортів рослин. Суть комбінаційної мінливості виражена законом Г. Менделя про незалежне успадкування ознак. Причиною комбінаційної мінливості може бути кросинговер, який визначає перекомбінацію генів у групах зчеплення батьківських хромосом.

У результаті поєднання спадкової інформації батьківських форм в онтогенезі потомка ознаки обох батьків проявляються у найрізноманітніших варіантах. Наприклад, при створенні нової породи ВРХ часто схрещують батьківські форми з альтернативними ознаками продуктивності. Так, для поліпшення молочних якостей у тварин комбінованого м'ясо-молочного типу їх схрещують з високоспеціалізованою молочною породою для одержання помісей з кращими ознаками молочності.

**Мутаційна мінливість.** Мутаціями називають зміну окремих ознак і властивостей або їхнього комплексу, які виникають у результаті дії мутагенних факторів на спадковий апарат клітини. Під впливом мутагенів змінюється спадкова інформація, що контролює відповідну ознаку. Процес виникнення мутацій називають мутагенезом, а організм, у якого виникла мутація тієї чи іншої ознаки, – мутантом. Мутації можуть виникати випадково (їх називають спонтанними) або вони можуть бути викликані дією на тварину чи рослину різними мутагенами (це – індуковані мутації).

Індукована мутаційна мінливість досить часто використовується у селекції рослин для створення нових сортів.

**Онтогенетична мінливість** – це сукупність послідовних змін ознак і властивостей особини в процесі її індивідуального розвитку (онтогенезу). В онтогенезі особини реалізується спадкова інформація, одержана від батьків, шляхом послідовної сумісної дії комплексів генів. В результаті у тварини чи рослини формуються органи, ознаки і властивості, характерні для даного виду і притаманні тільки цим особинам. В онтогенетичній мінливості важлива роль належить морфогенетичним кореляціям, завдяки яким зберігається чітко визначений тип розвитку органів і їхніх ознак. В процесі онтогенезу кожна ознака формується самостійно, але у чіткій відповідності з генетично детермінованим (визначеним) загальним планом розвитку даної особини.

**Кореляційна мінливість** – це взаємозалежність між розвитком двох ознак, коли із зміною однієї ознаки відповідною мірою змінюється інша. Термін “кореляція” походить від латинського *correlation* – *співвідношення*. В залежності від мінливості або ступеня розвитку корельованих ознак співвідносна мінливість (кореляція) може бути **позитивною** або **негативною**. Наприклад, із збільшенням живої маси зростає надій – це позитивна кореляція, із збільшенням надою, як правило, знижується вміст жиру – тут кореляція негативна.

У математичному відношенні кореляція зі знаком “+” буде **додатною**, коли із збільшенням однієї ознаки зростає інша, і **від’ємною** із знаком “–”, коли із збільшенням однієї ознаки інша зменшується, або навпаки.

У селекційному відношенні позитивна кореляція є бажаною, а негативна – ні, проте не завжди. Знак “–” перед коефіцієнтом кореляції не завжди адекватно характеризує негативну взаємозалежність між селекціонованими ознаками. Тобто, іноді від’ємна кореляція може мати бажаний селекційний ефект, наприклад від’ємна кореляція між інтенсивністю молоковіддачі та тривалістю доїння ( $r = -0,500$  і вище) свідчить, що із збільшенням інтенсивності молоковіддачі скорочуються витрати часу на видоювання корови. Із скороченням сервіс-періоду корови зростає надій за 305 днів лактації, хоча цей взаємозв’язок характеризується також від’ємною кореляцією, але він має бажане селекційне та, особливо, економічне значення.

У генетиці під кореляцією розуміють взаємозв’язок окремих ознак організмів, зумовлений генетичними факторами в поєднанні з умовами зовнішнього середовища. Найчастіше визначають фенотипову кореляцію, яка дозволяє встановити зв’язок між різними господарськи корисними ознаками і

використати її з метою селекції. Обчислюють також генетичну кореляцію, для визначення впливу спадкових факторів на ступінь фенотипової кореляції.

Корелятивна мінливість суттєво впливає на онтогенетичну, комбінаційну і мутаційну мінливість. Її вплив також поширюється і на ступінь та характер не спадкової, модифікаційної мінливості.

**Модифікаційна мінливість** – це не спадкові зміни ознаки або властивості організму в онтогенезі, що викликані впливом зовнішніх умов. Формування кожного органу або ознаки і властивостей контролюються сукупністю генів, але проходить в конкретних умовах зовнішнього середовища, яке здійснює досить суттєвий вплив на ступінь і характер реалізації спадкової інформації.

Модифікаційна мінливість має велике практичне значення, оскільки ріст і розвиток, продуктивність, відтворна здатність тварин істотно залежать від умов годівлі і утримання. Тому створення відповідних умов для реалізації спадкової інформації даного генотипу особини (годовля та утримання згідно їх фізіологічного стану) – основа підвищення продуктивності тварин.

Модифікаційна мінливість не успадковується.

Різкі зміни в будові органів і прояві ознак прийнято називати **морфозами**. Причинами їх появи бувають порушення процесу органогенезу в ембріональний період онтогенезу. Морфози не успадковуються і, як правило, носять патологічний характер.

Розрізняють **індивідуальну** мінливість, що виявляється у межах однієї породи і дає можливість відрізнити одну тварину від іншої, а також **групову**, за якою відрізняють тварин різних груп, наприклад однієї породи від іншої. При розведенні тварин найважливішою є індивідуальна мінливість, яка створює можливості поліпшення стад і порід. Мінливість є тим матеріалом, з якого селекціонер, використовуючи різні методи селекції, створює бажаний тип сільськогосподарських тварин.

## **2.1. Значення мінливості в селекції тварин**

У процесі росту і розвитку під впливом селекції та умов життя сільськогосподарські тварини змінюються, причому кожна особина згідно зі своєю генетичною програмою, створюючи всю різноманітність форм і властивостей, притаманних виду, породі. Тому **мінливість двобічна**: з одного боку вона відображає процес становлення особини, а з другого – стан цієї особини в даний момент, ступінь її різноманітності.

Мінливість особин у межах популяцій зумовлена їхньою гетерозиготністю (генотиповими відмінностями). Під впливом селекції відбувається постійна зміна спадковості: одні генотипи, більш пристосовані, через ряд поколінь набувають переваги, а інші, менш пристосовані, порівняно скорочуються, в результаті чого за рахунок зміни складу (частот) генів відбувається перебудова генофонду.

Мінливість селекційних ознак має вирішальне значення для реалізації селекційних програм. Вона створює резерв добору і має важливе еволюційне значення – забезпечує пристосованість організмів до змінних умов середовища, тобто *адаптацію*.

При відсутності мінливості для селекціонера виникає складна задача. Якщо між особинами немає різниці, то немає й необхідності відбирати або вибраковувати тварин при формуванні племінних груп, бо з генетичної точки зору ці особини будуть однаковими або ж генетичні відмінності між ними мало помітні.

Основною рушійною силою удосконалення спадкових якостей тварин є добір, тобто виділення у межах популяції окремих груп тварин, які розрізняються мінливістю за своїми господарськи корисними і біологічними ознаками, та завдяки чому по різному використовуються у подальшій племінній роботі: найбільш цінні – для відтворення маточного складу, гірші – для одержання надремонтного молодняку і найгірші вибраковуються.

Ефективність добору (селекції) визначається величиною мінливості. При малій мінливості селекціонеру дуже важко знайти у стаді особин, які задовольняють поставленим вимогам, або виявити таку їх кількість, яка не забезпечує необхідних темпів відтворення стада. Дуже велика мінливість також небажана, оскільки, проявляючись в кожному наступному поколінні, вона призводить до великої величини регресії, тобто повернення до середніх показників популяції у потомства тварин, відібраних за тією чи іншою ознакою.

## **2.2. Математичні параметри мінливості**

Визначення середніх характеристик вибірових сукупностей дає можливість порівнювати між собою окремі групи тварин за господарськи корисними ознаками. Хоча середні величини є важливими статистичними показниками, але вони не дозволяють оцінювати гетерогенність (різноманітність) селекціонованих ознак у

популяції. Середні величини нічого не говорять про основну властивість живих організмів – мінливість, без обліку якої неможливо скласти повну характеристику вибіркової сукупності. Ознаки, які досліджуються, при однакових середніх величинах можуть істотно відрізнятись за рівнем варіації.

Тому, поряд із середньою величиною ознаки визначаються параметри її мінливості. Одним із показників мінливості ознак є *ліміти*, що з латинської означає – межі, які вказують на мінімальні ( $V_{min}$ ) та максимальні ( $V_{max}$ ) варіанти, між якими знаходяться усі члени даної вибіркової сукупності.

Різниця між лімітами ознаки дозволяє визначити наступний показник мінливості – *амплітуду* або *розмах варіювання* ( $R = V_{max} - V_{min}$ ). Більший розмах свідчить про вищу мінливість однієї із оцінюваних ознак.

Показники лімітів конкретні та прості, але вони здатні істотно змінюватись при повторних вибірках із однієї й тієї ж генеральної сукупності, тому їхнє використання – обмежене.

Головними показниками, які характеризують мінливість селекційних ознак, є *дисперсія*, *середньоквадратичне відхилення* та *коефіцієнт варіації*. Ці показники вказують на ступінь різноманітності варіант у вибірковій сукупності.

**Дисперсія та середньоквадратичне відхилення.** Основним показником мінливості служить середнє квадратичне відхилення ( $\sigma$ , іноді використовують букву  $S$ ). Середнє квадратичне відхилення одержують при вирахуванні кореня квадратного із дисперсії ( $\sigma^2$ ).

Варіанти ( $V$ ) різняться між собою за рівнем своєї вираженості, а щоб величину кожної з них можна було б порівняти із середньою всієї вибіркової сукупності ( $M$ ) необхідно визначити середню величину їхнього відхилення від середньої всієї вибірки. Цього можна досягти, якщо від кожної варіанти відняти величину середньої арифметичної ( $V - M$ ). При цьому варіаційний ряд буде мати як від'ємні, так і позитивні числа, які у своїй сумі –  $\Sigma(V - M)$  завжди будуть дорівнювати нулю.

Саме з метою уникнення мінусових величин усі відхилення підносяться до квадрату  $\Sigma(V - M)^2$ . Щоб визначити середню величину відхилень потрібно розділити отриману суму на загальну кількість варіант –  $n$ .

В такому випадку дисперсія (варіанса) завжди має квадратичне значення і вираховується за наступною формулою:

$$\sigma^2 = \frac{\Sigma(V - M^2)}{n} \quad (5)$$

Таким чином, дисперсія вимірює мінливість ознак відносно своєї середньої арифметичної і виражає її через середній квадрат відхилення кожного члена сукупності в середньому. Дисперсія широко використовується у науково-дослідних роботах для поглибленого аналізу ознаки за допомогою методики дисперсійного аналізу. В усіх інших випадках в основному використовується середнє квадратичне відхилення –  $\sigma$ , яке визначається за формулою вирахування кореня квадратного із дисперсії:

$$\sigma = \pm \sqrt{\frac{\Sigma(V - M)^2}{n}} \quad (6)$$

Для малої вибірки формула має такий вид:

$$\sigma = \pm \sqrt{\frac{\Sigma(V - M)^2}{n - 1}} \quad (7)$$

Середнє квадратичне відхилення ще має назву *стандартного відхилення* і є величиною поіменованою бо виражається у тих самих одиницях виміру, що й оцінювана ознака.

Ця величина ( $\sigma$ ) може мати додатне і від’ємне значення, що свідчить про мінливість ознаки як у бік зменшення варіантів від середньої, так і в бік збільшення.

За кількісним значенням сигми у відносному порівнянні до середньої величини вибірки ( $M$  чи  $\bar{X}$ ) можна зробити висновок про ступінь мінливості ознаки. Чим більша сигма, тим більша мінливість і навпаки.

Можлива ситуація, коли дві вибірки, які порівнюються мають однакові величини  $V_{max}$ ,  $V_{min}$  і  $M$ , але за особливостями варіювання і величинами  $\sigma$  – розрізняються, тобто спостерігається неоднаковий ступінь мінливості (варіабельності) даної ознаки.

Середнє квадратичне відхилення служить основним показником різноманітності ознаки у вибірковій сукупності. Використовується сигма як самостійний параметр мінливості, так і в якості базового показника для визначення

багатьох інших показників біометрії: коефіцієнта варіації, помилок репрезентативності, різних показників розподілу варіант, коефіцієнтів кореляції та регресії, елементів дисперсійного аналізу.

Середнє квадратичне відхилення показує на стільки в середньому кожна варіанта відхиляється від своєї середньої арифметичної. У так званих нормальних варіаційних рядах весь розмах мінливості обмежений мінімальним та максимальним значенням варіюючої ознаки і містить у собі шестикратну величину середнього квадратичного відхилення. При цьому максимальна варіанта віддалена від середньої арифметичної на  $+3\sigma$ , а мінімальна на  $-3\sigma$ . Тому весь розмах мінливості прийнято виражати як  $M \pm 3\sigma$ . На основі цієї властивості мінливості у нормальних варіаційних рядах цілком можливо передбачати окремі показники без додаткових експериментальних спостережень.

Середнє квадратичне відхилення – важливий показник для зоотехніка-селекціонера, оскільки саме він визначає ефективність добору в популяціях. Так, при невеликих значеннях середньоквадратичного відхилення всі особини будуть близькими до середньої арифметичної і тому важко відібрати кращих. У цьому випадку поліпшення даної ознаки буде відбуватися повільно. Але при високій різноманітності ознаки створюються умови для постійного підвищення продуктивності. Селекційна стратегія повинна будуватися на тому, що на початкових етапах створення нових ліній або порід необхідна висока мінливість ознаки, а на завершальній стадії слід прагнути до її різкого зменшення з метою консолідації одержаної групи.

Слід враховувати, що підвищення мінливості у високо відселекціонованих популяціях є досить складним завданням і тому вдаються до схрещування ліній, порід, окремих видів тварин.

**Коефіцієнт мінливості або варіації.** Серед констант варіативності ознак досить наглядно і більш біологічно обґрунтовано характеризує мінливість коефіцієнт варіації ( $C_v$ ), тобто середнє квадратичне відхилення, виражене у відсотках від середнього показника оцінюваної групи тварин за даною ознакою:

$$C_v = \frac{\sigma}{M} \times 100 \quad (8)$$

Оскільки середнє квадратичне відхилення величина поіменована, використовувати її для порівняння ознак, які виражаються у різних одиницях вимірювання, неможливо. Використання коефіцієнта варіації дозволяє порівняти

ступінь мінливості різних за одиницями виміру ознак, наприклад, величини надою у кг і вмісту жиру в молоці у %.

Одним із суттєвих недоліків коефіцієнта варіації є його залежність від величини середньої арифметичної і чим вона вища, тим цей показник нижчий.

Вважають, що ознаки із слабкою мінливістю характеризуються коефіцієнтом варіації на рівні меншому за 10 %, з середньою – 11–20 і при значній мінливості ознаки –  $C_v$  становить більше 21 %.

### 2.3. Типи розподілу ознак. Нормальний розподіл

Більшості кількісних ознак сільськогосподарських тварин властивий нормальний тип розподілу. При ньому найбільше число варіант розміщено в центрі близько середніх значень ознаки. Чим більше відхиляються значення окремих варіант, тим рідше вони зустрічаються, тобто ймовірність повторюваності варіанти зменшується при віддаленні її від середніх величин. Тому в конкретній популяції особин з середніми значеннями продуктивності завжди більше, ніж з мінімальними й максимальними. Нормальний розподіл повністю характеризується середньою величиною і стандартним відхиленням.

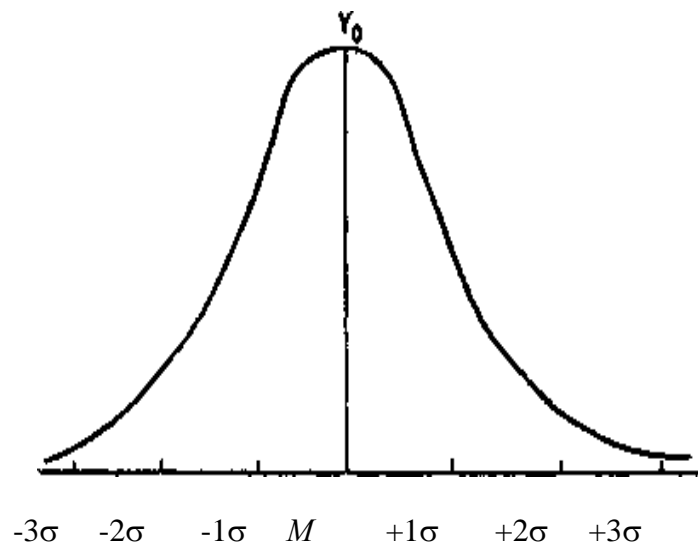
Існує правило трьох сигм або нормального розподілу ознак, яке ґрунтується на кривій Гауса (німецького вченого), який разом з Лапласом (французьким вченим) відкрили це правило.

Для нормального розподілу при  $n \rightarrow \infty$  характерним є наступне:

- теоретична крива має дзвіноподібний симетричний вигляд;
- фундаментом кривої служать класи варіюючої ознаки ( $x$ );
- гілки дзвіноподібної кривої утворені перпендикулярами (ординатами), відновленими з точок значень  $x$ ;
- кінці гілок не зливаються з віссю абсцис, а наближаються до неї в нескінченності (асимптотично).
- вершина нормальної кривої ( $y_0$ ) визначається перпендикуляром, спрямованим з точки  $x$ , яка відповідає величині середній арифметичній даної ознаки.
- $y_0$  відповідає найбільшій частоті зустрічі ( $p$ ) тих особин популяції, у яких величина ознаки дорівнює середній арифметичній;
- у нормальному розподілі точка  $x$  співпадає з величиною моди ( $Mo$ ) і медіани ( $Me$ );



- чим більша мінливість ознаки та відповідно значення сигми, тим більшою буде основа у кривій і тим нижче буде розміщуватися вершина.



**Рис. 1. Крива нормального розподілу частот (Гауса–Лапласа)**

Вся різноманітність ознаки як правило входить у розмах  $M \pm 3\sigma$ , тобто в інтервалі від самого мінімального до самого максимального значення входить шість сигм. При цьому в межах  $M \pm 1\sigma$  знаходиться 68,3% особин, у межах  $M \pm 2\sigma$  – 95,5 та  $M \pm 3\sigma$  – 99,7% (правило трьох сигм).

Виходячи з цього, можна так охарактеризувати значення селекційної ознаки – величина надою молочних корів племінного стада, коли  $M = 6550$  кг, а  $\sigma = \pm 600$  кг. Мінімальне значення варіанти у даній вибірці становить  $M - 3\sigma$ , тобто  $6550 - (3 \times 600) = 4750$  кг, а максимальне  $M + 3\sigma$ , тобто  $6550 + (3 \times 600) = 8350$  кг. Виходячи з цього, можна приблизно вирахувати середнє квадратичне відхилення, поділивши розмах мінливості на шість. Тобто сигма приблизно буде дорівнювати  $M_{\max} - M_{\min} / 6$ , або у нашому прикладі:  $8350 - 4750 = 3600 / 6 = 600$  кг.

