
МАТЕМАТИЧНІ МЕТОДИ В ХЕМІЇ

УДК 519.24/27:543/545

Г.О. Сіренко¹, Л.Я. Мідак¹, О.Г. Сіренко²

Методи лінійної множинної кореляції та регресії в хемічному матеріалознавстві

¹Прикарпатський національний університет імені Василя Стефаника,
вул. Шевченка, 57, м. Івано-Франківськ, 76025, Україна

²Національний ботанічний сад імені М.М. Гришка НАНУ,
вул. Тімірязєвська, 1, м. Київ, 01014, Україна

Приведені основні означення кореляційної аналізи, розрахункові формули коефіцієнта лінійної множинної кореляції, розглянуто процедури одержання рівнянь регресії та приклади застосування кореляційної аналізи в хемічній технології.

Ключові слова: кореляція, значущість, лінійний зв'язок, коефіцієнт кореляції, математична обробка результатів, хемія, хемічна технологія.

H.O. Sirenko¹, L.Ya. Midak¹, O.H. Sirenko²

Methods of Linear Plural Correlation and Regression in Chemical Materials Science

¹Vasyl Stefanyk' Precarpathian National University,
57, Shevchenko Str., Ivano-Frankivsk, 76025, Ukraine

²M.M. Gryshko National Botanic Garden NASU,
1, Timiryazevska Str., Kyiv, 01014, Ukraine

Basic determinations of cross-correlation analysis and calculation formulas of coefficient of linear plural correlation are resulted; procedures of receipt of equalizations of regression and the examples of application of cross-correlation analyses in chemical technology are considered.

Key words: correlation, meaningfulness, linear link, coefficient of correlation, mathematical treatment of results, chemistry, chemical technology.

Стаття поступила до редакції 28.03.2011; прийнята до друку 20.04.2011.

Вступ

Метою більшості досліджень в хемії та хемічній технології є вирішення складних багатофакторних експериментальних завдань, які пов'язані зі встановленням надійних лінійних і нелінійних зв'язків між випадковими величинами, пошуком оптимальних рішень якості матеріалів та оптимальних умов проведення хеміко-технологічних процесів.

Існують два різних підходи до вирішення таких завдань. Традиційні методи досліджень в хемії та хемічній технології пов'язані з «пасивним» експериментом, який потребує значних витрат часу та фінансів. У такому експерименті вирішенню екстремальних завдань передують всебічне дослідження механізму процесу

та властивостей речовини. «Пасивний» експеримент пов'язаний з почерговим варіюванням окремих змінних при сталих значеннях інших. Базуючись на результатах такого дослідження, можна створити теорію процесу, за допомогою якої можна вирішувати експериментальні завдання. Але точність і надійність таких результатів низька. Та й системи, які належить описати теоретично та оптимізувати є багатофакторними, багаторівневими та виявляються настільки складними, що не підлягають теоретичному вивченню у прийнятні терміни. Окрім того, у більшості випадків завдання вирішуються експериментально при неповному знанні механізмів процесів та явищ. Методологія знаходження таких рішень залишається неформалізованою.

У другому методі використовують теорію «активного» експерименту, що дозволяє вибрати оптимальну стратегію дослідження при неповних знаннях про об'єкт дослідження та багатофакторному завданні. При цьому на кожному етапі дослідження використовують математичне планування експерименту. Математичне планування експерименту базується на загальнометодичних концепціях: дисперсійної аналізи; регресійної аналізи; кореляційної аналізи; рандомізації (надання випадкового характеру реалізації дослідів та їх повторень); послідовного експерименту; оптимального використання факторного простору; компактності інформації; статистичних оцінок тощо.

Мета роботи полягала в узагальненні, систематизації матеріалу для створення навчально-методичного посібника «Математичні методи в хемії та хемічній технології», розділу «Кореляційна аналіза: методи лінійної множинної кореляції та регресії» для студентів спеціальності «Хемія» освітньо-кваліфікаційних рівнів магістра та спеціаліста.

I. Теоретична частина

Теоретичні основи кореляційної та регресійної аналізи викладено в [1–26].

1. У випадку k-змінних рівняння множинної регресії у натуральній шкалі має вигляд:

$$\bar{y}_1(2,3,\dots,i,\dots,k) = b_1 + b_2x_2 + b_3x_3 + \dots + b_ix_i + \dots + b_kx_k, \quad (1)$$

де $\bar{y}_1(2,3,\dots,i,\dots,k)$ – умовне середнє значення залежної величини y_1 , яке відповідає певним значенням незалежних величин $x_2, x_3, \dots, x_i, \dots, x_k$.