Слід враховувати, що в межах  $M \pm 0,67\sigma$  знаходиться 50 % особин з усієї популяції. Це так звані модальні, типові особини. За роботами Г.Я.Копиловської та І.В. Хорунжого (1966), зазначені особини відрізняються високою пристосованістю (адаптивною нормою) до конкретних факторів середовища, що проявляється в їх вищих відтворних якостях і збереженості.

Відбір особин модальних класів виявився ефективним засобом консолідації ліній, а також збереження генофонду. Емпіричний розподіл частот варіаційного ряду за найбільш важливими ознаками (жива маса, середньодобові прирости, молочна, м'ясна, вовнова, яєчна продуктивність) при

вивченні тієї чи іншої популяції рекомендується зображати графічно, відкладаючи на осі абсцис класи ознаки, а на осі ординат – його частоти. При цьому з великою точністю будуть виявлені такі явища, як зменшення або збільшення мінливості ознаки, яку визначають упродовж ряду поколінь. Знаючи параметри нормального розподілу ( $M$  та  $\sigma$ ), можна визначити частку особин у групі або стаді з певним рівнем продуктивності, чи вирішити зворотну задачу – за часткою особин, які не збігаються із середніми значеннями, вирахувати їх середню продуктивність або межу продуктивності, з якої необхідно вести добір чи вибракування тварин.

Використовуючи ЕОМ такий аналіз можна проводити автоматично для однієї або кількох ознак.

У селекції тварин нормоване відхилення є показником інтенсивності добору. Відбираючи особин, які переважають середні значення на  $1\sigma$ , можна залишити для розведення 15,85 % голів від загальної кількості оцінених тварин, а при  $1,5\sigma$  й вище таких особин буде тільки 6,68 %.

Чим вища інтенсивність добору, тим менше тварин залишають для відтворення наступного покоління. Враховуючи, що в свинарстві для забезпечення постійного відтворення стада необхідна інтенсивність добору 20–25 %, реальна величина нормованого відхилення відібраних свинок буде на рівні  $0,5\sigma$  і вище, для кнурів –  $1,0$ – $1,5\sigma$  і вище.

На практиці доводиться керуватися міркуваннями збереження розширеного відтворення і визначати межу добору при заданій мінімальній чисельності особин. Використання нормованого відхилення дає можливість визначити важливі для селекціонера параметри добору – нижню межу ознаки в популяції, та середнє відібраної групи. Це дозволяє планувати селекційну роботу, виявити оптимальну частку особин, яких добирають.

Експериментальне вивчення мінливості біологічних об'єктів показало, що в більшості випадків ми маємо їх нормальний розподіл. Збіг абсолютних величин середньої арифметичної ( $M$ ), моди ( $Mo$ ) та медіани ( $Me$ ) свідчить про нормальний розподіл варіаційного ряду.

### 3. РЕПРЕЗЕНТАТИВНІСТЬ ТА ПОХИБКИ ПОКАЗНИКІВ ВИБІРКОВИХ СУКУПНОСТЕЙ

Численні дослідження показали, що правильно сформована вибірка сукупність вірно відображає або *репрезентує* структуру, стан і властивості генеральної сукупності.

Існує загальна система добору об'єктів у вибірку сукупність. Головним для цієї системи є правило відповідної ймовірності вибору якого завгодно об'єкта із генеральної сукупності.

**Репрезентативність** – головна властивість вибірових сукупностей характеризувати генеральну сукупність з відповідною точністю та достатньою надійністю. Репрезентативністю позначають і ступінь відповідності вибірових показників генеральним параметрам.

Репрезентативність з'ясовується лише тоді, коли необхідно охарактеризувати усю велику сукупність особин, що цікавить дослідника, на основі вивчення лише якоїсь її частини і коли уже відібрана відповідна вибірка. При цьому можливі певні відхилення вибірових показників від параметрів генеральної сукупності. Такі відхилення в біометрії називаються *похибками репрезентативності*.

Похибки репрезентативності виникають тому, що вибірка сукупність не точно відповідає своїй генеральній сукупності за тими чи іншими ознаками. Ці похибки слід відрізняти від інших, які мають організаційний характер, оскільки при проведенні досліджень завжди існує небезпека створення похибок різного походження. Тому всі можливі похибки можна залежно від походження відповідно класифікувати.

**Методичні похибки:** виникають внаслідок методичних прорахунків при організації експериментів, неправильної організації спостережень, створенні різних умов для контрольної та дослідних груп;

**Систематичні похибки (точності):** пов'язані з використанням недосконалих, не стандартизованих та несправних приладів, проведення розрахунків з невідповідною точністю, неправильне використання алгоритмів;

**Випадкові (суб'єктивні) похибки:** описки, прорахунки, переплутування дослідних зразків, можуть виникати внаслідок низької кваліфікації і несумлінності виконавців та через незадовільних умов для їхньої роботи.

**Похибки вибіркового характеру:** добір особин, які неправильно характеризують генеральну сукупність; добір особин, що існували в умовах не властивих тваринам генеральної сукупності; добір без урахування пропорційності.

Для усунення помилок необхідно використовувати сучасні методики обліку й вимірювання, регулярно проводити перевірку приладів і враховувати поправку на виміри, вказану в паспортах.

Похибки репрезентативності можуть бути виявлені та вираховані спеціальними біометричними методами, але вони не можуть бути ліквідованими навіть при ідеальній організації експериментальних досліджень, оскільки похибки завжди ймовірні. Величина похибки значною мірою залежить від об'єму вибірки та способу добору об'єктів із генеральної сукупності.

Біометричні методи обліку похибок дозволяють визначити довірчі межі генеральних параметрів та достовірність різниці між вибірковими сукупностями.

### 3.1. Похибки біометричних показників

**Похибка середньої арифметичної.** Вибірка, мала чи велика, є лише часткою генеральної сукупності, тому вона не може з абсолютною точністю відобразити її параметри та властивості, тому усі статистичні показники будуть у тій чи іншій мірі відрізнятися від параметрів генеральної сукупності.

Стандартну статистичну похибку позначають буквою  $m$ , а під нею, як нижній індекс, записують знак того біометричного параметру, до якого вона належить.

Похибка середньої арифметичної для великої вибірки визначається за формулою:

$$m_M = \frac{\sigma}{\sqrt{n}} \quad (9)$$

Для малої вибірки формула наступна:

$$m_M = \frac{\sigma}{\sqrt{n-1}} \quad (10)$$

Похибка вибірки виражається у тих самих одиницях, що й оцінювані показники, які вона характеризує. Похибка має два об'єднані знаки – плюс і мінус ( $\pm$ ), які вказують на те, що відхилення вибірових показників можуть відбуватися як у бік більших, так і в бік менших значень по відношенню до параметра генеральної сукупності. Показник середньої арифметичної разом зі своєю похибкою записується таким чином:  $M \pm m_M$ , але з метою спрощення, як правило пишуть так:  $M \pm m$ .

Величина похибки залежить не лише від об'єму вибірки, але й від розмаху мінливості ознаки. Чим більша мінливість ознаки, тим більша величина похибки і навпаки.

*Похибка середнього квадратичного відхилення* обчислюється за наступною формулою:

$$m_{\sigma} = \frac{\sigma}{\sqrt{2n}} \quad (11)$$

*Похибка коефіцієнта варіації* вираховується за формулою:

$$m_{Cv} = \frac{Cv}{\sqrt{2n}} \quad (12)$$

Формули похибок інших біометричних та популяційних показників будуть наводитися у відповідних розділах.

### 3.2. Оцінка достовірності статистичних величин

В усіх випадках спостережень, експериментальних досліджень параметри генеральної сукупності невідомі, тому їх характеризують на основі аналізу вибірових сукупностей. Оскільки показники вибірки випадкові, тому й отримані результати вважаються гіпотетичними, або з певною часткою припущення. Для оцінки величини параметрів генеральної сукупності за даними вибірки у біології існує *нульова гіпотеза*. Вона ґрунтується на припущенні, що між показниками двох вибірових сукупностей достовірної різниці не існує, тобто припускається, що оцінювані групи складають однорідний матеріал, тобто, одну сукупність.

Виходячи із цього, математичний аналіз повинен підтвердити існування нульової гіпотези, або, навпаки, відхилити її та довести, що між показниками дійсно існує достовірна різниця. Якщо в процесі порівняльного аналізу результатів досліджень біометричні методи доводять достовірність різниці на користь експериментального варіанту, в такому разі розробку можна впроваджувати у виробництво. У разі, коли достовірність різниці не доведена, вона вважається випадковою і в дії залишається нульова гіпотеза, тому ніяких впроваджень результатів досліджень робити не можна.

Відхилення нульової гіпотези повинно бути обґрунтоване відповідним рівнем значущості. У біологічних дослідженнях приймають щонайменше 5%-й рівень, якому відповідає ймовірність  $P=0,95$ . У більш відповідальних дослідженнях, коли висновки повинні бути більш точними і надійними приймається величина значущості в 1,0% або 0,01%, яким відповідає рівень ймовірності – 0,99 та 0,999.

Найчастіше у науковій практиці використовується три пороги достовірної (довірчої) ймовірності, кожний із яких пов'язаний з відповідною величиною нормованого відхилення:

$P = 0,95$  відповідає  $-x = 1,96 \sigma$ ;

$P = 0,99$  відповідає  $-x = 2,58 \sigma$ ;

$P = 0,999$  відповідає  $-x = 3,29 \sigma$ .

Таким чином, завдяки рівню ймовірності, одержані характеристики вибірки переносяться на аналогічні показники генеральної сукупності з відповідним порогом достовірності. Наприклад ймовірність на рівні 0,95 ( $P \geq 0,95$ , більше або дорівнює) означає, що близькі величини до одержаних у досліді результативних показників повинні повторитися в 95 вибірках із 100.

### 3.3. Оцінка достовірності різниці вибірових середніх

Оскільки усе пізнається у порівнянні, у дослідній та практичній роботі велике значення має визначення достовірності отриманої різниці у порівнянні середніх показників ознак двох вибірових сукупностей.

Для визначення достовірності різниці між середніми значеннями окремих господарськи корисних ознак, які характеризують породу, селекційні групи, лінії тварин та кроси птиці визначають критерій вірогідності  $t_d$ .

При цьому виходять з двох передбачень:

1. Вибірки взяті з різних генеральних сукупностей і різниця між ними відображає характерні відмінності генеральних середніх.

2. Вибірки взяті з однієї генеральної сукупності і різниця між ними виникає у результаті похибки або дії випадкових факторів.

Рівень достовірності різниці між середніми значеннями двох вибірок обумовлюється трьома факторами:

**об'ємом вибірки**, тобто число особин вибіркової сукупності визначає достовірність різниці. Чим більша вибірка, тим більше вибіркові показники наближаються до генеральної сукупності, тим менша похибка репрезентативності, а значить, надійніша достовірність різниці;

**мінливістю ознак** – чим більша різноманітність вибіркових показників, тим менша достовірність різниці за однакових величин об'єму вибірки та самої різниці;

**величиною різниці** – чим більша різниця між середніми показниками, тим вона достовірніша за однаковими своїми об'ємом та мінливістю.

При вираховуванні достовірності різниці застосовують наступну формулу:

$$t_d = \frac{M_1 - M_2}{\sqrt{m^2_{M_1} + m^2_{M_2}}} \quad (13)$$

де:  $m^2_{M_1}$  та  $m^2_{M_2}$  – похибки середніх арифметичних ознак першої та другої груп тварин, що порівнюються.

Для вирішення питання про суттєвість або не суттєвість відмінностей між групами порівнюють значення  $t_d$  (критерій Стюдента) із стандартними при такому числі ступенів свободи  $v = n_1 + n_2 - 2$ ; де:  $v$  – ступінь свободи;  $n_1$  і  $n_2$  – кількість спостережень (дат) у групах. Рівень достовірності одержаних результатів позначається наступним чином:  $P > 0,95$ ;  $P > 0,99$  або  $P > 0,999$ , або  $P < 0,05$ ;  $P < 0,01$ ;  $P < 0,001$ ; або  $B = 0,95$ ;  $B = 0,99$  та  $B = 0,999$ .

Визначене у дослідженнях фактичне значення  $t_d$  порівнюємо з теоретично очікуваним за таблицею Стюдента і якщо воно дорівнює або вище, то говоримо про достовірність даних при певному порозі ймовірності.

## Приклади визначення біометричних показників

### 1. Розрахунок $M$ , $\sigma$ , $C_v$ , $m_M$ , $m_\sigma$ , $m_{C_v}$ , $t_M$ методом малих вибірок.

Для обчислення складається простий варіаційний ряд, наприклад, виписуються величини промірів або живої маси кожної тварини молодняку телиць при оцінці їхнього росту і розвитку. Варіанти ( $V$ ) розташовують у порядку їх збільшення або зменшення. Наприклад, потрібно обчислити середню арифметичну ( $M$ ) за живою масою 1000 телиць української червоно-рябої молочної породи. Для цього рендомізовано беруть живу масу 10 телиць, складають простий варіаційний ряд – виписують живу масу кожної особини (це варіанти  $V$ ) у порядку зростання, потім складають варіанти і ділять на кількість ( $n$ ) варіант, тобто  $M = \frac{\sum V}{n}$ .

$V$  – 350; 340; 365; 375; 380; 395; 397; 405; 350; 340.

1. Складаємо простий варіаційний ряд:

340; 340; 350; 350; 365; 375; 380; 395; 397; 405.

2. Додаємо варіанти і ділимо на кількість ( $n$ ) варіант:

$$M = \frac{\sum V}{n}$$

$$M = \frac{340 + 340 + 350 + 350 + 365 + 375 + 380 + 395 + 397 + 405}{n} = 369,7 \text{ кг.}$$

У даному випадку середня арифметична  $M=369,7$  кг показує, скільки б важила кожна первістка (з 1000 голів), якби не було мінливості.

Як відомо, усі біологічні ознаки змінюються. Тому потрібно обчислити середнє квадратичне відхилення  $\sigma$ , яке вказує на ступінь мінливості, тобто на яку величину в середньому кожна варіанта може відхилятися від середньої арифметичної у бік збільшення або зменшення. Для цього спочатку записують варіанти  $V$ , потім відхилення кожної варіанти від середньої арифметичної ( $V-M$ ), підносять ці відхилення до квадрату  $(V-M)^2$ , обчислюють суму квадратів відхилень  $\sum (V-M)^2$  і за формулою обчислюють  $\sigma$ , яка виражається в натуральних величинах:

1. Визначаємо середнє квадратичне відхилення  $\sigma$

$$\sigma = \pm \sqrt{\frac{\sum (V-M)^2}{n-1}}$$



Таблиця 1

№ тварини	Жива маса телиць (V)	Відхилення від середньої (V-M)	Квадрати відхилень (V-M) <sup>2</sup>
1	340	-29,7	882,09
2	340	-29,7	882,09
3	350	-19,7	388,09
4	350	-19,7	388,09
5	365	-4,7	22,09
6	375	5,3	28,09
7	380	10,3	106,09
8	395	25,3	640,09
9	397	27,3	745,29
10	405	35,3	1246,09
n = 10	M = 369,7	0	$\Sigma(V - M)^2 = 5328,1$

$$\sigma = \sqrt{\frac{5328,1}{10-1}} = \sqrt{592,01} = \pm 24,33 \text{ кг}$$

Отже, кожна варіанта, тобто жива маса кожної телиці, в середньому на 24,33 кг відхиляється від середньої арифметичної (M=369,7 кг) у бік збільшення та зменшення. Це вказує на різномірність поголів'я даної генеральної сукупності у генетичному аспекті.

Відомо, що все пізнається у порівнянні, в тому числі й ступінь мінливості. Проте порівнювати ступінь мінливості різних ознак, наприклад добового надою, вмісту жиру в молоці, температури тіла тварин тощо за середніми квадратичними відхиленнями ( $\sigma$ ) неможливо, оскільки вони в кожному конкретному випадку будуть виражатись у різних одиницях вимірювання (кг, %, °C) та різних розрядах чисел. Тому, щоб порівняти рівень мінливості ознак оцінених у різних величинах виміру обчислюють коефіцієнт мінливості  $C_v$ , виражений у відсотках, який дозволяє проводити такі порівняння:

Обчислюємо коефіцієнт мінливості  $C_v$ , %.

$$C_v = \frac{\sigma \times 100}{M}$$

$$C_v = \frac{\sigma \times 100}{M} = \frac{24,33 \times 100}{369,7} = \frac{2433}{369,7} = 6,58\%$$

Вирахувавши середню арифметичну, яка дорівнює  $M=369,7$  кг, середньоквадратичне відхилення –  $\sigma=\pm 24,33$  кг та коефіцієнт мінливості –  $C_v=6,58\%$  за живою масою 10 телиць (мала вибірка), ці показники переносяться на всю генеральну сукупність. При цьому виникають помилки узагальнення, оскільки частина не може точно характеризувати ціле. Помилки узагальнення, пов'язані з перенесенням даних ( $M$ ,  $\sigma$ ,  $C_v$  тощо), одержаних за допомогою вибірки на всю генеральну сукупність, називаються репрезентативними помилками.

Репрезентативні помилки обчислюють за такими формулами:

помилка середньої арифметичної

$$m_M = \frac{\sigma}{\sqrt{n}} = \frac{24,33}{3,16} = 7,699 \text{ кг}$$

помилка середнього квадратичного відхилення

$$m_\sigma = \frac{\sigma}{\sqrt{2n}} = \frac{24,33}{\sqrt{2 \times 10}} = \frac{24,33}{4,472} = 5,44 \text{ кг}$$

помилка коефіцієнта мінливості

$$m_{C_v} = \frac{C_v}{\sqrt{2n}} = \frac{6,58}{4,47} = 1,472\%$$

Одним із заключних етапів усіх підрахунків є запис даних:

$$M \pm m = 369,7 \pm 7,699 \text{ кг}$$

$$\sigma \pm m = 24,33 \pm 5,44 \text{ кг}$$

$$Cv \pm m = 6,58 \pm 1,472 (\%)$$

Далі ці помилки використовують для обчислення критерію вірогідності (достовірності)  $t$  цих показників:

$$t_M = \frac{M}{m_M} \quad (14)$$

$$t_M = \frac{369,7}{7,699} = 48,02 \quad (P > 0,999)$$

За величиною  $t_M$  судять про вірогідність, наприклад, середньої арифметичної, виходячи із зв'язку  $t_M$  з рівнем ймовірності.

Існує три рівні вірогідних меж ймовірності  $P$ :  $P=0,95$ ;  $P=0,99$ ;  $P=0,999$ . Ймовірність  $P=0,95$  означає, що в 95 випадках із 100 показники збігатимуться,  $P=0,99$  – в 99 випадках із 100,  $P=0,999$  – в 999 випадках із 1000.

Для цих вірогідних рівнів Стьюдентом обчислені стандартні значення критерію вірогідності у межах конкретних ступенів свободи, з якими і треба порівнювати визначений у кожному конкретному випадку критерій вірогідності. Ступінь свободи обчислюють за формулою  $\nu = n - 2$ , тобто  $\nu = 10 - 2 = 8$ . Потім у таблиці Стьюдента стандартних значень знаходять у першому стовпчику 8 і порівнюють значення  $t_M = 48$  із стандартними 2,3; 3,4; 5,0. В даному випадку критерій вірогідності середньої арифметичної більше, ніж стандартні значення 2,3; 3,4; 5,0. Це означає, що середня арифметична вірогідна при  $P > 0,999$ , тобто такою середньою арифметичною можна користуватись, характеризуючи генеральну сукупність.

Обчислення критерію вірогідності  $t$  цих показників.

Р – критерій вірогідності		
0,95	0,99	0,999
$v=n-2$ , тобто $v=10-2=8$		
2,3	3,4	5,0

*Якщо критерій вірогідності менше за стандартні значення, то користуватись такою середньою арифметичною не можна. Для цього потрібно взяти нову вибірку й обчислити її знову.*

## 2. Розрахунок статистичних показників $M$ , $\sigma$ , $C_v$ , $m_M$ , $m_\sigma$ , $m_{C_v}$ , $t_M$ методом великих вибірок

Якщо маємо справу з великою кількістю тварин ( $n > 30$ ), тобто з великою вибіркою, обробку цифрових даних проводять не прямим способом, а за допомогою складніших специфічних методів вирахування середньої арифметичної. Одним із таких найчастіше вживаних є **метод умовної середньої**.

В основі цього методу закладені математичні операції, у яких за основу узято модальний клас та відхилення від нього. При цьому модальний клас використовується як умовна середня. Остання здебільшого добирається за найбільшою частотою варіант. У тих випадках, коли модальних класів декілька або він знаходиться не посередині варіаційного ряду, за умовну середню добирають той клас, котрий розділяє варіаційний ряд на більш-менш рівні частини за своїми частотами. На конкретному прикладі розглянемо порядок визначення середньої арифметичної.

**Приклад виконання типового завдання.** Визначити середню величину живої маси корів української бурої молочної породи за наступними даними (кг): 646, 555, 692, 635, 610, 614, 597, 615, 610, 611, 585, 652, 654, 569, 669, 500, 665, 650, 673, 617, 592, 578, 672, 550, 605, 625, 645, 545, 552, 629, 598, 618, 621, 678, 540, 715, 568, 688, 612, 570, 685, 602, 670, 622, 596, 675, 618, 547, 638, 655, 530, 644, 599, 584, 657, 559, 580, 627, 567, 630, 590, 562, 660, 591, 628, 630, 586, 580, 635, 610, 567, 619, 576, 571, 538, 595, 655, 625, 522, 612, 636, 604, 625, 633, 630, 653, 708, 664, 572, 618, 634, 596, 612, 603, 644, 608, 576, 564, 535, 615, 652, 615, 637, 587, 601, 626, 565, 625, 630, 648, 639, 585, 612, 583, 604, 594, 590, 635, 617, 522, 574, 612, 620, 605, 573, 614, 647, 607, 584, 565, 695, 582, 614, 636, 658, 595, 606, 615, 588, 642, 604, 618, 603, 621, 647.

Для складання й опрацювання такого варіаційного ряду будемо класи згідно з табл. 3.

**Варіаційний ряд живої маси корів**

Класи ( $w_0 - w_t$ )	Частота варіант (P)	Умовне відхилення від модального класу (a)	Добуток $P a$	Добуток $P a^2$
500-519	1	- 5	- 5	25
520-539	5	- 4	- 20	80
540-559	8	- 3	- 24	72
560-579	14	- 2	- 28	56
580-599	26	- 1	- 26	26
<b>600-619</b>	<b>35</b>	<b>0</b>	<b>0</b>	<b>0</b>
620-639	25	+ 1	+ 25	25
640-659	16	+ 2	+ 32	64
660-679	9	+ 3	+ 27	81
680-699	4	+ 4	+ 16	64
700-719	2	+ 5	+ 10	50
$k = 20$	$\Sigma P = 145$		$\Sigma P a = 7$	$\Sigma P a^2 = 543$

1. Визначаємо ліміти – найбільше і найменше значення даної ознаки:

найменше значення – 500 кг;

найбільше значення – 715 кг.

2. Визначаємо розмах варіювання: від максимального значення у вибірці віднімаємо мінімальне:  $715 - 500 = 215$  кг.

3. Визначаємо кількість класів для побудови майбутнього варіаційного ряду. Класи – це градації, в які згруповані близькі за значенням варіанти. Кількість класів у кожному випадку береться довільно, але в межах від 5 до 15.

У даному разі взято 11 класів.

4. Визначаємо класовий проміжок (k) діленням величини розмаху

варіювання на кількість взятих класів.

$$k = \frac{215}{10} = 21,5 \approx 20,0 \text{ кг (краще заокруглити до цілого числа).}$$

5. Додаємо класовий проміжок до найменшого ліміту ( $500+20=520$ ), а потім до нових значень додають заплановану кількість класів і щоб в останній увійшла максимальна варіанта.

6. Знаходимо частоти  $P$ , тобто кількість корів, яких залежно від живої маси розміщують у відповідні класи.

7. Визначаємо умовний середній клас  $A$ . Виділяємо його жирними лініями, позначаємо нулем. Величина умовної середньої розраховується за формулою:

$$A = \frac{w_0 + w_t}{2} \quad (15)$$

$$A = \frac{600 + 619}{2} = 609,5$$

8. Знаходимо умовне відхилення ( $a$ ) від модального класу, що позначили нулем, усіх інших класів варіаційного ряду у цілих числах. Відхилення, що мають менше значення за умовну середню величину будуть із знаком мінус з відхиленням у бік зменшення на п'ять класових проміжків, а ті що мають більше значення – будуть із знаком плюс з відхиленням у бік зростання класів.

9. Знаходимо добуток частот на відхилення  $Pa$  та їх суму:  $\Sigma Pa = 7$

10. Визначаємо добуток частот на квадрат відхилень  $Pa^2$  та їх суму:

$$\Sigma P a^2 = 543$$

На цьому обробка варіаційного ряду закінчується. Підставляючи в формули дані з варіаційного ряду розрахуємо необхідні показники.

а) Вираховуємо середнє арифметичне значення за формулою:

$$M = A + kb \quad (16)$$

де:  $A$  – середнє значення умовного середнього класу;

$k$  – класовий проміжок;

$b$  – поправка, яка визначається за формулою:

$$b = \frac{\Sigma Pa}{n} \quad (17)$$

$$b = \frac{7}{145} = 0,04$$

Звідси:  $M = 609,5 + 20 \cdot 0,04 = 609,5 + 0,8 = 610,3$  кг.

б) Вираховуємо середнє квадратичне відхилення:

$$\sigma = \pm k \sqrt{\frac{\Sigma Pa^2}{n}} = 20 \sqrt{\frac{543}{145}} = 20 \cdot \sqrt{3,74} = 20 \cdot 1,93 = \pm 38,6 \text{ кг}$$

в) коефіцієнт варіації:

$$C_v = \frac{\sigma \cdot 100}{M} = \frac{38,6 \cdot 100}{610,3} = 6,32 \%$$

г) похибки середньої арифметичної:

$$m_m = \frac{\sigma}{\sqrt{n}} = \frac{38,6}{\sqrt{145}} = \frac{38,6}{12,04} = 3,21 \text{ кг}$$

д) похибку середньоквадратичного відхилення:

$$m_\sigma = \frac{\sigma}{\sqrt{2n}} = \frac{38,6}{\sqrt{145 \cdot 2}} = \frac{38,6}{17,03} = 2,27 \text{ кг}$$

ж) похибки коефіцієнта варіації:

$$m_{Cv} = \frac{Cv}{\sqrt{2n}} = \frac{6,32}{\sqrt{2 \cdot 145}} = \frac{6,32}{17,03} = 0,37\%$$

з) критерій достовірності середнього арифметичного значення:

$$t_m = \frac{M}{m} = \frac{610,3}{3,21} = 190,1$$

11. Остаточною є така форма запису статистичних даних:

$$M \pm m_M = 610,3 \text{ кг} \pm 3,21 \text{ кг}$$

$$\sigma \pm m_\sigma = 38,6 \text{ кг} \pm 2,27 \text{ кг}$$

$$Cv \pm m_{Cv} = 6,32 \% \pm 0,37 \%$$

12. Знаходимо ймовірність середньоарифметичного значення. При великій вибірці ступені свободи, практично, не мають значення, бо розподіл Стьюдента наближається до нормального. Тому кількість ступенів свободи буде  $\nu=n=145$ . Тому отримані фактично значення порівнюємо зі стандартними значеннями, які в таблиці (див. додаток 1) розміщені у рядку навпроти числа 143.

$P = 0,95$	$P = 0,99$	$P = 0,999$
1,96	2,58	3,29



Якщо порівняти фактичний критерій вірогідності ( $t_m = 190,1$ ) із найвищим стандартним значенням  $t = 3,29$ , що знаходяться у вірогідних межах  $P=0,999$ , помітно, що він істотно вище. Це означає, що середня арифметична вибірки вірогідно відображає рівень середньої арифметичної за живою масою корів генеральної сукупності.

### ***Задачі для самостійного розв'язування***

- 1.1. Розрахуйте статистичні показники  $M$ ,  $\sigma$ ,  $C_v$ ,  $m_M$ ,  $m_\sigma$ ,  $m_{C_v}$ ,  $t_M$ ,  $P$  методом малих вибірок користуючись додатком 2 (кожен студент отримує завдання згідно свого номеру в журналі).
- 1.2. Розрахуйте статистичні показники  $M$ ,  $\sigma$ ,  $C_v$ ,  $m_M$ ,  $m_\sigma$ ,  $m_{C_v}$ ,  $t_M$ ,  $P$  методом великих вибірок користуючись додатком 3.

### **3.4. Оцінка достовірності різниці вибірових середніх величин**

Оскільки усе пізнається у порівнянні, у науково-дослідній та практичній роботі велике значення має визначення достовірності отриманої різниці двох вибірових сукупностей. Наприклад, в експериментах при визначенні впливу якогось чинника на розвиток будь якої ознаки у тварин дослідної групи в порівнянні з контрольною групою завжди існує необхідність встановлення достовірності впливу досліджуваного чинника на основі отриманих середніх арифметичних. При порівнянні різниці між середніми показниками розвитку селекційних ознак підконтрольних груп тварин постає питання чи можна отриману різницю вважати закономірністю, а не випадковістю. На основі встановленої достовірності можна гарантовано стверджувати про корисний ефект застосованих методів та вмотивовано впроваджувати їх у виробництво.

Достовірність різниці обумовлюється трьома факторами: об'ємом вибірки, різноманітністю (рівнем мінливості) ознак та величиною показника різниці.

Об'єм вибірки, тобто число особин, які узяті до дослідження визначає достовірність різниці головним чином. Чим більше вивчено особин у вибірці, тим більше вибірові показники наближаються до генеральних параметрів, тим менша похибка репрезентативності, а значить, надійніша достовірність різниці.

Різнманітність величини ознаки призводить до більшої ймовірності отримати вибірові показники, що значно відрізнятимуться від параметрів генеральної сукупності і тим самим знижуватимуть достовірність різниці за однакових величин об'єму вибірки та самої різниці.

Чим більша величина різниці, тим вона достовірніша за умов однакових об'ємів та різноманітності ознаки.

При вирахуванні достовірності різниці використовують наступну формулу:

$$t_d = \frac{M_1 - M_2}{\sqrt{m_1^2 + m_2^2}}$$

**Приклад:** визначити достовірність різниці між середніми величинами надою дочок бугаїв-плідників двох ліній української червоно-рябої молочної породи. Продуктивність корів-первісток лінії Айвенго (n=155) становила 5523 кг молока за лактацію з похибкою  $\pm 58,6$  кг, а лінії Валіанта, відповідно – (n=135), 4855 та  $\pm 65,8$  кг.

Визначаємо достовірність за формулою:

$$t_d = \frac{5523 - 4855}{\sqrt{58,6^2 + 65,8^2}} = \frac{668}{\sqrt{3226,2 + 4329,6}} = \frac{668}{86,9} = 7,69$$

Цей критерій достовірності різниці досліджуваних середніх величин значно вищий за максимальне стандартне табличне значення Стьюдента (3,29), тому різниця надою між нащадками бугаїв-плідників двох ліній статистично вірогідна при  $P > 0,999$  на користь корів лінії Айвенго і залежить від спадковості плідників цієї лінії.

### **3.5. Визначення статистичних параметрів для якісних (альтернативних) ознак**

У практиці селекційної роботи виникає необхідність визначення параметрів для якісних ознак. Досить часто потребує визначення різниці між популяціями за життєздатністю, збереженістю, резистентністю й відтворними якостями. На відміну від кількісних, для якісних ознак середнє арифметичне визначається як частка або відсоток особин, що мають дану ознаку. За одержаними даними також визначають достовірність різниці між вибірковими частками.

## 4. КОРЕЛЯЦІЙНИЙ АНАЛІЗ

### (методи визначення сполученої мінливості ознак)

З метою підвищення ефективності селекції одночасно за кількома ознаками рекомендується насамперед враховувати їхню взаємну зумовленість, тобто кореляцію. Для цього виникла необхідність математичного усвідомлення такого явища та вираження його числовим показником, коефіцієнтом, за величиною якого можна було б говорити про тісноту чи силу зв'язку між окремими ознаками.

Для цього існує **коефіцієнт кореляції** – математичний вираз, за допомогою якого визначають ступінь зв'язку між окремими ознаками. Позначається коефіцієнт кореляції латинською буквою **r**.

Термін “кореляція” (від лат. *Correlatio* – співвідношення, зв'язок) вперше застосував Ж. Кюв'є ще у 1806 році. Запровадження у науку цього важливого математичного методу визвано нагальними потребами морфології, генетики, селекції та інших дисциплін.

**Кореляція** у генетиці – це зв'язок між середніми показниками окремих ознак організмів вибіркових сукупностей зумовлений генетичними факторами в поєднанні з умовами навколишнього середовища. Сам метод оцінки тісноти чи ступеня зв'язку носить назву **кореляційного аналізу**.

Кореляція між двома ознаками має назву *прості*, а коли до уваги береться більше оцінюваних ознак то говорять про множинну кореляцію та, відповідно, *множинний* коефіцієнт кореляції.

Існує також **функціональний зв'язок**, як правило – в неживій природі. За законами функціональної залежності, зміні одного показника на певну величину відповідає тільки одна конкретна величина іншого показника (наприклад, зменшення стовпчика ртуті в термометрах при зниженні температури).

Функціональна залежність (альтернативна) існує і в живій природі (заплідненість – народження). При кореляційній залежності між двома ознаками вони по відношенню одна до одної можуть приймати різні значення. Так, надій п'яти корів при зменшенні рівня протеїнового живлення також знизиться, але для кожної корови це зменшення буде різним. Якщо функціональні зв'язки однаково просто виявити як на одиничних, так і на групових об'єктах, то кореляційні можна встановити лише при вивченні групових об'єктів методами математичної статистики, з усвідомленням того, що кореляційний зв'язок типовий для об'єктів і процесів, які відбуваються у живій природі.

Зв'язки між мінливими селекціонованими ознаками мають різні тенденції. В одному випадку, коли із збільшенням першої ознаки збільшується друга, говорять про *прямий і позитивний* (додатний) зв'язок або про *пряму і позитивну* кореляцію, а коефіцієнт кореляції позначається із знаком плюс (+ $r$ ). В іншому випадку із збільшенням першої ознаки друга зменшується. У цьому разі говорять про *зворотний* або *від'ємний* зв'язок і коефіцієнт кореляції пишеться із знаком мінус (- $r$ ). Отже, коефіцієнт кореляції вказує не лише на величину сполучної мінливості ознак, але й на напрямок зв'язку.

За формою коефіцієнт кореляції поділяється на *лінійний* або *прямолінійний* та *нелінійний* або *криволінійний*.

Таким чином, кореляційний аналіз є кількісним методом з'ясування тісноти сполучної мінливості між ознаками, напрямку зв'язку та його форми. Проте необхідно пам'ятати, що кореляційний зв'язок усе ж таки залишається методом статистичного аналізу, а не біологічного. Тому, не дивлячись на велике значення цього методу в біометрії, все ж не слід його переоцінювати і тим більше підміняти формально статистичним методом глибокий біологічний аналіз результатів у дослідах чи спостереженнях.

Ступінь зв'язку між ознаками вимірюється за допомогою коефіцієнтів парної кореляції ( $r$ ) та регресії ( $R$ ).

Спрямованість і ступінь (сила) зв'язку визначаються у відносних величинах від -1,0 до +1,0.

Прийнято вважати, що залежність низька, якщо  $r=0,1-0,3$ ; середня при  $r=0,4-0,6$  і висока – при 0,7 і більше.

При коефіцієнті кореляції  $r=0,1$  мінливість другої ознаки лише на 1% залежить від мінливості першої, а в 99% випадках – від випадкових факторів.

Розрізняють фенотипову, генотипову та середовищну кореляції.

**Фенотипова кореляція** дає змогу встановити зв'язок між різними господарськи корисними ознаками і використати їх з метою селекції, а також для ранньої діагностики продуктивності тварин за ознаками, корелятивно пов'язаними з продуктивністю.

**Генотипова кореляція** – це форма зв'язку між двома ознаками, зумовлена адитивною дією і взаємодією генів (плейотропія, епістаз тощо). Генотипова кореляція визначає вплив спадкових факторів на ступінь фенотипової кореляції.

**Паратипова (середовищна) кореляція** зумовлена силою і спрямованістю впливу умов середовища на дві ознаки, що вивчаються.

В усіх випадках напрям кореляції (+ чи –) буде визначатися відповідно подібною або різною дією генів і середовища на пару ознак, яку вивчають.

Генетична кореляція відіграє важливу роль у генетиці і селекції тварин, оскільки вона успадковується і у випадках, якщо це небажано (наприклад, негативна кореляція між надоєм та вмістом жиру), перебороти її можна тільки шляхом на зменшення.

У селекційно-генетичних дослідженнях при аналізі повторюваності ознак, тобто їх спадкової стабільності у віковому аспекті, стійкості передачі в поколіннях (мати–дочка, батько–син–онук і т. д.), використовується розрахунок детермінанти для кількісних ознак, або ранговий коефіцієнт кореляції Спірмена для якісних ознак.

У тваринництві нерідко якість тварин характеризують рангом (ієрархією), який вони займають серед інших тварин. При створенні порід, ліній, кросів стежать, яке місце (ранг) займатимуть з них тварини на конкурсних випробуваннях. При оцінці плідників за якістю потомства також визначають їх ранг (в балах, індексах).

Високий коефіцієнт рангової кореляції показує повторюваність оцінок плідників.

#### **4.1. Розрахунок коефіцієнта кореляції для малих вибірок**

Для розрахунку коефіцієнта кореляції в малих вибірках найчастіше використовують наступну формулу:

$$r = \frac{\Sigma xy - \frac{\Sigma x \cdot \Sigma y}{n}}{\sqrt{\sigma_x \cdot \sigma_y}} \quad (18)$$

де:  $n$  – кількість тварин вибірки;

$x$  та  $y$  – значення варіант першої та другої ознак;

$$\sigma_x = \Sigma x^2 - \frac{(\Sigma x)^2}{n} \quad (19)$$

$$\sigma_y = \Sigma y^2 - \frac{(\Sigma y)^2}{n} \quad (20)$$

Для визначення вірогідності кореляційного зв'язку розраховують критерій достовірності коефіцієнта кореляції за формулою:

$$t_r = \frac{r}{m_r} \quad (21)$$

де:  $t_r$  – критерій достовірності коефіцієнта кореляції;

$r$  – коефіцієнт кореляції;

$m_r$  – похибка коефіцієнта кореляції.