2. Проведемо процедуру нормування: перейдемо до нової випадкової нормалізованої змінної t_i :

$$\left. \begin{aligned} t_i &= \frac{y_j - \bar{y}_i}{S_{y_i}}; \\ t_j &= \frac{x_j - \bar{x}_i}{S_{x_i}} \end{aligned} \right\} \quad (2)$$

тоді, всі величини y_i та всі їх залежності знайдуть вираження у стандартній шкалі.

3. Таким чином, рівняння регресії (1) в нормованому вигляді набуде форми:

$$\bar{t}_1(2,3,\dots,i,\dots,k) = \beta_2t_2 + \beta_3t_3 + \dots + \beta_it_i + \dots + \beta_kt_k, \quad (3)$$

де $\bar{t}_1(2,3,\dots,i,\dots,k)$ – умовне середнє значення нормованої (стандартної) залежної величини t_1 , яке відповідає певним значенням нормованих незалежних величин $t_2, t_3, \dots, t_i, \dots, t_k$; $t_2, t_3, \dots, t_i, \dots, t_k$ – значення нормованих (стандартних) незалежних величин $y_2, y_3, \dots, y_i, \dots, y_k$; $\beta_2, \beta_3, \dots, \beta_i, \dots, \beta_k$ – нормовані (стандартні) коефіцієнти множинної регресії за рівнянням (3).

4. Нормовані (стандартні) коефіцієнти множинної регресії визначимо за системою лінійних рівнянь:

$$\left. \begin{aligned} r_{1,2} &= \beta_2r_{22} + \beta_3r_{32} + \beta_4r_{42} + \dots + \beta_ir_{i2} + \dots + \beta_kr_{k2} \\ r_{1,3} &= \beta_2r_{23} + \beta_3r_{33} + \beta_4r_{43} + \dots + \beta_ir_{i3} + \dots + \beta_kr_{k3} \\ r_{1,4} &= \beta_2r_{24} + \beta_3r_{34} + \beta_4r_{44} + \dots + \beta_ir_{i4} + \dots + \beta_kr_{k4} \\ &\dots\dots\dots \\ &\dots\dots\dots \\ r_{1,i} &= \beta_2r_{2i} + \beta_3r_{3i} + \beta_4r_{4i} + \dots + \beta_ir_{ii} + \dots + \beta_kr_{ki} \\ &\dots\dots\dots \\ &\dots\dots\dots \\ r_{1,k} &= \beta_2r_{2k} + \beta_3r_{3k} + \beta_4r_{4k} + \dots + \beta_ir_{ik} + \dots + \beta_kr_{kk} \end{aligned} \right\} \quad (4)$$

де $r_{12}, r_{13}, r_{14}, \dots, r_{1i}, \dots, r_{1k}, \dots, r_{22}, r_{23}, r_{2i}, r_{2k}, r_{32}, r_{33}, r_{3i}, r_{3k}, r_{42}, r_{43}, \dots, r_{ik}, \dots, r_{kk}$ – коефіцієнти парної лінійної кореляції між змінними $y_1 \sim x_2$; $y_1 \sim x_3$; $x_2 \sim x_3$; $x_2 \sim x_4$; \dots ; $x_i \sim x_j$; $x_i \sim x_k$; \dots ; $x_k \sim x_k$.

5. Щільність зв'язку змінної y_1 із сукупністю змінних $x_2, x_3, \dots, x_i, \dots, x_k$ у випадку лінійної множинної кореляції визначається за коефіцієнтом множинної кореляції:

$$r_1(2,3,\dots,i,\dots,k) = \sqrt{\beta_2r_{12} + \beta_3r_{13} + \dots + \beta_ir_{1i} + \dots + \beta_kr_{1k}}. \quad (5)$$

6. Коефіцієнт множинної кореляції після корекції (врахування числа коефіцієнтів рівняння (1) – числа параметрів рівняння регресії в натуральній шкалі дорівнює:

$$\bar{r}_1(2,3,\dots,i,\dots,k) = \sqrt{1 - [1 - r_1^2(2,3,\dots,i,\dots,k)] \cdot \left(\frac{N-1}{N-k} \right)}, \quad (6)$$

де N – число спостережень;

k – число параметрів (коефіцієнтів) математичної моделі (1).

7. Розрахунок коефіцієнтів математичної моделі (1) в натуральній шкалі:

$$b_i = \beta_1 \frac{S_{y1}}{S_i}, \text{ де } i = 2, 3, \dots, i, \dots, k; \quad (7)$$

$$b_1 = \bar{y}_1 - [b_2\bar{x}_2 + b_3\bar{x}_3 + \dots + b_i\bar{x}_i + \dots + b_k\bar{x}_k], \quad (8)$$

де $\bar{y}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_i, \dots, \bar{x}_k$ – середнє значення відповідних величин.

8. Оцінка середньої квадратичної помилки розрахунку величини $\bar{y}_1(2,3,\dots,i,\dots,k)$ в рівнянні моделі (1) дорівнює:

$$\delta_{y1} = \sqrt{S^2(y_1)} = S(y_1) \sqrt{[1 - \bar{r}_1^2(2,3,\dots,i,\dots,k)] \cdot \left(\frac{N-1}{N-2} \right)}. \quad (9)$$

II. Числовий приклад

В основі цього розділу покладений числовий приклад, взятий з [24].

Середнє значення та середні квадратичні

відхилення границі міцності зразків алюмінієвого стопу АМг6 (N = 26 топів) та концентрації основних інгредієнтів Mg, Fe, Si у цьому стопі приведені в табл. 1. Визначимо коефіцієнти моделі (1) у натуральній скалі (емпіричної регресії) для границі міцності стопу (y_1) як функції від концентрації основних інгредієнтів: Mg (x_2), Fe (x_3), Si (x_4) та коефіцієнт множинної кореляції.

1. Розрахунок коефіцієнтів парних кореляцій r_{ij} між границею міцності стопу y_1 та концентраціями первнів x_2, x_3, x_4 (Mg, Fe, Si) зведені у табл. 2 (у загальному вигляді) та у табл. 3 (у числових значеннях).

2. За формулою (1) рівняння емпіричної регресії в натуральній скалі для даного прикладу має вигляд:

$$\bar{y}_1(2,3,4) = b_1 + b_2x_2 + b_3x_3 + b_4x_4. \quad (10)$$

3. За формулою (3) у нормованій скалі рівняння (10) набуває вигляду:

$$\bar{t}_1(2,3,4) = \beta_2t_2 + \beta_3t_3 + \beta_4t_4. \quad (11)$$

4. Знайдемо коефіцієнти $\beta_2, \beta_3, \beta_4$, розв'язуючи систему лінійних рівнянь:

$$\left. \begin{aligned} r_{1,2} &= \beta_2r_{22} + \beta_3r_{32} + \beta_4r_{42}; \\ r_{1,3} &= \beta_2r_{23} + \beta_3r_{33} + \beta_4r_{43}; \\ r_{1,4} &= \beta_2r_{24} + \beta_3r_{34} + \beta_4r_{44}. \end{aligned} \right\} \quad (12)$$

Таблиця 1

Статистичні характеристики границі міцності та концентрацій Mg, Fe, Si в алюмінієвому стопі АМг6

Випадкова величина	Натуральне позначення	Розмірність	Кодоване позначення	Середня \bar{y}_1, \bar{x}_i	Середнє квадратичне відхилення S_i
границя міцності	σ_B	МПа	y_1	359,7079	5,8860
концентрація Mg	C (Mg)	%	x_2	6,34	0,1136
концентрація Fe	C (Fe)	%	x_3	0,30	0,0200
концентрація Si	C (Si)	%	x_4	0,22	0,0500

Таблиця 2

Матриця вибірових коефіцієнтів парних кореляцій r_{ij}

Випадкова величина	Коефіцієнт парної кореляції			
	$y_1 [\sigma_B]$	$x_2 [C (Mg)]$	$x_3 [C (Fe)]$	$x_4 [C (Si)]$
$y_1 [\sigma_B]$	r_{11}	r_{12}	r_{13}	r_{14}
$x_2 [C (Mg)]$	r_{21}	r_{22}	r_{23}	r_{24}
$x_3 [C (Fe)]$	r_{31}	r_{32}	r_{33}	r_{34}
$x_4 [C (Si)]$	r_{41}	r_{42}	r_{43}	r_{44}

Таблиця 3

Значення вибірових коефіцієнтів парних кореляцій r_{ij} між границею міцності та концентраціями Mg, Fe, Si

Випадкова величина	Коефіцієнт парної кореляції			
	$y_1 [\sigma_B]$	$x_2 [C (Mg)]$	$x_3 [C (Fe)]$	$x_4 [C (Si)]$
$y_1 [\sigma_B]$	1	0,5352	-0,4273	-0,2659
$x_2 [C (Mg)]$	0,5352	1	-0,4286	-0,6458
$x_3 [C (Fe)]$	-0,4273	-0,4286	1	0,6154
$x_4 [C (Si)]$	-0,2659	-0,6458	0,6154	1