Похибка коефіцієнта кореляції визначається за формулою:

$$m_r = \sqrt{\frac{1-r^2}{n-2}} \quad (22)$$

#### 4.1.1. Приклади розв'язування задач

Визначити спрямованість і величину кореляційного зв'язку між віком та багатоплідністю свиноматок. Вік матки виражається порядковим номером опоросу, а багатоплідність – кількістю поросят за опорос. Врахована багатоплідність 10 свиноматок.

У табл. 3 приведені вихідні дані про вік та багатоплідність свиноматок, з яких одержують необхідні для кожної форми значення суми.

Одержані значення  $\Sigma x$ ,  $\Sigma y$ ,  $\Sigma x^2$ ,  $\Sigma y^2$  дозволяють розрахувати  $\sigma_x$  та  $\sigma_y$ :

Підставивши одержані значення  $\sigma_x$  та  $\sigma_y$ ,  $\Sigma xy$ ,  $\Sigma x$ ,  $\Sigma y$ ,  $n$  у формулу коефіцієнта кореляції для малої вибірки, розраховуємо його:

Таблиця 3

**Визначення зв'язку між віком свиноматок та  
кількістю поросят за опорос**

Номер тварини	Кількість опоросів ( $x$ )	Кількість поросят за опорос ( $y$ )	$x \cdot y$	$x^2$	$y^2$
1	2	9	18	4	81
2	1	7	7	1	49
3	5	11	55	25	121
4	7	12	84	49	144
5	3	9	27	9	81
6	2	7	14	4	49
7	6	11	66	36	121
8	1	6	6	1	36
9	4	12	48	16	144
10	3	14	42	9	196
$\Sigma n = 10$	$\Sigma x = 34$	$\Sigma y = 98$	$\Sigma xy = 367$	$\Sigma x^2 = 154$	$\Sigma y^2 = 1022$

$$\sigma_x = \Sigma x^2 - \frac{(\Sigma x)^2}{10} = 154 - \frac{34^2}{10} = 154 - 115,6 = 38,4;$$

$$\sigma_y = \Sigma y^2 - \frac{(\Sigma y)^2}{10} = 1022 - \frac{98^2}{10} = 1022 - 960,4 = 61,6;$$

$$r = \frac{\Sigma xy - \frac{\Sigma x \cdot \Sigma y}{n}}{\sqrt{\sigma_x \cdot \sigma_y}} = \frac{367 - \frac{34 \cdot 98}{10}}{\sqrt{38,4 \cdot 61,6}} = \frac{367 - 333,2}{\sqrt{2365,4}} = \frac{33,8}{48,6} = +0,695.$$

**Висновок.** Між віком та багатоплідністю свиноматок встановлений достатньо високий позитивний кореляційний зв'язок, тобто з віком багатоплідність у свиноматок підвищується.

Після визначення коефіцієнта кореляції потрібно встановити його достовірність ( $t_r$ ), але спочатку розраховується похибка коефіцієнта кореляції за формулою:

$$m_r = \sqrt{\frac{1-r^2}{n-2}} = \sqrt{\frac{1-0,695^2}{10-2}} = \sqrt{\frac{0,305}{8}} = \sqrt{0,038} = 0,195;$$

Достовірність коефіцієнта кореляції між віком та багатоплідністю свиноматок становить:

$$t_r = \frac{r}{m_r} = \frac{0,695}{0,195} = 3,56.$$

Визначивши  $t_r$ , порівнюємо одержане його значення з величиною  $t$  стандартного за таблицею Ст'юдента (додаток 12). Якщо розраховане  $t_r$  буде більше або дорівнювати теоретично очікуваній величині  $t$ , тоді коефіцієнт кореляції є достовірним при відповідно взятому рівні Р.

Для розрахунку прикладу табличні значення  $t$  при  $v = n - 2 = 10 - 2 = 8$  становили: 2,26; 3,36; 5,04. Розраховане значення  $t_r$  (3,56) менше табличного 5,04 та більше за 3,36, тобто отримана вірогідність на рівні  $P < 0,01$ , що підтверджує достовірність кореляційного зв'язку та говорить про його не випадковий характер.

#### **4.2. Порядок побудови та обробки кореляційної решітки для великої вибірки**

Кореляційна решітка являє собою сукупність двох варіаційних рядів для першої та другої корельованих ознак. Тому послідовність побудови та обробки кореляційної решітки така ж сама як і варіаційного ряду.

1. Обробку варіант ознак починають з їх умовного позначення – одна з ознак позначається через  $x$ , друга через  $y$ .
2. Вираховують кількість варіант (тварин) у вибірці ( $n$ ).
3. Визначають розмах мінливості за ознаками  $x$  та  $y$ : це різниця між максимальним та мінімальним значенням за кожною ознакою  $\lim_x, \lim_y$ .
4. Розраховують величину класового проміжку для ознак  $x$  та  $y$ .
5. Будують кореляційну решітку (табл. 4).
6. Будують класи за ознаками  $x$  та  $y$ .
7. Розносять попарно варіанти ознак  $x$  та  $y$  по вічках кореляційних ґраток, які утворюються перетином відповідних класів.
8. Підраховують кількість варіант за класами ознаки  $x$ , заповнюючи рядок  $P_x$  та за класами ознаки  $y$ , заповнюючи графу  $P_y$ .
9. Виділяють модальні класи та визначають числові відхилення для кожного класу від модальних. За графою  $a_x$  угору від модального класу будуть відхилення зі знаком мінус, вниз – з позитивним знаком, у рядку  $a_y$  ліворуч від модального класу – зі знаком мінус, праворуч – із знаком плюс. В результаті виділення модальних класів решітка розбивається на чотири квадранти.
10. Після заповнення графи  $P_x a_x$  і рядка  $P_y a_y$  одержують суму цих добутків.

Таблиця 4

**Кореляційна решітка**

Межі класів ознаки $x$	Межі класів ознаки $y$								
						$P_x$	$a_x$	$P_x a_x$	$P_x a_x^2$
	I			II					
	III			IV					
$P_y$								$\Sigma P_x a_x$	$\Sigma P_x a_x^2$
$a_y$									
$P_y a_y$						$\Sigma P_y a_y$			
$P_y a_y^2$						$\Sigma P_y a_y^2$			

11. Одержують суму добутків  $P_x a_x^2$  ( $P_x a_x \cdot a_x$ ) та  $P_y a_y^2$  ( $P_y a_y \cdot a_y$ ).



12. Визначають  $\Sigma Pa_x a_y$  (суму частот варіант на відхилення за кожною частотою зокрема за ознаками  $x$  та  $y$ ).

Для цього слід зробити наступне:

а) у кожному вічку, яке має частоту варіант, поставити добуток, який одержаний в результаті множення значень  $a_x$  на  $a_y$ . Добутки за частотами записують у вигляді ступеня, наприклад,  $1^{25}$  (1 – частота,  $^{25}$  – добуток);

б) у кожному квадранті решітки помножити частоти варіант на їхні добутки і одержати суму цих добутків за кожним квадрантом;

в) за сумами добутків чотирьох квадрантів визначають загальну суму і одержують  $\Sigma Pa_x a_y$ .

13. Визначають поправки до ознак  $x$  та  $y$ :

$$b_x = \frac{\Sigma P_x \cdot a_x}{n} \quad (23)$$

$$b_y = \frac{\Sigma P_y \cdot a_y}{n} \quad (24)$$

14. Розраховують середньоквадратичне відхилення (сигму) за ознаками  $x$  та  $y$ :

$$\sigma_x = \sqrt{\frac{\Sigma P_x \cdot a_x^2}{n} - b_x^2} \quad (25)$$

$$\sigma_y = \sqrt{\frac{\Sigma P_y \cdot a_y^2}{n} - b_y^2} \quad (26)$$

15. Підставимо одержані значення у формулу коефіцієнта кореляції, розраховують його значення:

$$r = \frac{\Sigma Pa_x \cdot a_y - n \cdot b_x \cdot b_y}{n \cdot \sigma_x \cdot \sigma_y} \quad (27)$$

16. Розраховують помилку коефіцієнта кореляції  $m_r$ :

$$m_r = \frac{1 - r^2}{\sqrt{n}} \quad (28)$$

17. Встановлюють достовірність кореляційного зв'язку. Для цього розраховують критерій вірогідності  $t_r$ :

$$t_r = \frac{r}{m_r}$$

### 4.3. Розрахунок коефіцієнта кореляції для великої вибірки

Для великої вибірки коефіцієнт кореляції розраховується за формулою:

$$r = \frac{\Sigma Pa_x a_y - nb_x b_y}{n \sigma_x \sigma_y} \quad (29)$$

де:  $x$  та  $y$  – значення варіант першої та другої ознаки;

$\Sigma Pa_x a_y$  – сума добутків частот на відхилення за двома ознаками;

$n$  – кількість тварин у вибірці;

$b_x$  та  $b_y$  – поправка до умовних середніх арифметичних першої та другої ознаки;

$\sigma_x$  та  $\sigma_y$  – середньоквадратичне відхилення за обома ознаками.

Елементи формули коефіцієнта кореляції  $\Sigma Pa_x a_y$ ,  $b_x$  та  $b_y$ ,  $\sigma_x$  та  $\sigma_y$  – визначаються у процесі побудови та обробки кореляційної решітки.

#### 4.3.1. Приклад виконання типового завдання

Визначити напрям та величину кореляційного зв'язку між величиною проміру обхвату грудей за лопатками та живою масою у корів симентальської породи за показниками, що наведені у табл. 5.

Позначаємо селекційні ознаки та відому величину вибірки:

$x$  – жива маса;  $y$  – обхват грудей;  $n = 130$

Для визначення класового інтервалу вираховуємо розмах мінливості у межах максимального та мінімального значень ліміту:

$$\lim_x = 864 \text{ кг} - 443 \text{ кг} = 421 \text{ кг};$$

$$\lim_y = 218 \text{ см} - 180 \text{ см} = 38 \text{ см}.$$

Таблица 5

**Дані про живу масу (кг,  $V_x$ ) та проміру – обхвату грудей за лопатками (см  $V_y$ ) корів симентальської породи**

№ п/п	$V_x$	$V_y$	№ п/п	$V_x$	$V_y$	№ п/п	$V_x$	$V_y$	№ п/п	$V_x$	$V_y$	№ п/п	$V_x$	$V_y$
1	602	196	27	615	202	53	590	204	79	610	200	105	580	197
2	680	210	28	670	201	54	500	<b>180</b>	80	697	211	106	713	200
3	610	200	29	705	202	55	590	192	81	660	202	107	495	188
4	591	205	30	590	202	56	612	202	82	540	196	108	635	203
5	715	206	31	494	196	57	515	187	83	725	205	109	690	216
6	650	205	32	600	203	58	750	201	84	597	208	110	615	200
7	635	195	33	555	198	59	663	210	85	630	204	111	674	198
8	<b>443</b>	191	34	675	210	60	555	194	86	785	215	112	670	199
9	685	211	35	659	195	61	770	206	87	598	205	113	570	197

10	625	197	36	735	210	62	678	205	88	566	188	114	860	215
11	680	206	37	695	200	63	568	195	89	680	207	115	553	190
12	821	216	38	470	180	64	601	200	90	640	200	116	670	197
13	600	205	39	600	193	65	643	197	91	560	194	117	570	194
14	670	205	40	680	214	66	645	210	92	680	202	118	650	198
15	735	216	41	685	208	67	560	201	93	670	212	119	625	204
16	615	194	42	637	204	68	650	212	94	533	195	120	655	202
17	615	198	43	625	205	69	740	210	95	828	210	121	575	200
18	593	192	44	600	200	70	675	201	96	590	197	122	600	194
19	582	193	45	634	192	71	760	202	97	849	<b>218</b>	123	736	208
20	787	210	46	745	206	72	<b>864</b>	214	98	732	206	124	654	206
21	687	197	47	715	206	73	640	200	99	750	203	125	686	203
22	730	207	48	600	202	74	642	198	100	721	206	126	520	190
23	590	200	49	653	203	75	670	204	101	695	210	127	667	205
24	597	197	50	578	192	76	532	192	102	645	203	128	560	197
25	605	200	51	625	200	77	660	210	103	527	192	129	670	208
26	666	205	52	645	205	78	573	197	104	620	195	130	681	201

Визначаємо класові інтервали для оцінюваних ознак:

$$k_x = \frac{\lim_x}{\text{класи}} = \frac{421}{10} = 42,1 \approx 40 \text{ кг};$$

$$k_y = \frac{\lim_y}{\text{класи}} = \frac{38}{10} = 3,8 \approx 4,0 \text{ см}.$$

Будуємо кореляційну решітку (табл. 6) таким чином, по горизонталі відкладаємо класи обхвату грудей, а по вертикалі – живу масу корів, розносимо у її клітини інформаційні дані із таблиці 6 та розраховуємо значення:  $\Sigma P_x a_x$ ,  $\Sigma P_y a_y$ ,  $\Sigma P_x a_x^2$ ,  $\Sigma P_y a_y^2$ .

Проводимо обрахування найбільш складного багаточлена  $\Sigma P a_x a_y$ .

Оскільки кореляційна решітка розділена умовними середніми арифметичними (модальними класами) на чотири частини, які називаються квадрантами, обчислення проводяться по кожному із них. Квадранти нумерують зліва на право і зверху вниз.

Для цього:

а) у клітинках кожного окремого квадранту, в яких проставлені частоти, ставимо добутки одержані при множенні  $a_x \times a_y$ . Наприклад, клітинка першого квадранту із чотирьох, який утворений класами 441-480 та 180-183, містить

частоту 1, для класу 180-183  $a_y = -5$ , для класу 441-480  $a_x = -5$ . Отже, добуток буде становити:  $(-5) \times (-5) = 25$ , який і ставиться до частоти у вигляді ступеня і т.д.;

б) у кожному квадранті перемножуємо частоти на їх добутки, одержані при множенні  $a_x \times a_y$ , і визначаємо загальну суму цих добутків за кожним квадрантом:

1 квадрант:  $25+20+16+15+24+9+6+30+32+8+4+9+10+3=211$ ;

2 квадрант:  $(-8)+(-4)+(-8) = -20$ ;

3 квадрант:  $-1$ ;

4 квадрант:  $2+10+3+8+12+6+8+9+30+4+8+16+20=136$

в)  $\Sigma P a_x \cdot a_y = I + II + III + IV$  квадранти  $= 211 + (-20) + (-1) + 121 = 326$ .

Таблиця 6

**Кореляційна решітка для розрахунку коефіцієнта кореляції між проміром обхвату грудей та живою масою корів**

Жива маса, x	Обхват грудей, y										$P_x$	$a_x$	$P_x a_x$	$P_x a_x^2$
	180-183	184-187	188-191	192-195	196-199	200-203	204-207	208-211	212-215	216-219				
441-480	$1^{25}$		$1^{15}$								2	-5	-10	50
481-550	$1^{20}$	$1^{16}$	$2^{12}$		$1^4$						5	-4	-20	60
521-560		<b>I-й</b>	$1^9$	$5^6$	$3^3$	1			<b>II-й</b>		10	-3	-30	90
561-600			$1^6$	$8^4$	$5^2$	6	$4^{-2}$	$2^{-4}$			26	-2	-52	104
601-640				$4^2$	$3^1$	11	$4^{-1}$				22	-1	-22	22
641-680				1	6	7	10	5	3		32	0	0	0
681-720					$1^{-1}$	5	$2^1$	$4^2$		$1^4$	13	1	13	13
721-760						3	$5^2$	$3^4$		$1^8$	12	2	24	48
761-800							$1^3$	$1^6$	$1^9$		3	3	9	27
801-840		<b>III-й</b>						$1^8$	<b>IV-й</b>	$1^{16}$	2	4	8	32
841-880									$2^{15}$	$1^{20}$	3	5	15	75
$P_y$	2	1	5	18	19	33	26	16	6	4	n=130; $\Sigma P_x a_x = -65$ ; $\Sigma P_x a_x^2 = 521$			
$a_y$	-5	-4	-3	-2	-1	0	1	2	3	4				
$P_y a_y$	-10	-4	-15	-36	-19	0	26	32	18	16	$\Sigma P_y a_y = 8$			
$P_y a_y^2$	50	16	45	72	19	0	26	64	54	64	$\Sigma P_y a_y^2 = 410$			

Визначаємо складові формули розпочинаючи з обчислення кореляції окремо за кожною ознакою:

$$b_x = \frac{\Sigma P_x \cdot a_x}{n} = \frac{-65}{130} = -0,500 \text{ кг};$$

$$b_y = \frac{\Sigma P_y \cdot a_y}{n} = \frac{8}{130} = 0,062 \text{ см};$$

$$\sigma_x = \sqrt{\frac{\Sigma P_x \cdot a_x^2}{n}} - b_x^2 = \sqrt{\frac{521}{130}} = 2,00 - (-0,500)^2 = 2,00 - 0,25 = \pm 1,75 \text{ кг};$$

$$\sigma_y = \sqrt{\frac{\Sigma P_y \cdot a_y^2}{n}} - b_y^2 = \sqrt{\frac{410}{130}} - 0,062^2 = 1,776 - 0,004 = \pm 1,772 \text{ см};$$

$$r = \frac{\Sigma P a_x \cdot a_y - n \cdot b_x \cdot b_y}{n \cdot \sigma_x \cdot \sigma_y} = \frac{326 - 130 \cdot (-0,500) \cdot 0,062}{130 \cdot 1,75 \cdot 1,772} = \frac{322,0}{403,1} = +0,799;$$

**Висновок:** Між проміром обхвату грудей та живою масою корів симентальської породи встановлений високий, позитивний кореляційний зв'язок, тобто із збільшенням живої маси тіла корів у них істотно зростає промір обхвату грудей за лопатками.

Щоб довести, що вирахований кореляційний зв'язок не є випадковістю, визначаємо критерій достовірності. Оскільки для формули з його визначення невідома похибка, знаходимо її значення:

$$m_r = \frac{1 - r^2}{\sqrt{n}} = \frac{1 - 0,799^2}{\sqrt{130}} = \frac{1 - 0,638}{11,4} = \frac{0,362}{11,4} = 0,032;$$

$$t_r = \frac{r}{m_r} = \frac{0,799}{0,032} = 24,9.$$

Зіставляємо фактично розраховане значення  $t_r$  з теоретичною величиною  $t$  таблиці Ст'юдента при  $v=130$ . Розраховане  $t_r$  істотно більше максимального табличного (3,29). Отже, кореляційний зв'язок між живою масою та обхватом грудей у корів є високодостовірним при  $P < 0,001$ .

## 5. РЕГРЕСІЙНИЙ АНАЛІЗ

Поряд з кореляційним аналізом у селекції тварин широко використовують регресійний аналіз як метод вивчення зв'язку між ознаками, що вказує, на яку величину зміниться одна із залежних ознак (функція) при зміні іншої ознаки (аргумента) на одиницю виміру (кг, мл, см, бал).