5. Для системи лінійних рівнянь (12) знайдемо визначники:

$$\Delta_0 = \begin{vmatrix} r_{22} & r_{32} & r_{42} \\ r_{23} & r_{33} & r_{43} \\ r_{24} & r_{34} & r_{44} \end{vmatrix} = \begin{vmatrix} 1 & -0,4286 & -0,6458 \\ -0,4286 & 1 & 0,6154 \\ -0,6458 & 0,6154 & 1 \end{vmatrix} = +0,3611; \quad (13)$$

$$\Delta_2 = \begin{vmatrix} r_{12} & r_{32} & r_{42} \\ r_{13} & r_{33} & r_{43} \\ r_{14} & r_{34} & r_{44} \end{vmatrix} = \begin{vmatrix} 0,5352 & -0,4286 & -0,6458 \\ -0,4273 & 1 & 0,6154 \\ -0,2659 & 0,6154 & 1 \end{vmatrix} = +0,2176; \quad (14)$$

$$\Delta_3 = \begin{vmatrix} r_{22} & r_{12} & r_{42} \\ r_{23} & r_{13} & r_{43} \\ r_{24} & r_{14} & r_{44} \end{vmatrix} = \begin{vmatrix} 1 & 0,5352 & -0,6458 \\ -0,4286 & -0,4273 & 0,6154 \\ -0,6458 & -0,2659 & 1 \end{vmatrix} = -0,1424; \quad (15)$$

$$\Delta_4 = \begin{vmatrix} r_{22} & r_{32} & r_{12} \\ r_{23} & r_{33} & r_{13} \\ r_{24} & r_{34} & r_{14} \end{vmatrix} = \begin{vmatrix} 1 & -0,4286 & 0,5352 \\ -0,4286 & 1 & -0,4273 \\ -0,6458 & 0,6154 & -0,2659 \end{vmatrix} = +0,1320. \quad (16)$$

6. Тоді коефіцієнти рівняння (12) будуть дорівнювати:

$$\beta_2 = \frac{\Delta_2}{\Delta_0} = \frac{+0,2176}{+0,3611} = +0,6026; \quad (17)$$

$$\beta_3 = \frac{\Delta_3}{\Delta_0} = \frac{-0,1424}{+0,3611} = -0,3944; \quad (18)$$

$$\beta_4 = \frac{\Delta_4}{\Delta_0} = \frac{+0,1320}{+0,3611} = +0,3655. \quad (19)$$

7. Рівняння (11) у нормованій скалі набуває вигляду:

$$\bar{t}_1(2,3,4) = 0,6026t_2 - 0,3944t_3 + 0,3655t_4. \quad (20)$$

8. Розрахунок коефіцієнтів моделі (10) у натуральній скалі за формулами (7), (8), (21)–(23) та табл. 1 привів до таких результатів:

$$b_2 = \beta_2 \frac{S_{y1}}{S_2} = 0,6026 \frac{5,8860}{0,1136} = 31,2227; \quad (21)$$

$$b_3 = \beta_3 \frac{S_{y1}}{S_3} = (-0,3944) \frac{5,8860}{0,0200} = -116,0719; \quad (22)$$

$$b_4 = \beta_4 \frac{S_{y1}}{S_4} = 0,3655 \frac{5,8860}{0,0500} = 43,0267. \quad (23)$$

$$b_1 = \bar{y}_1 - [b_2\bar{x}_2 + b_3\bar{x}_3 + b_4\bar{x}_4] = 359,7079 - [31,2227 \cdot 6,34 - 116,0719 \cdot 0,30 + 43,0267 \cdot 0,22] = 187,1117. \quad (24)$$

Тоді модель (10) буде мати вигляд:

$$\bar{y}_1(2,3,4) = 187,1117 + 31,2227x_2 - 116,0719x_3 + 43,0267x_4.$$

9. Рівняння залежності границі міцності алюмінієвого стопу АМг6 від концентрації основних компонентів у натуральній скалі та у натуральному позначенні має вигляд:

$$\sigma_B = 187,1117 + 31,2227 \cdot C(\text{Mg}) - 116,0719 \cdot C(\text{Fe}) + 43,0267 \cdot C(\text{Si}). \quad (25)$$

10. За формулою (5) множинний коефіцієнт кореляції:

$$r_1(2,3,4) = \sqrt{\beta_2 r_{12} + \beta_3 r_{13} + \beta_4 r_{14}} = \sqrt{0,6026 \cdot 0,5352 + (-0,3944) \cdot (-0,4273) + 0,3655 \cdot (-0,2659)} = 0,62758. \quad (26)$$