**Регресія** (від лат. *regressio* – рух назад) – це зміна функцій від зміни одного або кількох аргументів.

Зв'язок між ознаками можна виражати не тільки коефіцієнтом кореляції, але й іншими показником – коефіцієнтом регресії (R).

На відміну від коефіцієнта кореляції, який показує величину у відносних одиницях, коефіцієнт регресії виражає зв'язок в поіменованих величинах (у тих в яких вимірюється ознака). Коефіцієнт регресії показує наскільки в середньому зміниться одна із ознак, якщо інша, яка корелює з нею, зміниться на якусь певну величину (одиницю вимірювання).

Коефіцієнт регресії може не тільки показувати наскільки змінлася одна із ознак при зміні іншої на одиницю, але й вказати на наявність та напрямок зв'язку між досліджуваними ознаками (знак “+” або “–” перед величиною).

Коефіцієнт регресії можна розрахувати для малої та великої вибірок за формулами:

$$R_{x/y} = r \frac{\sigma_x}{\sigma_y} \quad (30)$$

$$R_{y/x} = r \frac{\sigma_y}{\sigma_x} \quad (31)$$

де:  $x$  – умовне позначення однієї з вивчаючих ознак;

$y$  – умовне позначення другої, яка корелює з першою ознакою;

$\sigma_x$  та  $\sigma_y$  – середнє квадратичне відхилення за кожною із оцінюваних ознак;

$r$  – коефіцієнт кореляції, який відображає зв'язок між вивчаючими ознаками.

Однак для малих вибірок більш зручніші розгорнуті формули розрахунку коефіцієнта регресії:

$$R_{x/y} = \frac{\Sigma x \cdot y - \frac{\Sigma x \cdot \Sigma y}{n}}{\Sigma y^2 - \frac{(\Sigma y)^2}{n}} \quad (32)$$

$$R_{y/x} = \frac{\Sigma x \cdot y - \frac{\Sigma x \cdot \Sigma y}{n}}{\Sigma x^2 - \frac{(\Sigma x)^2}{n}} \quad (33)$$

Величина  $R_{x/y}$  показує наскільки змінюється ознака, позначена через “х”, якщо інша, яка корелює з нею (у), зміниться на одиницю (тобто на 1 кг, 1 см, 1 % та ін.).

Величина  $R_{y/x}$ , навпаки, показує наскільки зміниться ознака, яка позначена через (у), при зміні корелюючої з нею ознаки (х) на одиницю (тобто на 1 %, 1 см, 1 кг та ін.).

### Приклади розв’язування задач

Розрахувати коефіцієнт регресії живої маси у корів голштинської породи та обхвату грудей і навпаки за наступними даними:

жива маса, кг – 467, 441, 491, 433, 534, 488, 390, 391, 421, 447.

обхват грудей, см – 186, 182, 190, 183, 196, 189, 172, 174, 180, 179.

Для розрахунку коефіцієнта регресії позначаємо одну ознаку – живу масу корів – через “х”, іншу – обхват грудей – через “у”.

Так як дана вибірка мала, тоді більш доцільно використати для розрахунку регресії розгорнуті формули.

Щоб одержати всі необхідні для розрахунку регресії величини, зручніше всі розрахунки звести у таблицю, аналогічну таблиці для визначенні коефіцієнта кореляції для малої вибірки (табл. 7).

$$R_{x/y} = \frac{\Sigma x \cdot y - \frac{\Sigma x \cdot \Sigma y}{n}}{\Sigma y^2 - \frac{(\Sigma y)^2}{n}} = \frac{827456 - \frac{4503 \cdot 1831}{10}}{335747 - \frac{1831^2}{10}} = \frac{827456 - 824499,3}{335747 - 335256,1} = 6,02 \text{ кг}$$

$$R_{y/x} = \frac{\Sigma x \cdot y - \frac{\Sigma x \cdot \Sigma y}{n}}{\Sigma x^2 - \frac{(\Sigma x)^2}{n}} = \frac{827456 - \frac{4503 \cdot 1831}{10}}{2046471 - \frac{4503^2}{10}} = \frac{827456 - 824499,3}{2046471 - 2027700,9} = 0,157 \text{ см}$$

**Визначення коефіцієнта регресії (жива маса та обхват грудей у корів)**

Жива маса, кг ( $x$ )	Обхват грудей, см ( $y$ )	$x \cdot y$	$x^2$	$y^2$
467	186	86862	218089	34586
441	182	80262	194481	33124
491	190	93290	241081	36100
433	183	79239	187489	33489
534	196	104664	285156	38416
488	189	92232	238144	35721
390	172	67080	152100	29584
391	174	68034	152881	30276
421	180	75780	177241	32400
447	179	80013	199809	32041
$\Sigma x=4503$	$\Sigma y=1831$	$\Sigma xy=827456$	$\Sigma x^2=2046471$	$\Sigma y^2=335747$

**Висновок:** Між живою масою корів голштинської породи та обхватом грудей існує позитивний кореляційний зв'язок. Підвищення обхвату грудей на 1 см призведе до підвищення живої маси на 6,02 кг; підвищення живої маси корів на 1 кг викликає у них збільшення проміру обхвату грудей на 0,16 см.



## 6. ДИСПЕРСІЙНИЙ АНАЛІЗ

Поряд з кореляційним та регресійним аналізами при вивченні причинно-наслідкових відносин між явищами особливо цінним виявився метод дисперсійного аналізу. Розроблений англійським математиком і біологом Р.Е. Фішером і опублікований у 1925 р., дисперсійний аналіз у даний час знайшов широке застосування при обробці матеріалів науково-дослідних робіт з сільськогосподарськими тваринами і рослинами.

Закономірно, що біологічні властивості живих організмів та рівень розвитку селекціонованих ознак тварин знаходяться у великій залежності від дії спадкових та середовищних факторів. У зв'язку з цим і виникла необхідність відокремлено визначати ступінь впливу кожного із цих чинників у загальній мінливості взятих для дослідження показників.

Головне призначення дисперсійного аналізу – це розкласти загальну варіативність ознаки на часткову мінливість, що виникає у членів популяції під впливом багатьох чинників. В процесі аналізу встановлюють частку мінливості, зумовлену кожним фактором, що враховується у дослідженні, частку мінливості, викликану сумісною дією цих факторів, а також виявляють ту частку мінливості, яка є результатом дії багатьох не врахованих у досліді факторів, що утворюють так звану випадкову, або залишкову, мінливість ознаки.

*Отже, сутність дисперсійного аналізу полягає у тому, що він дозволяє встановити ступінь впливу окремих факторів на загальну мінливість тієї чи іншої ознаки.*

Для пізнання закономірностей мінливості об'єктів якої-небудь сукупності важливо визначити частку впливу окремих факторів, вивчення дії яких на дані об'єкти може служити метою спеціальних дослідів. Обробляючи матеріали досліджень дисперсійним методом, можна одержати математичний вираз і визначити величину мінливості, обумовлену дією врахованих в досліді факторів, і мінливість залишкову, що виникла під впливом всіх інших, не врахованих в досліді, факторів. У цьому полягає перша і найбільш важлива властивість дисперсійного аналізу.

Друга властивість дисперсійного аналізу полягає в тому, що він дозволяє визначити статистичну достовірність частки впливу факторів, що вивчаються. Важливою особливістю дисперсійного методу є те, що його можна проводити не тільки на малих і великих вибірках, але й, що особливо важливо, він дозволяє обробляти сукупності, які включають неоднорідний матеріал.

Наприклад, сукупність може складатися з тварин різної статі, в яку входить група самок і самців. Або дисперсійному аналізу піддають сукупність, що складається з груп тварин різного генетичного походження (чистопорідні та помісні) і т.д.

Ознаки, що змінюються під дією тих чи інших причин, називаються *результативними*, а причини, що викликають зміну величини результативної ознаки чи ознак, – *факторами*. Наприклад, жива маса, продуктивність або лінійні ознаки екстер'єру, за якими ми складаємо уяву про розвиток та потенціальні можливості організму – це все ознаки, на які впливають різні фактори: рівень годівлі, елементи живлення, технологія утримання тощо.

Факторів, що впливають на одну і ту ж ознаку, багато. Проте у досліді регулюються лише окремі з них, вони так і називаються *регульованими* або *організованими* факторами на відміну від тих, які не піддаються регулюванню, хоча й впливають на величину результативної ознаки. Як правило кожний регульований фактор досліджується серійно, тобто у вигляді відокремлених одна від одної груп, які називаються *градаціями*. Кількість градацій того чи іншого фактора визначаються умовами досліду, наприклад поживністю чи складом раціону годівлі

Користуючись дисперсійним аналізом можна вирішити частину основних задач:

- визначити силу впливу закономірних та випадкових факторів;
- визначити достовірність впливу досліджуваних факторів;
- з'ясувати зв'язок між окремими факторами;
- виміряти ступінь успадкованості ознак при передачі генетичної інформації із покоління в покоління;
- провести порівняльний аналіз спадкових комбінаційних властивостей племінних плідників, як основу масового добору

Частка впливу окремих факторів виражається, так званою дисперсією, тобто сумою квадратів відхилення і позначається буквою  $C$ . Дисперсія математично це не що інше як  $\sigma^2$ . Тому прийнято визначати загальну дисперсію або мінливість, як суму квадратів відхилення кожного варіанта від середньої арифметичної, що дозволяє записати її наступною формулою:

$$C_y = \Sigma(V-M)^2 \quad (34)$$

Загальну дисперсію можна розкласти на відповідні складові – факторіальну дисперсію ( $C_x$ ), яка визначає частку впливу факторів, що вивчаються (організовані фактори) та залишкову дисперсію ( $C_z$ ), яка виникає під дією різних випадкових не врахованих факторів.

Таким чином, у загальному вигляді дисперсія, тобто різноманітність або мінливість ознаки може бути записана наступною формулою:

$$C_y = C_x + C_z \quad (35)$$

Завданням дисперсійного аналізу є вирахування факторіальної ( $C_x$ ) та залишкової ( $C_z$ ) дисперсій.

За допомогою дисперсійного аналізу визначають не лише величину кожної дисперсії, але й вираховують і частку чи силу впливу  $C_x$  та  $C_z$  на варіюючи ознаку. Для цього обчислюють величину відношення між дисперсіями і позначають її через  $\eta^2$ .

Таким чином, частка чи ступінь впливу усіх взятих до уваги організованих факторів виразиться формулою:

$$\eta^2_x = \frac{C_x}{C_y} \quad (36)$$

Ступінь впливу випадкових факторів буде виражена відповідною формулою:

$$\eta^2_z = \frac{C_z}{C_y} \quad (37)$$

Дисперсійний аналіз проводять у декілька етапів шляхом математичної обробки статистичного комплексу та складання зведеної таблиці.

## 6.1. Статистичні комплекси

Розрізняють статистичні комплекси за тим, скільки факторів включено до обчислення, тому вони поділяються на – *однофакторні, двохфакторні, трьохфакторні та багатфакторні*.

Фактори впливу поділяються на класи і в залежності від співвідношення частот у цих класах статистичні комплекси можуть бути *рівномірними*, *нерівномірними* та *пропорційними*. У рівномірних комплексах число спостережень, тобто частоти знаходяться у співвідношеннях, як 1:1:1 і т.п., у пропорційних – як 2:4 і т.п., у нерівномірних – як завгодно, табл. 8.

Математичні формули та алгоритми обрахунків залежать від характеру комплексів. Найскладніші обрахунки існують для нерівномірних багатфакторних комплексів.

Таблиця 8

**Статистичні комплекси залежно від  
співвідношення частот у класах**

Рівномірні					Пропорційні					Нерівномірні				
показ- ники	A <sub>1</sub>		A <sub>2</sub>		показ- ники	A <sub>1</sub>		A <sub>2</sub>		показ- ники	A <sub>1</sub>		A <sub>2</sub>	
	B <sub>1</sub>	B <sub>2</sub>	B <sub>1</sub>	B <sub>2</sub>		B <sub>1</sub>	B <sub>2</sub>	B <sub>1</sub>	B <sub>2</sub>		B <sub>1</sub>	B <sub>2</sub>	B <sub>1</sub>	B <sub>2</sub>
n	10	10	10	10	n	5	10	7	21	n	10	12	13	14
B <sub>1</sub> :B <sub>2</sub>	1:1		1:1		B <sub>1</sub> :B <sub>2</sub>	1:2		1:3		B <sub>1</sub> :B <sub>2</sub>	1:1,2		1:1,5	

**6.2. Основні терміни і величини,  
що використовуються у дисперсійному аналізі**

При проведенні дисперсійного аналізу, користуються наступними термінами і величинами:  $C_y$  – загальна дисперсія;  $C_x$  – часткова, або факторіальна, дисперсія;  $C_z$  – випадкова, або залишкова, дисперсія.

Визначення дисперсій ( $C_y$ ,  $C_x$ ,  $C_z$ ) ґрунтується в процесі дисперсійного аналізу на обчисленні суми квадратів різниці між величиною варіативної ознаки ( $V$ ) і загальній середній арифметичній ( $M_0$ ) або частковими середніми арифметичними ( $M_i$ ), а також на обчисленні суми квадратів різниці між частковою ( $M_i$ ) і загальною ( $M_0$ ) середніми арифметичними.

Результативна ознака ( $y$ ) – варіативна ознака, заради якої проводиться дисперсійний аналіз мінливості, що виникає під впливом якихось факторів, що враховуються в експерименті;

*враховані (або організовані) фактори ( $x$ )* – чинники, які вивчаються і контролюються у досліді;

*невраховані (або випадкові) фактори ( $z$ )*, які не враховуються і не контролюються схемою досліду, але чинять вплив на варіативність ознаки;

*градації або класи ( $l$ ) фактора*, що впливає на варіативність ознаки;

*варіанта результативної ознаки ( $V$ )*;

*статистичний комплекс* – вибіркова сукупність, за якою обраховується дисперсійний аналіз і організовані за градаціями фактори, що впливають на варіативність ознаки;

*дисперсії*:

*загальна дисперсія*  $C_y = \Sigma(V-M_0)^2$ ;

*часткова або факторіальна дисперсія*  $C_x = \Sigma(M_i-M_0)^2$ ;

*залишкова або випадкова дисперсія*  $C_z = \Sigma(V-M_i)^2$ ;

*ступені свободи*: для загальної дисперсії ( $v_y$ ); для факторіальної дисперсії ( $v_x$ ) і для залишкової дисперсії ( $v_z$ );

*варіанси, девіати (або вагові дисперсії) ( $\sigma^2$ )*:

*доля мінливості, зумовлена впливом фактора  $\eta^2$* :

$$\text{факторіальна } \eta_x^2 = \frac{C_x}{C_y}; \text{ залишкова } \eta_z^2 = \frac{C_z}{C_y};$$

*показник криволінійного зв'язку ( $\eta$ )* між результативною ознакою і фактором, що впливає на варіативність ознаки;

*критерій достовірності Фішера ( $F$ )*.

Для проведення дисперсійного аналізу формують малу або велику вибірку за принципом випадкового добору членів сукупності. Правила добору при цьому залишаються ті ж самі, що були використані й для інших статистичних методів. Це означає, що дисперсійний аналіз вимагає оформлення **рандомізованої** вибірки, для якої добір членів будується за принципом випадковості. Недотримання цього принципу призводить до невірних

висновків, оскільки одержані величини  $C$ ,  $\sigma^2$  і  $\eta^2$  не відображатимуть властивостей генеральної сукупності.

Одержану рендомізовану вибірку оформляють у вигляді так званого статистичного комплексу, що представляє таблицю, структура якої може бути більш простою або складнішою, залежно від числа організованих чинників, що вивчаються, від числа градацій за кожний фактор і від об'єму вибірки.

Таблиця статистичного комплексу утворюється градаціями (класами) фактора, що впливає (одного або декількох) і градаціями результативної ознаки. При вирішенні питання про те, які чинники можуть бути включені до структури статистичного комплексу, необхідно дотримуватися низки обов'язкових правил: при вивченні двох або більшого числа організованих факторів допускаються до включення в статистичний комплекс тільки такі чинники, які не знаходяться у кореляційному або функціональному зв'язку. Якщо ж між факторами існує зв'язок, то таку сукупність не можна обробляти звичайним способом дисперсійного аналізу, а необхідно використовувати інші методи.

Фактори, що підлягають вивченню методом дисперсійного аналізу, можуть мати декілька градацій (або рівнів), що враховуються у структурі статистичного комплексу. Наприклад, вивчається вплив статі та типу годівлі на добові прирости телят. Тут можуть бути такі градації фактора: для статі тварин ( $A$ ) дві градації (самки  $LA_1$  і самці  $LA_2$ ), для чинника годівлі ( $B$ ) — різні раціони: концентратний ( $LB_1$ ), соковитий ( $LB_2$ ), грубий ( $LB_3$ ). Отже, в комплексі по чиннику  $A$  будуть дві градації ( $LA = 2$ ) і по чиннику  $B$ , три — ( $LB = 3$ ).

За характером і особливостям фактора їх градації можуть бути наступними.

*Фіксовані* (визначені), коли градації характеризують певний стан чинника, наприклад за статтю (самки, самці), за типом мутацій (хромосомні, генні) і т.д.

*Варіативні* градації відносяться до кількісних ознак і оформляються у вигляді класів.

*Ієрархічні*, коли градації одного чинника жорстко супідрядні з градаціями іншого, наприклад ієрархічні градації можуть бути між тваринами за їх походженням в такій послідовності: вид, порода, лінія, плідники, нащадки.

Залежно від того, яким типом градацій характеризується фактор, структура статистичного комплексу і його обробка певною мірою

видозмінюються. Градації фактора можуть бути виражені рівновеликими кількісними величинами або якісними угрупованнями. Так, наприклад, градації по віковому фактору можуть бути підрозділені на вікові групи в послідовному порядку: 1 рік, 2 роки, 3 роки і т.д. Але, окрім кількісної відмінності, градації фактора можуть виражатися якісними підрозділами. Так, для вікового фактора градаціями можуть бути наступні якісні показники: молоді, середньовікові, старі. Якісні градації по генетичному фактору можна розділити на наступні: чистопородні, помісні, інбредні, гомозиготні, гетерозиготні тощо.

При складанні статистичного комплексу, коли в нього включені два або три фактори, число градацій по кожному з факторів може бути різним. Необхідно мати на увазі, що прийнятий рівень градації може впливати на результат дисперсійного аналізу. Тому одержані підсумкові дані аналізу можуть бути поширені тільки на той тип комплексу і на той рівень градацій, які піддавалися обробці.

На результатах дисперсійного аналізу позначається також рівень неврахованих, тобто неорганізованих і випадкових факторів, які впливають на результативну ознаку спільно з врахованими факторами. Наприклад, вивчається сила гетерозису, що виявляється у збільшенні величини живої маси при міжлінійних внутрішньопородних кросах і промислового схрещуванні.

Неврахованими, випадковими факторами будуть різноманітні фактори, що чинять з свого боку позитивний або негативний вплив на дисперсію живої маси. Так, наприклад, на тлі низького або високого рівня годівлі, яка входить до складу неврахованих факторів ( $z$ ), можна зробити різні висновки з дисперсійного аналізу про силу гетерозису за живою масою при різних типах генетичного походження тварин. Тому при проведенні дисперсійного аналізу необхідно враховувати рівень і специфіку можливої дії елементів, які створюють невраховані, випадкові чинники, і мати ясне уявлення про можливий їхній вплив на результативну ознаку. Інакше можна зробити неправильні висновки з даних, одержаних за допомогою дисперсійного аналізу.