11. Коефіцієнт множинної кореляції після корекції – врахування числа параметрів рівняння (10) $k = 4$; числа експериментів $N = 26$ – за рівнянням (6) дорівнює:

$$\bar{r}_1(2,3,4) = \sqrt{1 - [1 - r_1^2(2,3,4)] \cdot \left(\frac{N-1}{N-k} \right)} = \sqrt{1 - [1 - 0,62758^2] \cdot \left(\frac{26-1}{26-4} \right)} = 0,55785.$$

12. Оцінка середньої квадратичної помилки під час розрахунку границі міцності зразка алюмінієвого стопу АМг6 за рівнянням моделі (1) у натуральній скалі за (9) становить:

$$\delta_{y1} = S(y_1) \sqrt{[1 - \bar{r}_1^2(2,3,4)] \cdot \left(\frac{N-1}{N-2} \right)} = 5,8860 \sqrt{[1 - 0,55785^2] \cdot \left(\frac{26-1}{26-2} \right)} = 4,9858 \text{ МПа}. \quad (28)$$

13. Висунемо нульову гіпотезу H_0 відносно генерального множинного коефіцієнта кореляції $\rho_1(2,3,4)$, оцінкою якого є вибіровий множинний коефіцієнт кореляції $r_1(2,3,4)$, розрахований за формулою (5):

$$\left. \begin{array}{l} H_0: \rho_1(2,3,4) = 0 \\ \uparrow \\ r_1(2,3,4) \neq 0 \end{array} \right\} \quad (29)$$

14. Перевірка нульової гіпотези H_0 (29) за критичним значенням коефіцієнта кореляції, за критерієм Стюдента та z-перетворенням Фішера привела до таких результатів:

1) за критичним значенням коефіцієнта кореляції [1, 16]:

- для $\alpha=0,05$ $|r_1(2,3,4)|=0,62758 > r_{кр.} \{q=1-0,05/2=0,975; f=N-2=24\}=0,3882$ [16];

- для $\alpha=0,01$ $|r_1(2,3,4)|=0,62758 > r_{кр.} \{q=1-0,01/2=0,995; f=N-2=24\}=0,4958$ [16];

тобто, з ймовірностями $p=0,95$ та $p=0,99$ при рівнях значущості $\alpha=0,05$ та $\alpha=0,01$ нульова гіпотеза відкидається. Тоді стверджуємо, що за $r_{кр.}$ між границею міцності стопу АМг6 та концентраціями основних компонентів С(Мг), С(Fe) та С(Si) існує статистично надійний множинний лінійний зв'язок, при цьому **ступінь лінійності множинного зв'язку** дорівнює:

- для $\alpha=0,05$

$$\xi_1(r)_{0,05} = \frac{|r_1(2,3,4)|}{r_{кр.}(\alpha=0,05)} = \frac{0,62758}{0,3882} = 1,6166 > 1;$$

- для $\alpha=0,01$

$$\xi_1(r)_{0,01} = \frac{|r_1(2,3,4)|}{r_{кр.}(\alpha=0,01)} = \frac{0,62758}{0,4958} = 1,2658 > 1,$$

а ступінь залишків нелінійності у лінійному множинному зв'язку дорівнює:

- для $\alpha=0,05$

$$\xi_2(r)_{0,05} = \frac{r_{кр.}(\alpha=0,05)}{|r_1(2,3,4)|} = \frac{0,3882}{0,62758} = 0,6186 < 1;$$

- для $\alpha=0,01$

$$\xi_2(r)_{0,01} = \frac{r_{кр.}(\alpha=0,01)}{|r_1(2,3,4)|} = \frac{0,4958}{0,62758} = 0,7900 < 1;$$

загальна ступінь щільності множинного зв'язку (лінійного + нелінійного) за $r_{кр.}$:

- для $\alpha=0,05$

$$\xi_{12}(r)_{0,05} = \xi_1(r)_{0,05} + \xi_2(r)_{0,05} = 1,6166 + 0,6186 = 2,2352;$$

- для $\alpha=0,01$

$$\xi_{12}(r)_{0,01} = \xi_1(r)_{0,01} + \xi_2(r)_{0,01} = 1,2658 + 0,7900 = 2,0558;$$

2) за критерієм Стюдента [16, 24]:

- для $\alpha=0,05$

$$|t_p| = \frac{|r_1(2,3,4)|}{\sqrt{1-r_1^2(2,3,4)}} \sqrt{N-2} = \frac{0,62758}{\sqrt{1-0,62758^2}} \times \sqrt{26-2} = 3,9490 > t_T \{q=1-0,05/2=0,975, (36) \\ f = N-2 = 26-2 = 24\} = 2,064 [16];$$

- для $\alpha=0,01$ $|t_p| = 3,9490 > t_T \{q=1-0,01/2=0,995; f=24\} = 2,797 [16],$ (37)

тобто, з ймовірностями $p=0,95$ та $p=0,99$ при рівнях значущості $\alpha=0,05$ та $\alpha=0,01$ нульова гіпотеза H_0 відкидається. Тоді стверджуємо, що за t-критерієм між границею міцності стопу АМг6 та концентраціями основних компонентів С(Мг), С(Fe) та С(Si) існує статистично надійний множинний лінійний зв'язок, при цьому **ступінь лінійності множинного зв'язку** дорівнює:

- для $\alpha=0,05$

$$\xi_1(t)_{0,05} = \frac{|t_p|}{t_T(\alpha=0,05)} = \frac{3,9490}{2,064} = 1,9133 > 1;$$

- для $\alpha=0,01$

$$\xi_1(t)_{0,01} = \frac{|t_p|}{t_T(\alpha=0,01)} = \frac{3,9490}{2,797} = 1,4119 > 1,$$

а ступінь залишків нелінійності у лінійному множинному зв'язку дорівнює:

- для $\alpha=0,05$

$$\xi_2(t)_{0,05} = \frac{t_T(\alpha=0,05)}{|t_p|} = \frac{2,064}{3,9490} = 0,5227 < 1;$$

- для $\alpha=0,01$

$$\xi_2(t)_{0,01} = \frac{t_T(\alpha=0,01)}{|t_p|} = \frac{2,797}{3,9490} = 0,7083 < 1;$$

загальна ступінь щільності множинного зв'язку (лінійного+нелінійного) за t-критерієм Стюдента:

- для $\alpha=0,05$

$$\xi_{12}(t)_{0,05} = \xi_1(t)_{0,05} + \xi_2(t)_{0,05} = 1,9133 + 0,5227 = 2,4360;$$

- для $\alpha=0,01$

$$\xi_{12}(t)_{0,01} = \xi_1(t)_{0,01} + \xi_2(t)_{0,01} = 1,4119 + 0,7083 = 2,1202;$$

3) за Z-перетворенням Фішера [16, 24]:

- для $\alpha=0,05$

$$|z_p| = \frac{1}{2} \ln \frac{1+r_1(2,3,4)}{1-r_1(2,3,4)} = \frac{1}{2} \ln \frac{1+0,62758}{1-0,62758} = 0,7374 > [(z_T \{q=1-\alpha/2=1-0,05/2=0,975\} = 1,96 [24]) \cdot (\sigma_z = \frac{1}{\sqrt{N-3}} = \frac{1}{\sqrt{26-3}} = 0,2085) = 0,4087];$$

- для $\alpha=0,01$ $|z_p| = 0,7374 >$
 $>[(z_{\tau}\{q=0,995\}=2,58 [24]) \cdot (\sigma_z=0,2085)=0,5379],$ (43)

тобто, з ймовірностями $p=0,95$ та $p=0,99$ при рівнях значущості $\alpha=0,05$ та $\alpha=0,01$ нульова гіпотеза H_0 відкидається. Тоді стверджуємо, що за Z-перетворенням Фішера між границею міцності стопу АМг6 та концентраціями основних компонентів C(Mg), C(Fe) та C(Si) існує статистично надійний множинний лінійний зв'язок, при цьому **ступінь лінійності множинного зв'язку** дорівнює:

- для $\alpha=0,05$

$$\xi_1(z)_{0,05} = \frac{|z_p|}{[z_{\tau}(q=0,975) \cdot \sigma_z]} = \frac{0,7374}{0,4087} = 1,8043 > 1;$$
 (44)

- для $\alpha=0,01$

$$\xi_1(z)_{0,01} = \frac{|z_p|}{[z_{\tau}(q=0,995) \cdot \sigma_z]} = \frac{0,7374}{0,5379} = 1,3709 > 1,$$
 (45)

а **ступінь залишків нелінійності у лінійному множинному зв'язку** дорівнює:

- для $\alpha=0,05$

$$\xi_2(z)_{0,05} = \frac{[z_{\tau}(q=0,975) \cdot \sigma_z]}{|z_p|} = \frac{0,4087}{0,7374} = 0,5542 < 1;$$
 (46)