Вказані особливості випадкових факторів та їхня роль впливу на дисперсію результативної ознаки складають слабку сторону дисперсійного методу і вимагають ретельного і свідомого підходу до матеріалу.

Отже, підсумовуючи наведений матеріал, нам відомо, що мінливість ознак у тварин зумовлена дією різноманітних факторів внутрішнього (спадкового) та зовнішнього середовища. Ці фактори діють на ознаки організму тварин з різною силою, а іноді й у різних напрямках. У результаті такої дії ознака набуває певної величини. Частка впливу різних, одночасно задіяних факторів, на мінливість тієї чи іншої ознаки, неоднакова. Виявляти частку впливу кожного, окремо взятого, фактора на мінливість ознак можна лише за

допомогою дисперсійного аналізу, основне призначення якого – розділити загальну мінливість на загальну, факторіальну і залишкову.

### Приклади розв’язування задач

Необхідно виявити вплив віку свиноматок великої білої породи на їхню молочність (загальну масу поросят у віці 21 дня); при цьому враховується наступний вік свиноматок – 16, 26, 38, та 50 місяців. З цією метою для кожної вікової групи підбирають по декілька особин.

Так, у 16-ти місячному віці молочність п’яти свиноматок склала 90, 88, 79, 74, 75 кг; у 26 місяців – 81, 87, 95, 86, 85 кг; у 38 місяців – 95, 78, 90, 85, 84 кг та у 50 місяців – 73, 82, 92, 90, 85 кг.

Спочатку визначаємо *фактор*, який необхідно вивчити. Для нашого прикладу – це вік свиноматки.

*Градаціями* фактора у нас є вік свиноматок, які позначаються буквою “*i*”. Загальна кількість врахованих градацій позначають як “*r*”. У нашому прикладі таких градацій фактора чотири – 16, 26, 38 та 50 місяців ( $\Sigma r=4$ ).

Тварини, які підібрані для кожної градації, складають її обсяг (*n*). У кожній особини, яка входить до комплексу вимірюється результативна ознака (у нашому прикладі молочність), тобто визначаються абсолютні значення варіанти (*Vi*). Градації фактору і відповідні їм показники варіант складають дисперсійний комплекс. Дисперсійний комплекс, а також розрахунок його допоміжних величин для визначення частки впливу вивчаючого фактору на мінливість ознаки оформляється у вигляді таблиці. Такий однофакторний дисперсійний комплекс для наведеного прикладу представлений у табл. 9.

Таблица 9

### Однофакторный дисперсионный комплекс

Градації фактора ( <i>i</i> )	16 місяців	26 місяців	38 місяців	50 місяців	$\Sigma$
<i>Vi</i> (варіанти градації)	90, 88, 79, 74, 75	81, 87, 95, 86, 85	95, 78, 90, 85, 84	73, 82, 92, 90, 85	
<i>n</i>	5	5	5	5	$N = 20$
$\Sigma Vi$	406	434	432	422	$\Sigma V = 1694$
$Mi = \frac{\Sigma Vi}{n}$	81,6	86,8	86,4	84,4	$M = \frac{\Sigma V}{n} = 84,7$
$Vi^2$	8100, 7744, 6241, 5476, 5625 $\Sigma_1 = 33186$	6561, 7569, 9025, 7396, 7225 $\Sigma_2 = 37775$	9025, 6084, 8100, 7225, 7056 $\Sigma_3 = 37490$	5329, 6724, 8464, 8100, 7225 $\Sigma_4 = 35842$	$\Sigma Vi^2 = 144293$
$(\Sigma Vi)^2$	164836,2	188356,1	186624,1	178084,2	
$H_i = \frac{(\Sigma Vi)^2}{n}$	32967,2	38019,2	37324,8	35616,8	$\Sigma H_i = 143928,0$



### 6.3. Порядок складання дисперсійного комплексу

#### та розрахунку допоміжних величин

1. У таблицю заносять градації фактору ( $i$ ) та проставляють їх загальну кількість ( $r$ ). У даному прикладі – це різні вікові групи свиноматок – 16, 26, 38 та 50 місяців. Всього 4 градації фактору ( $r = 4$ ).

2. Проставляємо абсолютні значення варіант ( $Vi$ ), які складають об'єм кожної градації (показники молочності свиноматок в кг).

3. Підраховуємо об'єм градації ( $n$ ) та кількість особин комплексу в цілому ( $N$ ). У кожному випадку об'єм кожної градації  $n = 5$  тварин,  $N = 20$ .

4. Знаходимо суму варіант кожної градації комплексу ( $\Sigma Vi$ ) і комплексу в цілому ( $\Sigma V$ ). Так, по 1-й градації  $\Sigma Vi = 90+88+79+74+75=406$ ; по 2-й градації – 434; по 3-й – 432 та по 4-й 422. Сума варіант комплексу в цілому  $\Sigma V = 406+434+432+422=1694$ .

5. Визначаємо середню арифметичну варіант кожної градації ( $Mi$ ), а також середню арифметичну варіант всього комплексу ( $M$ ):

$$Mi = \frac{\Sigma Vi}{n} \quad (38)$$

далі складаємо за градаціями:

$$\text{перша} - \frac{406}{5} = 81,2 \text{ кг};$$

$$\text{друга} - \frac{434}{5} = 86,8 \text{ кг};$$

$$\text{третя} - \frac{432}{5} = 86,4 \text{ кг};$$

$$\text{четверта} - \frac{422}{5} = 84,4 \text{ кг}.$$

$$M = \frac{\Sigma V}{N} \quad (39)$$

$$M = \frac{1694}{20} = 84,7 \text{ кг}.$$

6. Розраховуємо квадрати варіант градацій ( $Vi$ )<sup>2</sup> і знаходимо суму квадратів варіант всього комплексу ( $\Sigma Vi^2$ ). Це досягається піднесенням кожної варіанти у квадрат та наступним додаванням цих величин.

$$\Sigma Vi^2 = \Sigma_1 + \Sigma_2 + \Sigma_3 + \Sigma_4 = 33186 + 37775 + 37490 + 35842 = 144293.$$

7. Розраховуємо квадрат суми варіант градацій ( $\Sigma Vi$ )<sup>2</sup>, для чого підносимо показники графі  $\Sigma Vi$  до квадрату:  $406^2 = 164836,2$ ;  $434^2 = 188356,1$  та ін.

8. Визначаємо спеціальні допоміжні величини ( $Hi$ ) за формулою:

$$Hi = \frac{(\Sigma Vi)^2}{n} \quad (40)$$

які потім додають для одержання величини  $\Sigma Hi$ .

Так для 1-ої градації  $Hi = \frac{164836,2}{5} = 32967,2$ . Подібні розрахунки робимо для інших градацій.

$$\Sigma Hi = 32967,2 + 38019,2 + 37324,8 + 35618,8 = 143928,0.$$

9. Обраховуємо загальну додаткову величину ( $H_{\Sigma}$ ), для чого показник графі суми варіант всього комплексу ( $\Sigma V$ ) підносимо до квадрату і ділимо на показник графі числа особин комплексу, тобто його об'єм ( $N$ ):

$$H_{\Sigma} = \frac{(\Sigma V)^2}{N} \quad (41)$$

$$H_{\Sigma} = \frac{1694^2}{20} = \frac{2869366}{20} = 143481,8$$

Після повторення комплексу і знаходження всіх допоміжних величин, розраховуємо три дисперсії: загальну ( $C_y$ ), факторіальну ( $C_x$ ) та залишкову ( $C_z$ ). Робочі формули розрахунку цих величин наступні:

$$C_y = \Sigma Vi^2 - \Sigma H_{\Sigma} \quad (42)$$

$$C_x = \Sigma Hi - H_{\Sigma} \quad (43)$$

$$C_z = \Sigma Vi^2 - Hi \quad (44)$$

Підставивши числові значення у формули, одержуємо:

$$C_y = 144293 - 143481,8 = 811,2;$$

$$C_x = 143928 - 143481,8 = 446,2;$$

$$C_z = 144293 - 143928 = 365,0.$$

Вірність розрахунку величин дисперсій можна перевірити за формулою  $C_y = C_x + C_z = 811,2 = 446,2 + 365,0 = 811,2$ .

На підставі трьох дисперсій можна розраховуємо заключні показники дисперсійного аналізу:

10. Показник сили (частки) впливу фактору, який вивчається ( $\eta^2_x$ ):

$$\eta^2_x = \frac{C_x}{C_y}$$

Цей показник можна виражати як у частках одиниці, так і в відсотках. У наступних випадках формула набуває наступного виду:

$$\eta^2_x = \frac{C_x}{C_y} \cdot 100 \% \quad (45)$$

11. Показник вірогідності впливу ( $F$ ):

$$F = \frac{C_x}{C_z} \cdot \frac{N-r}{r-1} \quad (46)$$

Для нашого прикладу частка впливу віку свиноматок на їхню молочність дорівнює:

$$\eta^2_x = \frac{C_x}{C_y} \cdot 100 \% = \frac{446,2}{811} \cdot 100 = 0,55 \cdot 100 = 55\% .$$

Одержаний показник  $\eta^2_x$  говорить про достатньо сильний вплив віку свиноматок на їх молочність. З усіх факторів, які визначають різноманіття молочності, на частку віку свиноматок припадає 55%. Наскільки вірогідний цей висновок можна судити за показником вірогідності впливу ( $F$ ).

$$F = \frac{446,2}{365} \cdot \frac{20-4}{4-1} = 1,22 \cdot 5,33 = 6,50$$

Для того, щоб вияснити, вірогідний чи невірогідний вплив фактору, який ми вивчаємо у вибіркового дисперсійному комплексі, необхідно порівняти розраховане значення  $F$  зі стандартним значенням  $F_{st}$ , критерію Фішера, наведені у додатку 13.

У цьому додатку  $v_1$   $v_2$  – кількості ступенів свободи, які становлять відповідно  $v_1 = r - 1$ ;  $v_2 = N - r$ .

Для нашого прикладу:  $v_1 = 4 - 1 = 3$ ;  $v_2 = 20 - 4 = 16$ .

Розраховане значення  $F$  перевищує друге стандартне значення:

$F = 6,50 > 6,4$ , тобто вплив фактору, що вивчається при рівні вірогідності  $P < 0,01$ .

## 7. УСПАДКОВУВАНІСТЬ ОЗНАК ТА МЕТОДИ ЇЇ ВИЗНАЧЕННЯ

Селекційно-племінна робота із сільськогосподарськими тваринами постійно спрямована на удосконалення їхніх господарськи корисних ознак у бажаному напрямі. Добір та підбір цінних у племінному відношенні тварин проводиться за показниками оцінки фенотипу. Проте необхідно відмітити, що фенотип тварини ніколи не ідентичний його генотипу. Це пов'язано з тим, що кількісні ознаки проявляють явище фенотипової мінливості, яка обумовлена двома причинами: спадковою (генотиповою) різноманітністю тварин певного стада, популяції та впливом зовнішніх (паратипових) факторів.

Разом з тим схожі за фенотипом тварини мають несхожі генотипи. Від групи схожих за продуктивністю корів одержують потомство з різними продуктивними властивостями. Отже, у свою чергу, фенотипова схожість не гарантує генотипової схожості. Практично це означає, що не всі дочки корів-рекордисток за надоєм стануть відповідно рекордистками за цією ознакою, навіть якщо вони вирощені в схожих умовах середовища. Причина цього — генотипові відмінності схожих за фенотипом корів-матерів.

Не можна одержати точного уявлення про спадковість (генотип) тварини, судячи лише за однією його продуктивністю (за фенотипом); не можна знати напевно, які з кращих відібраних за фенотипом тварин будуть в той же час кращими і за генотипом. Правда, певна закономірність існує: кращі за фенотипом тварини будуть в середньому (а не кожна особина окремо), як правило, і кращими за генотипом, тобто вони даватимуть краще потомство. Селекціонери поступають абсолютно вірно, коли в своїй практичній роботі кращі генотипи шукають серед кращих фенотипів.

*Ступінь точності* оцінки генотипу за фенотипом істотно коливається залежно від того, наскільки мінливість тієї чи іншої ознаки, що селекціонується, обумовлена більшою чи меншою мірою умовами зовнішніх факторів. Адже різні ознаки з неоднаковою мірою змінюються під впливом умов середовища.

Так величина надою значно більше залежить від рівня і повноцінності годівлі ніж вміст жиру в молоці.

Отже, чим більше дана ознака змінюється під дією середовища (чим більша частка паратипової мінливості в загальній фенотиповій різноманітності

ознаки), тим важче визначити існуючі відмінності в племінній цінності (генотипові відмінності) тварин, тим менш ефективним буде масовий добір тварин за такою ознакою. І навпаки, ефективність добору особин на племінні цілі за тією чи іншою ознакою зростає у міру зменшення її мінливості залежно від умов зовнішнього середовища.

Таким чином, щоб підвищити ефективність добору, потрібно перш за все встановити, якою мірою та чи інша ознака, за якою ведеться селекція, залежить від спадкових особливостей тварини і якою – від середовища. Або, мовою популяційної генетики, виникає необхідність встановити частку генотипової обумовленості селекціонованої ознаки в загальній фенотиповій мінливості тварин.

Для цього необхідно визначити відносну частку впливу спадковості та факторів зовнішнього середовища на ступінь фенотипової мінливості ознаки в окремому стаді або популяції.

Взаємозв'язок між фенотипом і середовищем можна представити таким рівнянням:  $\Phi = \Gamma + \Pi$ .

Отже, наведене рівняння показує, що загальну фенотипову мінливість ознаки у стаді тварин можна розділити на дві частини. Перша – це та частина фенотипової мінливості, яка обумовлена генотиповою різноманітністю, а друга – факторами зовнішнього середовища.

Термін успадковуваність був уперше запропонований С. Лашем у 1939 році. Він запозичив його у С. Райта, який позначав цим терміном детермінацію ознаки спадковістю. Термін походить від англійського слова (*heritability*).

Дуже важливо зрозуміти, що успадковуваність – це явище популяційне, вона вивчається не в межах окремих особинах, а на рівні групи тварин. І тому його потрібно чітко відрізняти від таких понять як *спадковість* і *успадкування*.

**Успадковуваність** – це показник, який виражає частку генетичної різноманітності у загальній фенотиповій мінливості ознаки у групі, стаді, популяції тварин.

Це саме той показник, який дозволяє вирішити завдання, поставлене М.Ф.Івановим – “Шукайте кращі генотипи серед кращих фенотипів”. Звідси витікає, що той, хто приділяє всю увагу лише одній тварині, втрачає із виду всю популяцію. Слід зазначити, що чисто генетичної частки загальної мінливості немає, як немає мінливості, обумовленої тільки умовами зовнішнього

середовища. Будь-яка ознака є продуктом сукупного впливу спадковості та середовища.

Проте залежність різних ознак від дії середовища, як це вже наголошувалося, все ж таки відрізняється – мінливість кількісних ознак істотно залежить від середовища, мінливість же якісних ознак майже повністю контролюється спадковістю. Розмежувати з абсолютною точністю вплив спадковості та середовища на ту чи іншу ознаку окремо взятої особини неможливо.

Ці теоретичні судження і лягли в основу розробки поняття успадковуваності та тієї кількості методів визначення коефіцієнта успадковуваності, якими користуються в даний час для практичних цілей. Ці методи дозволяють вивчити не окремих тварин як таких, не окремі зміни кожної особини окремо, а мінливість тієї чи іншої *ознаки* у даної групи тварин.

Ось чому виявлення ступеня залежності між фенотипом і генотипом є однією з важливих завдань селекції, яка вирішується за допомогою визначення *коефіцієнта успадковуваності* ( $h^2$ ).

Розрізняють коефіцієнт успадковуваності у широкому та вузькому розумінні цього слова. Коефіцієнт успадковуваності у широкому розумінні враховує вплив усієї генетичної різноманітності, яка обумовлена як адитивною дією генів, так і впливом домінування, наддомінування, епістазу та плейотропії. Коефіцієнт успадковуваності у вузькому розумінні враховує генетичну різноманітність, яка обумовлена лише адитивною дією генів.

Методи визначення коефіцієнта успадковуваності, які наразі використовуються, можна поділити на дві групи:

- вирахування коефіцієнта успадковуваності на основі прямолінійної генетичної кореляції або регресії між продуктивністю батьків і нащадків;
- вирахування коефіцієнта успадковуваності на основі дисперсійного аналізу як силу впливу батьків на їхніх нащадків.

Методу вирахування коефіцієнта успадковуваності першої групи на основі прямолінійних кореляційних зв'язках між показниками продуктивності батьків та їхніх нащадків або показниками продуктивності матерів та їхніх дочок було засновано на схемі С. Райта, що з'єднує кореляціями чотири спадкові групи (генотип батьків, генотип дітей, фенотип батьків і фенотип дітей). Для вирахування коефіцієнтів успадковуваності за цим методом використовуються такі формули:

$$h^2 = 2 \cdot r_{M/D} \quad (47)$$

(подвоєний коефіцієнт кореляції між продуктивністю матерів (М) та їхніх дочок (Д);

$$h^2 = 2 \cdot r_{B/H} \quad (48)$$

(подвоєний коефіцієнт кореляції між продуктивністю батьків (Б) та їхніх нащадків (Н).

Однак при цьому методі об'єктивні дані про ступінь успадкованості ознак можна одержати лише за таких умов: популяція тварин є великою за об'ємом; популяція знаходиться в генетичній рівновазі, тобто між тваринами відбувається вільне парування; прояв кількісних ознак обумовлений лише адитивною дією генів.

Проте відомо, що у популяціях сільськогосподарських тварин вільне парування відсутнє, генетична рівновага не зберігається і має місце не лише адитивна дія генів, а й явище домінування, наддомінування, епістаз, тому прямолінійна кореляція між ознаками втрачається і коефіцієнт кореляції може бути більше одиниці. Крім того цим методом можна користуватися лише у тих випадках, коли інтенсивність добору серед батьків та їхніх нащадків є однаковою, а різниця у мінливості ознак батьків та нащадків є незначною. Величина коефіцієнта успадкованості, визначена цим методом, залежить від препотентності плідника.

При різній інтенсивності добору серед батьків і нащадків, коли вони різко відрізняються за ступнем мінливості ознак, доцільніше вираховувати коефіцієнт успадкованості на основі коефіцієнта регресії, використовуючи наступні формули:

$$h^2 = 2 \cdot R_{D/M} \quad (49)$$

(подвоєний коефіцієнт регресії між продуктивністю дочок та матерів);

$$h^2 = 2 \cdot R_{H/B} \quad (50)$$

(подвоєний коефіцієнт регресії між продуктивністю нащадків та батьків).

Зв'язок між генотипами різних родичів різний, тому і формули  $h^2$ , виражені через коефіцієнти кореляції ( $r$ ) або регресії ( $R$ ) між фенотипами, будуть різними.