- для $\alpha=0,01$

$$\xi_2(z)_{0,01} = \frac{[z_{\tau}(q=0,995) \cdot \sigma_z]}{|z_p|} = \frac{0,5379}{0,7374} = 0,7295 < 1;$$
 (47)

загальна ступінь щільності множинного зв'язку (лінійного + нелінійного) за z-перетворенням Фішера:

- для $\alpha=0,05$

$$\xi_{12}(z)_{0,05} = \xi_1(z)_{0,05} + \xi_2(z)_{0,05} = 1,8043 + 0,5542 = 2,3585;$$
- для $\alpha=0,01$

$$\xi_{12}(z)_{0,01} = \xi_1(z)_{0,01} + \xi_2(z)_{0,01} = 1,3709 + 0,7295 = 2,1004.$$

15. Висунемо нульову гіпотезу H_0 відносно генерального множинного коефіцієнта кореляції $\bar{\rho}_{1(2,3,4)}$, оцінкою якого є вибіркового множинного коефіцієнта кореляції $r_{1(2,3,4)}$, підданий корекції за рівнянням (6) з врахування числа параметрів рівняння (60): $k=4$ та числа спостережень $N=26$:

$$\left. \begin{array}{l} H_0: \bar{\rho}_{1(2,3,4)} = 0 \\ \uparrow \\ \bar{r}_{1(2,3,4)} \neq 0 \end{array} \right\} \quad (48)$$

16. Здійснено перевірку нульової гіпотези H_0 (48) за процедурою (30)–(47):

1) за критичним значенням коефіцієнта кореляції [1, 16]:

- для $\alpha=0,05$ $|\bar{r}_1(2,3,4)| = 0,5579 >$
 $>r_{кр.}\{q=0,975; f=24\}=0,3882 [16];$ (49)

- для $\alpha=0,01$ $|\bar{r}_1(2,3,4)| = 0,5579 >$
 $>r_{кр.}\{q=0,995; f=24\}=0,4958 [16],$ (50)

тобто, з ймовірностями $p=0,95$ та $p=0,99$ H_0 відкидається. Тоді стверджуємо, що за $r_{кр.}$ між σ_b та концентраціями основних компонентів C(Mg), C(Fe) та C(Si) існує статистично надійний множинний лінійний зв'язок після корекції, при цьому **ступінь лінійності зв'язку** дорівнює:

- для $\alpha=0,05$

$$\xi_1(\bar{r})_{0,05} = \frac{|\bar{r}_1(2,3,4)|}{r_{кр.}(\alpha=0,05)} = \frac{0,5579}{0,3882} = 1,4372 > 1;$$
 (51)

- для $\alpha=0,01$

$$\xi_1(\bar{r})_{0,01} = \frac{|\bar{r}_1(2,3,4)|}{r_{кр.}(\alpha=0,01)} = \frac{0,5579}{0,4958} = 1,1253 > 1,$$
 (52)

а **ступінь залишків нелінійності у лінійному множинному зв'язку** дорівнює:

- для $\alpha=0,05$

$$\xi_2(\bar{r})_{0,05} = \frac{r_{кр.}(\alpha=0,05)}{|\bar{r}_1(2,3,4)|} = \frac{0,3882}{0,5579} = 0,6958 < 1;$$
 (53)

- для $\alpha=0,01$

$$\xi_2(\bar{r})_{0,01} = \frac{r_{кр.}(\alpha=0,01)}{|\bar{r}_1(2,3,4)|} = \frac{0,4958}{0,5579} = 0,8887 < 1;$$
 (54)

загальна ступінь щільності множинного зв'язку (лінійного + нелінійного) за $r_{кр.}$:

- для $\alpha=0,05$

$$\xi_{12}(r)_{0,05} = \xi_1(r)_{0,05} + \xi_2(r)_{0,05} = 1,4372 + 0,6958 = 2,1330;$$
- для $\alpha=0,01$

$$\xi_{12}(r)_{0,01} = \xi_1(r)_{0,01} + \xi_2(r)_{0,01} = 1,1253 + 0,8887 = 2,0140;$$

2) за критерієм Стюдента [16, 24]:

- для $\alpha=0,05$

$$|t_p| = \frac{|\bar{r}_1(2,3,4)|}{\sqrt{1-r_1^2(2,3,4)}} \sqrt{N-2} = \frac{0,55785}{\sqrt{1-0,55785^2}} \cdot \sqrt{26-2} =$$

 $= 3,2929 > t_{\tau}\{q=0,975, f=24\} = 2,064 [16];$ (55)

- для $\alpha=0,01$ $|t_p| = 3,2929 > t_{\tau}$
 $>t_{\tau}\{q=0,995; f=24\}=2,797 [16],$ (56)