Коефіцієнт успадковуваності можна вирахувати на основі подвоєного коефіцієнта кореляції між продуктивністю повних братів або сестер (повних сибсів) за формулою:

$$h^2 = 2 \cdot r_{C/C} \quad (51)$$

та за використання перемноженого на чотири коефіцієнта кореляції між показниками продуктивності напівсестер (напівсибсів) за батьком або матір'ю, використовуючи таку формулу:

$$h^2 = 4 \cdot r_{HC/HC} \quad (52)$$

Коефіцієнт успадковуваності може бути вирахований шляхом дисперсійного аналізу, який у порівнянні з кореляцією або регресією має більшу перевагу, оскільки дозволяє розчленувати всю суму ознак ( $C_y$ ), які визначають загальну фенотипову різноманітність на генотипову або факторіальну ( $C_x$ ) дисперсію і випадкову ( $C_z$ ), або паратипову дисперсію.

Співвідношення генотипової дисперсії до загальної і характеризує ступінь успадковуваності ознак, яке можна виразити формулою:

$$h^2 = \frac{C_x}{C_y} \quad (53)$$

Коефіцієнт успадковуваності – це величина відносна і виражається або у частках одиниці дробовим числом (0-1), або у відсотках (0-100).

Ступінь успадковуваності ознаки умовно ділять на високу ( $h^2 > 0,4$ ), середню ( $h^2 = 0,2—0,4$ ) і низьку ( $h^2 < 0,2$ ). Встановлено, що ефективність добору пропорційна ступеню успадковуваності. Тому масова селекція тварин на підвищення жирномолочності ( $h^2 = 0,5-0,6$ ) за інших рівних умов дає швидший ефект, чим селекція на високу молочність ( $h^2 = 0,2-0,3$ ).

На величину коефіцієнта успадковуваності впливає:

- ступінь гомозиготності тварин у популяції – чим вона вища, тим менша генотипова мінливість, тим вищий коефіцієнт успадковуваності ;
- ступінь генетичної однорідності популяції – чим однорідніша популяція, тим вище коефіцієнт успадковуваності ;
- генетична детермінація ознаки – чим меншим числом генів вона зумовлена, тим вищою буде частка генотипової мінливості;



– вплив зовнішніх умов (рівень годівлі, умови утримання, сезон року, вік тварин та ін.). Коливання в годівлі та низький її рівень не тільки знижує величину  $h^2$ , а й збільшує частку паратипової мінливості ознаки. Належні та стабільні умови підвищують рівень коефіцієнта успадковуваності, відповідно збільшується генетична зумовленість фенотипової мінливості ознаки, що дозволяє реально визначити ефективність селекції;

– мінливість ознаки – чим вища варіабельність ознаки, тим нижчий  $h^2$ . Це стосується таких кількісних ознак як: величина надою, вага курячих яєць, яйценосність, відгодівельні якості.

Яке ж практичне значення мають показники успадковуваності ознак для селекційної роботи?

1. Ефективність селекції в популяції тварин знаходиться в залежності від ступеня успадковуваності ознаки, що селекціонується. Це означає, що ознаки, які мають високу успадковуваність, можна швидше удосконалити *прямим добором*, тобто добором кращих тварин.

2. Користуючись коефіцієнтом успадковуваності, можна передбачити майбутню продуктивність потомства, тобто передбачити інтенсивність поліпшення селекціонованої ознаки в процесі селекції. Проте, для такого прогнозування, необхідно знати наскільки батьки, що використані для відтворення, перевершують за своєю продуктивністю середні показники тварин того стада, з якого вони були відібрані, тобто необхідно заздалегідь розрахувати селекційний диференціал.

3. З урахуванням коефіцієнта успадковуваності можливий критичніший підхід до добору тварин (особливо плідників) за родоводом, особливо за ознаками з високим коефіцієнтом успадковуваності.

5. Коефіцієнт успадковуваності допомагає правильно вибрати для даного стада метод селекції тварин за селекціонованою ознакою. Якщо селекція ведеться за ознакою з високим коефіцієнтом успадковуваності, то досить застосувати масовий добір (на перших етапах селекції він буде досить ефективним). Якщо ж ознака має низький коефіцієнт успадковуваності, то проводять надійніший поглиблений індивідуальний добір тварин з оцінкою плідників за якістю потомства.

### ***Задачі для самостійного розв'язування***

7.1.1. Розрахуйте статистичні показники  $M$ ,  $\sigma$ ,  $C_v$ ,  $m_M$ ,  $m_\sigma$ ,  $m_{C_v}$ ,  $t_M$ ,  $P$  методом малих вибірок користуючись, додатком 13 (кожен студент отримує

завдання згідно свого номеру за журналом).

7.1.2. Розрахуйте статистичні показники  $M$ ,  $\sigma$ ,  $C_v$ ,  $m_M$ ,  $m_\sigma$ ,  $m_{C_v}$ ,  $t_M$ ,  $P$  методом великих вибірок користуючись, додатком 14.

## 8. ПЕРЕВІРКА ГЕНЕТИЧНИХ ГІПОТЕЗ

### МЕТОДОМ ХІ-КВАДРАТ

Поряд із співставленням вибірових середніх у біології досить часто виникає необхідність порівнювати емпіричні (експериментальні) частоти з теоретичними, або очікуваними частотами, оцінювати результати експериментів стосовно якоїсь норми, стандарту чи гіпотези.

Іноді є необхідність встановлення якою мірою емпіричний варіаційний ряд відповідає нормальному розподілу і з яким ступенем ймовірності можливо вважати його нормальним, або порівнювати альтернативні розподіли ознак.

Метод  $\chi^2$  (хі-квадрат) запропонований К. Пірсоном для порівняння емпіричних рядів з теоретичними, виявлення між ними різниці та визначення її достовірності.

Величина  $\chi^2$  може приймати різні значення, від малих до найбільших, але вони завжди позитивні. Якщо  $\chi^2 = 0$ , то різниця між рядами, що порівнюються відсутня і нульова гіпотеза не відкидається. Проте якщо різниця між емпіричними і теоретичними частотами велика і величина  $\chi^2$  вказує на її достовірність, то це свідчить про неприйнятність нульової гіпотези, а значить про достовірність впливу фактора, який взятий до вивчення чи спостереження.

Величина  $\chi^2$  залежить від різниці між частотами та величини ступеня свободи.

Критерій відповідності  $\chi^2$  розраховують за формулою:

$$\chi^2 = \sum \frac{(O - E)^2}{E} \quad (54)$$

де:  $O$  – фактична кількість особин (емпірична частота);

$E$  – теоретично очікувана кількість особин (теоретична чи очікувана частота).

Для остаточного судження про відповідність, чи не відповідність фактично одержаних величин з теоретично очікуваними, одержане значення  $\chi^2$  (хі-квадрат) порівнюють з табличним з урахуванням числа ступенів свободи.

Число ступенів свободи ( $\nu$ ) для різних способів застосування  $\chi^2$  має особливості розрахунку. При гібридологічному аналізі число ступенів свободи

дорівнює числу фенотипів мінус одиниця. Так, у  $F_2$  моногібридного схрещування при повному домінуванні розщеплення 3:1 (2 фенотипи,  $v = 2-2=1$ ), при дигібридному 9:3:3:1 (4 фенотипи,  $v = 4-1=3$ ).

При вивченні поліморфізму білків та ферментів число ступенів свободи дорівнює числу генотипових класів мінус число алелів. При складанні чотири – або багатопільних решіток для визначення числа ступенів свободи користуються наступною формулою:

$$N = (l_x - 1) \cdot (l_y - 1) \quad (55)$$

де:  $l_x$  – число горизонтальних рядків;

$l_y$  – число вертикальних стовпчиків.

### Приклади розв'язування задач

При схрещуванні абердин-ангуських чорних комолих бугаїв з червоними коровами у  $F_1$  одержали усіх чорних комолих потомків, у  $F_2$  одержано: 640 потомків, які були чорними комолими – 363, чорними рогатими – 113, червоними комолими – 123 і червоними рогатими – 36.

Встановити, чи відповідає фактична кількість народжених потомків теоретично очікуваній?

Розрахунок починаємо із побудови спеціальної таблиці (табл. 10), у яку записують фактично одержані при схрещуванні дані ( $O$ ), теоретично очікувані ( $E$ ), арифметичні розрахунки: відхилення ( $O-E$ ), квадрати відхилень  $(O-E)^2$  та частки від ділення квадратів відхилень на теоретично очікувану величину за кожним класом. Сума часток за усіма класами і буде показником  $\chi^2$  (хі-квадрата).

Розрахунок теоретично очікуваної кількості особин різних фенотипів проводять таким чином. Знаходять загальну фактичну ( $O$ ) кількість особин різних фенотипів:  $363+113+123+35=640$ .

Теоретично очікувану ( $E$ ) кількість особин різних фенотипів розраховують виходячи з того, що у  $F_2$  при дигібридному схрещуванні згідно теорії ймовірності можливі 16 варіантів поєднання гамет, у результаті чого утворюється 4 фенотипи у співвідношенні 9:3:3:1.

Отже, для розрахунку теоретично очікуваної кількості особин різних фенотипів необхідно загальну кількість нащадків  $F_2$  помножити на відповідну частку кожного з фенотипів та ділити на число 16:

чорні комолі:  $640 \times 9 : 16 = 360$ ;

чорні рогаті:  $640 \times 3 : 16 = 120$ ;

червоні комолі:  $640 \times 3 : 16 = 120$ ;

червоні рогаті:  $640 \times 1 : 16 = 40$ .

Таблиця 10

Класи (фенотипи)	Число особин		Відхи- лення ( $O-E$ )	$(O-E)^2$	$\frac{(O-E)^2}{E}$
	фактичне ( $O$ )	очікуване ( $E$ )			
чорні комолі	363	360	8	64	$64:360=0,17$
чорні рогаті	113	120	7	49	$49:120=0,41$
червоні комолі	123	120	3	9	$9:120=0,08$
червоні рогаті	36	40	4	16	$16:40=0,40$
Всього	640	640	-	-	$\Sigma \frac{(O-E)^2}{E} = 1,06$

Після цього від фактичної кількості особин віднімаємо теоретично очікувану і знаходимо відхилення ( $O-E$ ), які необхідно піднести до квадрату. Квадрати відхилень за кожним із фенотипів ділимо на кількість теоретично очікуваних особин:  $64 : 360 = 0,17$ ;  $49:120=0,41$ ;  $9:120=0,08$ ;  $16:40=0,40$ .

Склавши суму усіх класів, одержуємо показник  $\chi^2$  (хі-квадрат):

$$\chi^2 = \Sigma \frac{(O-E)^2}{E} = 0,17 + 0,41 + 0,08 + 0,40 = 1,06$$

Щоб зробити висновок про відповідність чи не відповідність кількості фактично одержаних нащадків різних фенотипів з теоретично очікуваними,

розраховане значення  $\chi^2 = 1,06$  порівнюємо зі стандартним табличним значенням  $\chi^2$  при різних рівнях ймовірності (додаток 2).

Для цього необхідно визначити число ступенів свободи ( $\nu$ ). При гібридологічному аналізі число ступенів свободи дорівнює числу фенотипів мінус одиниця:

$$\nu = 4 - 1 = 3.$$

Стандартне значення  $\chi^2$  при  $\nu = 3$  для різних рівнів ймовірності становить:

$$P > 0,95 - 7,8; \quad P > 0,99 - 11,3; \quad P > 0,999 - 16,3.$$

Отримане значення  $\chi^2$  (1,06) істотно нижче табличного і підтверджує нульову гіпотезу.

**Висновок.** Фактичний розподіл кількості різних фенотипів одержаних від схрещування чорних комолих бугаїв абердин-ангуської породи з червоними рогатими коровами відповідає теоретично очікуваному з високим ступенем достовірності ( $P > 0,999$ ).

### Задачі для самостійного розв'язування

*Підготувати схему схрещування, визначити співвідношення фенотипів, розрахувати  $\chi^2$  і дати оцінку відповідності фактичного і теоретично очікуваного розщеплення щодо наведених нижче варіантів схрещування.*

8.1.1. Схрещували  $F_1$ , які походили від свиноматок великої білої та кнурів української степової породи. Одержано 680 поросят української степової і 204 великої білої.

8.1.2. Гетерозиготний помісний кнур білої масті, який походить від матки великої білої породи і беркширського кнура, схрещувався з чорними чистопородними матками беркширської породи. Одержано 83 чорних і 102 білих поросят.

8.1.3. Сірі барани спаровувалися з гетерозиготними чорними самками. Одержано 96 сірих і 87 чорних ягнят.

8.1.4. Схрещували комолих  $F_1$ , які походять від абердин-ангуських комолих бугаїв і рогатих швіцьких корів. Одержали 62 комолих і 24 рогатих телят.

8.1.5. Від схрещування чистопородного рогатого бугая герефордської породи з помісними комолими гетерозиготними коровами, які походять від комолих абердин-ангуських бугаїв і рогатих швіцьких корів, одержано 24 рогатих і 32

комолих телят.

8.1.6. Схрещували між собою  $F_1$  від чорного гемпширського кнура і маток породи дюррок. Одержано 93 поросят чорної масті і 37 червоної.

8.1.7. Від схрещування гетерозиготних курок, які мали трояндоподібний гребінь, з такими ж півнями одержано 215 курчат з трояндоподібним і 64 курчат з листоподібним гребенем.

8.1.8. Схрещували між собою  $F_1$  від довгошерстих ангорських самок і самців породи шиншила. Одержано 86 короткошерстих і 34 довгошерстих кроленят.

8.1.9. Від схрещування гетерозиготних сірих кролів між собою одержано 240 сірих і 78 білих кроленят.

8.1.10. Від схрещування помісних корів і бугаїв червоної масті з білими плямами, які походять від симентальських маток і бугаїв червоної степової породи, одержано 27 червоно-рябих телят, 43 червоних з білими плямами і 20 червоних телят.

8.1.11. Від схрещування гетерозиготних червоно-рябих корів і бугаїв голштинської породи між собою одержали 73 чорно-рябих і 22 червоно-рябих телят.

8.1.12. Від схрещування чистопородних чалих шортгорнських корів з білими шортгорнськими бугаями одержано 37 чалих і 26 білих телят.

8.1.13. При схрещуванні довговухих овець з безвухими баранами всі потомки в  $F_1$  мають короткі вуха. Від схрещування  $F_1$  між собою одержано 37 безвухих, 50 довговухих та 95 коротковухих ягнят.

8.1.14. Дикі кролі мають підшкірний жир білого кольору, а деякі свійські породи – жовтого. Від схрещування самців  $F_1$  зі свійськими кролематками 185 нащадків мали білий жир, а 209 – жовтий.

8.1.15. Від схрещування між собою гетерозиготних кролів з жовтим підшкірним жиром одержано 215 кроленят з жовтим жиром і 87 – з білим.

8.1.16. При схрещуванні норок з темним забарвленням хутра з білими норками одержують норок кохінурового забарвлення (світлі з темним хрестом на спині). Від схрещування кохінурових норок  $F_1$  з білими норками одержано 52 білих і 68 кохінурових нащадків.

### ***Контрольні запитання***

1. Дайте визначення термінів біометрія, генеральна сукупність, вибірка, середнє арифметичне значення, помилки репрезентативності.
2. Поясніть суть біометричних методів оцінки мінливості ознак.
3. Поясніть суть методу Хі-квадрат ( $\chi^2$ ).
4. Опишіть практичне значення використання методу Хі-квадрат.

### **Бібліографічний список**

1. **Близнюченко О.Г.** Біометрія: Монографія / Близнюченко О.Г. – Полтава: Редакційно-видавничий відділ “Тетра” Полтавської державної аграрної академії, 2003. – 346 с.
2. **Засуха В.А., Лисенко В.П.** прикладна математика – К.: Арістей, 2005, 342с.
3. **Меркурьєва Е.К.** Биометрия в селекции и генетике сельскохозяйственных животных / Е.К. Меркурьєва. – М. : Колос, 1969. – С. 120 – 124.
4. **Меркурьєва Е.К.** Генетика с основами биометрии / Е.К. Меркурьєва, Г.Н. Шангин-Березовский. – М. : Колос, 1983. – С. 243 – 260.
5. **Петухов В.Л.** Ветеринарная генетика с основами вариационной статистики / В.Л. Петухов, А.И. Жигачев, Г.А. Назаров. – М. : Агропромиздат, 1985. – С. 157 – 161.
6. **Проценко Н.Е.** Генетика с основами биометрии / Н.Е. Проценко. – Киев: Изд-во УСХА, 1991. – С. 92 – 95.



**Стандартні значення критерію достовірності  
(хі-квадрат)**

v	Рівні ймовірності P			v	Рівні ймовірності P		
	0,95	0,99	0,999		0,95	0,99	0,999
1	3,8	6,6	10,8	26	38,9	45,6	54,1
2	6,0	9,2	13,8	27	40,1	47,0	55,5
3	7,8	11,3	16,3	28	41,3	48,3	56,9
4	9,5	13,3	18,5	29	42,6	49,6	58,3
5	11,1	15,1	20,5	30	43,8	50,9	59,7
6	12,6	16,8	22,5	32	46,2	53,5	62,4
7	14,1	18,5	24,3	34	48,6	56,0	65,2
8	15,5	20,1	26,1	36	51,0	58,6	67,9
9	16,9	21,7	27,9	38	53,4	61,6	70,7
10	18,3	23,2	29,6	40	55,8	63,7	73,4
11	19,7	24,7	31,3	42	58,1	66,2	76,1
12	21,0	26,2	32,9	44	60,5	68,7	78,7
13	22,4	27,7	34,5	46	62,8	71,2	81,4
14	23,7	29,1	36,1	48	65,2	73,7	84,0
15	25,0	30,6	37,7	50	67,5	76,2	86,7
16	26,3	32,0	39,3	55	73,3	82,3	93,2
17	27,6	33,4	40,8	60	79,1	88,4	99,6
18	28,9	34,8	42,3	65	84,8	94,4	106,0
19	30,1	36,3	43,8	70	90,5	100,4	112,3
20	31,4	37,6	45,3	75	96,2	106,4	118,5
21	32,7	38,9	46,8	80	101,9	112,3	124,8
22	33,9	40,3	48,3	85	107,5	118,2	131,0
23	35,2	41,6	49,7	90	113,1	124,1	137,1
24	36,4	43,0	51,2	95	118,7	130,0	143,3
25	37,7	44,3	52,6	100	124,3	135,8	149,4

**Стандартні значення критерію достовірності  $t_d$   
для статистичних показників (за Стьюдентом)**

Число ступенів свободи, $\nu$	Ймовірність (P)				
	<b>0,90 (0,10)</b>	<b>0,95 (0,05)</b>	<b>0,98 (0,02)</b>	<b>0,99 (0,01)</b>	<b>0,999 (0,001)</b>
1	6,31	12,7	31,82	63,66	-
2	2,92	4,30	6,97	9,93	31,60
3	2,35	3,18	4,54	5,84	12,94
4	2,13	2,78	3,75	4,60	8,61
5	2,02	2,57	3,37	4,03	6,86
6	1,94	2,45	3,14	3,71	5,96
7	1,90	2,37	3,00	3,50	5,41
8	1,86	2,31	2,90	3,36	5,04
9	1,83	2,26	2,82	3,25	4,78
10	1,81	2,23	2,76	3,17	4,50
11	1,80	2,20	2,72	3,11	4,44
12	1,78	2,18	2,68	3,06	4,32
13	1,77	2,16	2,65	3,01	4,22
14	1,76	2,15	2,62	2,98	4,14
15	1,75	2,13	2,60	2,95	4,07
16	1,75	2,12	2,58	2,92	4,02
17	1,74	2,11	2,57	2,90	3,97
18	1,73	2,10	2,55	2,88	3,92
19	1,73	2,09	2,54	2,86	3,88
20	1,73	2,09	2,53	2,85	3,85
21	1,72	2,08	2,52	2,83	3,82
22	1,72	2,07	2,51	2,82	3,79
23	1,71	2,07	2,50	2,81	3,77
24	1,71	2,06	2,49	2,80	3,75
25	1,71	2,06	2,49	2,79	3,73
26	1,71	2,06	2,48	2,78	3,71
27	1,70	2,05	2,47	2,77	3,69
28	1,70	2,05	2,47	2,76	3,67
29	1,70	2,05	2,46	2,76	3,66
30	1,70	2,04	2,46	2,75	3,65
31 і до $\infty$	1,64	1,96	2,33	2,58	3,29

Додаток 3

Стандартні значення критерію Фішера ( $P>0.95$ ;  $P>0.99$ ;  $P>0.999$ )

$v_2$	$v_1$												
	1	2	3	4	5	6	7	8	9	10	50	100	500
3	167,5	148,5	141,1	137,1	134,6	132,9	131,8	130,6	130,0	129,5	125,0	124,4	123,8
	34,1	30,8	29,5	28,7	28,2	27,9	27,7	27,7	27,4	27,5	26,4	26,2	26,1
	10,1	9,6	9,3	9,1	9,0	8,9	8,8	8,8	8,8	8,8	8,6	8,6	8,5
4	74,1	61,2	56,1	53,4	51,7	50,5	49,8	49,0	48,6	48,2	45,2	44,7	44,3
	21,2	18,8	16,7	16,0	15,5	15,2	15,0	14,8	14,7	14,7	13,7	13,5	13,5
	7,7	6,9	6,6	6,4	6,3	6,2	6,1	6,0	6,0	6,0	5,7	5,7	5,6
5	47,0	36,6	33,2	31,1	29,8	28,8	28,2	17,6	27,3	27,0	24,6	24,3	23,8
	16,3	13,3	12,1	11,4	11,0	10,7	10,5	10,3	10,2	10,1	9,1	9,1	9,0
	6,6	5,8	5,4	5,2	5,1	5,0	4,9	4,8	4,8	4,7	4,4	4,4	4,4
6	35,5	27,0	23,7	21,9	20,8	20,0	19,5	19,0	18,8	18,5	16,5	16,2	15,9
	13,4	10,9	9,8	9,2	8,8	8,5	8,3	8,1	8,0	7,9	7,1	7,0	6,9
	6,0	5,1	4,8	4,5	4,4	4,3	4,2	4,1	4,1	4,1	3,8	3,7	3,7
7	29,2	21,7	18,8	17,2	16,2	15,5	15,1	14,6	14,4	14,2	12,3	12,1	11,8
	12,3	9,6	8,5	7,9	7,5	7,2	7,0	6,8	6,7	6,6	5,9	5,8	5,7
	5,6	4,7	4,4	4,1	4,0	3,9	3,8	3,7	3,7	3,6	3,3	3,3	3,2
8	25,4	18,5	15,8	14,4	13,5	12,9	12,5	12,0	11,8	11,6	10,0	9,7	9,5
	11,3	8,7	7,6	7,0	6,6	6,4	6,2	6,0	5,9	5,8	5,1	5,0	4,9
	5,3	4,6	4,1	3,8	3,7	3,6	3,5	3,4	3,4	3,3	3,0	3,0	2,9
9	22,9	16,4	13,9	12,6	11,7	11,1	10,8	10,4	10,2	10,0	8,4	8,1	7,9
	10,6	8,0	7,0	6,4	6,1	5,8	5,6	5,5	5,4	5,3	4,5	4,4	4,3
	5,1	4,3	3,6	3,6	3,5	3,4	3,3	3,2	3,2	3,1	2,8	2,8	2,7
10	21,0	14,9	12,3	11,3	10,5	9,9	9,6	9,2	9,0	8,9	7,3	7,1	6,9
	10,0	7,9	6,6	6,0	5,6	5,4	5,2	5,1	5,0	4,9	4,1	4,0	3,9
	5,0	4,1	3,7	3,5	3,3	3,2	3,1	3,1	3,0	3,0	2,6	2,6	2,6
50	12,2	8,0	6,4	5,4	4,9	4,5	4,3	4,0	3,9	3,7	2,6	2,3	2,1
	7,2	5,1	4,2	3,7	3,4	3,2	3,0	2,9	2,8	2,7	1,9	1,8	1,7
	4,0	3,2	2,8	2,6	2,4	2,3	2,2	2,1	2,1	2,0	1,6	1,5	1,5
100	11,5	7,4	5,9	5,0	4,1	4,1	3,9	3,7	3,6	3,4	2,2	1,9	1,7
	6,9	4,8	4,0	3,5	3,2	3,0	2,8	2,7	2,6	2,5	1,7	1,6	1,5
	3,9	3,1	2,7	2,5	2,3	2,2	2,1	2,0	2,0	1,9	1,5	1,4	1,3
1000	10,9	7,0	5,5	4,7	4,2	3,8	3,6	3,4	3,3	3,1	1,8	1,6	1,3
	6,7	4,6	3,8	3,4	3,3	3,0	2,8	2,7	2,5	2,4	1,6	1,4	1,2
	3,8	3,0	2,6	2,4	2,2	2,1	2,0	1,9	1,9	1,8	1,4	1,3	1,1

### Додаток 4. Завдання для оцінки мінливості кількісних ознак тварин методом малих вибірок

№ п.п.	Вид продуктивності	Показники продуктивність с.-г. тварин																	
		n1	n2	n3	n4	n5	n 6	n 7	n 8	n 9	n 10	n11	n 12	n 1	n 14	n 15	n 16	n17	n18
1	Надій за лактацію, кг	3500	3600	3700	3800	3900	4000	4300	4600	5000	5200	3900	3700	5000	3600	3700	4000	3900	4600
2	Добовий надій, кг	8	9	10	10	11	12	13	15	16	18	11	10	16	9	10	12	11	15
3	Тривалість сервіс-період, днів	80	60	70	74	64	84	45	85	65	55	64	70	65	60	70	84	64	85
4	Тривалість лактації, днів	305	280	320	295	310	316	310	300	295	310	310	320	295	280	320	316	310	300
5	Жива маса, первісток кг	350	340	365	375	380	395	397	405	350	340	380	365	350	340	365	395	380	405
6	Вік першого отелення, міс	27	28	29	30	29	28	27	29	28	27	29	29	28	28	29	28	29	29
7	Багатоплідність у свиней, голів	9	10	11	14	14	13	10	12	14	11	14	11	14	10	11	13	14	12
8	Вік досягнення 100 кг, діб	163	160	161	162	172	175	180	170	165	164	172	161	165	160	161	175	172	170
9	Жива маса дорослих свиноматок	237	252	240	250	240	245	247	239	238	246	240	240	238	252	240	245	240	239
10	Маса гнізда в два місяці, кг	224	222	214	215	220	225	225	215	216	219	220	214	216	222	214	225	220	215
11	Несучість у курей, яєць	220	230	236	225	234	236	237	225	228	231	234	236	228	230	236	236	234	225
12	Настриг чистої вовни, кг	4,0	4,1	4,2	3,9	4,1	4,1	5,0	4,1	5,1	3,8	4,1	4,2	5,1	4,1	4,2	4,1	4,1	4,1
13	Настриг немитої вовни, кг	6,9	7,1	6,3	6,8	7,5	8,1	7,5	7,6	7,3	7,4	7,5	6,3	7,3	7,1	6,3	8,1	7,5	7,6
14	Жива маса дорослих корів, кг	540	545	560	570	580	595	610	<del>765</del> 615	620	610	580	560	500	545	560	595	580	615

15	Жива маса дорослих бугаїв, кг	950	850	975	980	960	990	1005	1150	980	950	850	950	850	980	990	990	1005	975
16	Жива маса телят при народженні, кг	25	30	27	28	26	31	33	27	30	32	30	25	30	28	31	31	33	27
17	Молочність м'ясних корів, кг (6 міс)	160	186	170	175	180	187	175	180	170	175	186	160	186	175	187	187	175	170
18	Жива маса баранів-плідників, кг	85	90	86	91	92	93	94	95	91	90	90	85	90	91	93	93	94	86
19	Жива маса ягнят при відлученні, кг	25	30	26	27	28	29	26	25	27	29	30	25	30	27	29	29	26	26
20	Висота в холці жеребців, см	159	158	159	160	161	158	161	162	161	159	158	159	158	160	158	158	161	159
21	Еритроцити крові	6,3	6,5	7,0	6,7	7,1	6,0	5,5	6,1	5,9	6,3	6,5	6,3	6,5	6,7	6,0	6,0	5,5	7,0

## Додаток 5. Завдання для оцінки мінливості кількісних ознак с.-г. тварин методом великих вибірок

№ п/п	Вид продуктивності	Показники продуктивності с.-г. тварин																	
		n1	n2	n3	n4	n5	n 6	n 7	n 8	n 9	n 10	n11	n 12	n 13	n 14	n 15	n 16	n 17	n 18
1	Надій за лактацію, кг	3500	3600	3700	3800	3900	4000	4300	4600	5000	5200	3900	3700	5000	3600	3700	4000	3900	4600
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		4300	4600	5000	3700	5000	4000	4300	3700	5000	3700	3900	4600	3700	3800	3900	3700	3900	4600
2	Добовий надій, кг	8	9	10	10	11	12	13	15	16	18	11	10	16	9	10	12	11	15
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		9	10	10	11	16	9	10	12	15	16	18	15	15	16	16	16	9	10
3	Тривалість сервіс-періоду, днів	80	60	70	74	64	84	45	85	65	55	64	70	65	60	70	84	64	85
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		45	85	65	55	64	70	65	60	70	45	85	64	55	64	70	65	60	85
4	Тривалість лактації, днів	305	280	320	295	310	316	310	300	295	310	310	320	295	280	320	316	310	300
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		295	310	320	295	310	316	310	295	280	310	310	320	300	300	320	316	310	300
5	Жива маса,первісток кг	350	340	365	375	380	395	397	405	350	340	380	365	350	340	365	395	380	405
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		340	365	395	380	405	350	340	365	350	380	395	380	395	380	395	397	405	350
6	Вік першого отелення, міс	27	28	29	30	29	28	27	29	28	27	29	29	28	28	29	28	29	29
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		28	29	28	29	30	28	27	27	28	29	29	30	29	30	29	28	27	29
7	Багатоплідність у свиней, голів	9	10	11	14	14	13	10	12	14	11	14	11	14	10	11	13	14	12
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		13	14	12	13	14	12	9	10	11	14	13	14	12	9	10	11	14	9
8	Вік досягнення 100 кг, діб	163	160	161	162	172	175	180	170	165	164	172	161	165	160	161	175	172	170
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		165	164	172	161	162	165	160	161	162	172	175	165	160	161	161	162	161	161
9	Жива маса дорослих свиноматок	237	252	240	250	240	245	247	239	238	246	240	240	238	252	240	245	240	239
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		240	245	240	239	245	240	239	246	240	240	245	247	239	240	240	239	238	246
10	Маса гнізда в два місяці, кг	224	222	214	215	220	225	225	215	216	219	220	214	216	222	214	225	220	215
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		215	216	219	220	220	225	215	216	219	222	214	222	214	225	220	215	225	224
11	Несучість у курей, яєць	220	230	236	225	234	236	237	225	228	231	234	236	228	230	236	236	234	225
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		225	228	231	237	225	228	231	220	230	236	237	225	228	231	237	225	228	231

12	Настриг чистої вовни, кг	4,0	4,1	4,2	3,9	4,1	4,1	5,0	4,1	5,1	3,8	4,1	4,2	5,1	4,1	4,2	4,1	4,1	4,1
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		4,1	5,0	4,1	5,1	3,8	4,1	4,2	3,9	4,1	4,1	5,0	4,1	4,2	3,9	4,2	5,1	4,1	4,2
13	Настриг немітої вовни, кг	6,9	7,1	6,3	6,8	7,5	8,1	7,5	7,6	7,3	7,4	7,5	6,3	7,3	7,1	6,3	8,1	7,5	7,6
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		6,3	8,1	7,5	7,6	7,5	6,3	7,3	6,3	8,1	7,5	7,6	6,3	6,8	7,5	8,1	7,5	7,6	7,3
14	Жива маса дорослих корів, кг	540	545	560	570	580	595	610	615	620	610	580	560	500	545	560	595	580	615
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		570	580	595	595	580	615	580	595	610	615	580	500	545	560	580	595	610	615
15	Жива маса дорослих бугаїв, кг	950	850	975	980	960	990	1005	1150	980	950	850	950	850	980	990	990	1005	975
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		1005	1150	980	950	850	980	950	850	980	990	1005	975	850	975	980	960	990	1005
16	Жива маса телят при народженні, кг	25	30	27	28	26	31	33	27	30	32	30	25	30	28	31	31	33	27
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		28	26	31	33	27	30	32	30	25	32	33	27	30	28	26	31	33	27
17	Молочність м'ясних корів, кг (6 міс)	160	186	170	175	180	187	175	180	170	175	186	160	186	175	187	187	175	170
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		180	187	175	180	170	184	187	175	170	187	187	175	170	187	186	170	175	180
18	Жива маса баранів-плідників, кг	85	90	86	91	92	93	94	95	91	90	90	85	90	91	93	93	94	86
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		86	91	92	93	94	95	91	90	85	90	86	93	94	86	84	85	90	86
19	Жива маса ягнят при відлученні, кг	25	30	26	27	28	29	26	25	27	29	30	25	30	27	29	29	26	26
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		27	29	25	30	26	26	26	27	26	29	29	26	30	25	30	25	27	29
20	Висота в холці жеребців, см	159	158	159	160	161	158	161	162	161	159	158	159	158	160	158	158	161	159
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		161	162	161	159	160	158	158	160	158	158	159	160	161	159	159	158	159	160

21	Добовий надій, кг	11	9	10	10	11	12	13	15	16	19	11	10	16	9	10	12	11	15
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		9	10	10	11	16	9	10	12	15	16	19	15	15	16	16	16	9	10
22	Тривалість сервіс-періоду, днів	55	60	70	74	64	84	45	85	65	55	64	70	65	60	70	84	64	87
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		45	85	65	55	64	70	65	60	70	45	85	64	55	64	70	65	60	85
23	Тривалість лактації, днів	310	280	320	295	310	316	310	300	295	310	310	325	295	280	320	316	310	300
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		295	310	320	295	310	316	310	295	280	310	310	320	300	300	320	316	310	300
24	Жива маса,первісток кг	345	340	365	375	380	395	397	407	350	340	380	365	350	340	365	395	380	405
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		340	365	395	380	405	350	340	365	350	380	395	380	395	380	395	397	405	350
25	Вік першого отелення, міс	29	28	29	31	29	28	27	29	28	27	29	29	28	28	29	28	29	29
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		28	29	28	29	30	28	27	27	28	29	29	30	29	30	29	28	27	29
26	Багатоплідність у свиней, голів	9	10	11	14	15	13	10	12	14	11	14	11	14	10	11	13	14	12
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		13	14	12	13	14	12	9	10	11	14	13	14	12	9	10	11	14	9
27	Вік досягнення 100 кг, діб	164	160	161	162	172	175	180	170	165	164	174	161	165	160	161	175	172	170
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		165	164	172	161	162	165	160	161	162	172	175	165	160	161	161	162	161	161
28	Жива маса дорослих свиноматок	238	252	240	250	240	245	247	239	238	246	240	240	238	252	240	248	240	239
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		239	245	240	239	245	240	239	246	240	240	245	247	239	240	240	239	238	246
29	Маса гнізда в два місяці, кг	224	222	214	215	220	230	225	215	216	219	220	214	216	222	214	225	220	215
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		215	216	219	220	220	225	215	216	219	222	214	222	214	225	220	215	225	224
30	Несучість у курей, яєць	236	230	236	225	234	236	237	225	228	231	234	238	228	230	236	236	234	225
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		225	228	231	237	225	228	231	220	230	236	237	225	228	231	237	225	228	231
31	Настриг чистої вовни, кг	4,0	4,1	4,2	3,9	4,1	4,1	5,0	4,1 80	5,2	3,8	4,1	4,2	5,1	4,1	4,2	4,1	4,1	4,1
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		4,1	5,0	4,1	5,1	3,8	4,1	4,2	3,9	4,1	4,1	5,0	4,1	4,2	3,9	4,2	5,1	4,1	4,2
32	Настриг немитої	6,9	7,1	6,3	6,8	7,5	8,3	7,5	7,6	7,3	7,4	7,5	6,3	7,3	7,1	6,3	8,1	7,5	7,6
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36



	вовни, кг	6,3	8,1	7,5	7,6	7,5	6,3	7,3	6,3	8,1	7,5	7,6	6,3	6,8	7,5	8,1	7,5	7,6	7,3
33	Жива маса дорослих корів, кг	580	545	560	570	580	595	610	617	620	610	580	560	500	545	560	595	580	615
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		570	580	595	595	580	615	580	595	610	615	580	500	545	560	580	595	610	615
34	Жива маса дорослих бугаїв, кг	980	850	975	980	960	990	1005	1155	980	950	845	950	850	980	990	990	1005	975
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		1005	1150	980	950	850	980	950	850	980	990	1005	975	850	975	980	960	990	1005
35	Жива маса телят при народженні, кг	25	30	27	28	26	31	33	27	30	33	30	25	30	28	31	31	33	27
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		28	26	31	33	27	30	32	30	25	32	33	27	30	28	26	31	33	27
36	Молочність м'ясних корів, кг (6 міс)	175	186	170	175	180	190	175	180	170	175	186	160	186	175	187	187	175	170
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		180	187	175	180	170	184	187	175	170	187	187	175	170	187	186	170	175	180
37	Жива маса баранів- плідників, кг	93	90	86	91	92	93	94	95	91	90	90	85	90	91	93	93	95	86
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		86	91	92	93	94	95	91	90	85	90	86	93	94	86	84	85	90	86
38	Жива маса ягнят при відлученні. кг	27	30	26	27	28	29	26	25	27	29	32	25	30	27	29	29	26	26
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		27	29	25	30	26	26	26	27	26	29	29	26	30	25	30	25	27	29
39	Висота в холці жеребців, см	159	158	159	160	161	158	161	162	161	159	158	159	158	160	158	158	163	159
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		161	162	161	159	160	158	158	160	158	158	159	160	161	159	159	158	159	160
40	Еритроцити крові	5,5	6,5	7,0	6,7	7,3	6,0	5,5	6,1	5,9	6,3	6,5	6,3	6,5	6,7	6,0	6,0	5,5	7,0
		n19	n20	n 21	n 22	n 23	n 24	n 25	n 26	n 27	n 28	n 29	n 30	n 31	n 32	n 33	n 34	n 35	n 36
		6,7	6,0	5,5	7,0	5,5	6,1	5,9	6,5	5,5	6,1	5,9	7,0	6,7	7,1	6,3	6,5	7,0	6,7