тобто, з ймовірностями $p=0,95$ та $p=0,99$ H_0 відкидається. Тоді стверджуємо, що за t-критерієм між σ_b та концентраціями основних компонентів C(Mg), C(Fe) та C(Si) існує статистично надійний множинний лінійний зв'язок після корекції, при цьому **ступінь лінійності зв'язку** дорівнює:

- для $\alpha=0,05$

$$\xi_1(t)_{0,05} = \frac{|t_p|}{t_{\tau}(\alpha=0,05)} = \frac{3,2929}{2,064} = 1,5954 > 1;$$
 (57)

- для $\alpha=0,01$

$$\xi_1(t)_{0,01} = \frac{|t_p|}{t_r(\alpha=0,01)} = \frac{3,2929}{2,797} = 1,1773 > 1, \quad (58)$$

а ступінь залишків нелінійності у лінійному множинному зв'язку дорівнює:

- для $\alpha=0,05$

$$\xi_2(t)_{0,05} = \frac{t_r(\alpha=0,05)}{|t_p|} = \frac{2,064}{3,2929} = 0,6268 < 1; \quad (59)$$

- для $\alpha=0,01$

$$\xi_2(t)_{0,01} = \frac{t_r(\alpha=0,01)}{|t_p|} = \frac{2,797}{3,2929} = 0,8494 < 1; \quad (60)$$

загальна ступінь щільності множинного зв'язку (лінійного + нелінійного) за t-критерієм Стюдента:

- для $\alpha=0,05$

$$\xi_{12}(t)_{0,05} = \xi_1(t)_{0,05} + \xi_2(t)_{0,05} = 1,5954 + 0,6268 = 2,2222;$$

- для $\alpha=0,01$

$$\xi_{12}(t)_{0,01} = \xi_1(t)_{0,01} + \xi_2(t)_{0,01} = 1,1773 + 0,8494 = 2,0267;$$

3) за Z-перетворенням Фішера [16, 24]:

- для $\alpha=0,05$

$$|z_p| = \frac{1}{2} \ln \frac{1 + \bar{r}_1(2,3,4)}{1 - \bar{r}_1(2,3,4)} = \frac{1}{2} \ln \frac{1 + 0,55785}{1 - 0,55785} = 0,6297 > [(z_T \{q = 0,975\} = 1,96 [24]) \times (\sigma_z = 0,2085) = 0,4087]; \quad (61)$$

- для $\alpha=0,01$ $|z_p| = 0,6297 >$

$$> [(z_T \{q = 0,995\} = 2,58 [24]) (\sigma_z = 0,2085) = 0,5379], \quad (62)$$

тобто, з ймовірностями $p=0,95$ та $p=0,99$ H_0 відкидається. Тоді стверджуємо, що за Z-перетворенням Фішера між σ_b та концентраціями

основних компонентів C(Mg), C(Fe) та C(Si) існує статистично надійний множинний лінійний зв'язок після корекції, при цьому ступінь лінійності зв'язку дорівнює:

- для $\alpha=0,05$

$$\xi_1(z)_{0,05} = \frac{|z_p|}{[z_T(q = 0,975) \cdot \sigma_z]} = \frac{0,6297}{0,4087} = 1,5407 > 1; \quad (63)$$

- для $\alpha=0,01$

$$\xi_1(z)_{0,01} = \frac{|z_p|}{[z_T(q = 0,995) \cdot \sigma_z]} = \frac{0,6297}{0,5379} = 1,1707 > 1, \quad (64)$$

а ступінь залишків нелінійності у лінійному множинному зв'язку дорівнює:

- для $\alpha=0,05$

$$\xi_2(z)_{0,05} = \frac{[z_T(q = 0,975) \cdot \sigma_z]}{|z_p|} = \frac{0,4087}{0,6297} = 0,6491 < 1; \quad (65)$$

- для $\alpha=0,01$

$$\xi_2(z)_{0,01} = \frac{[z_T(q = 0,995) \cdot \sigma_z]}{|z_p|} = \frac{0,5379}{0,6297} = 0,8542 < 1; \quad (66)$$

загальна ступінь щільності множинного зв'язку (лінійного + нелінійного) за z-перетворенням Фішера:

- для $\alpha=0,05$

$$\xi_{12}(z)_{0,05} = \xi_1(z)_{0,05} + \xi_2(z)_{0,05} = 1,5407 + 0,6491 = 2,1898;$$

- для $\alpha=0,01$

$$\xi_{12}(z)_{0,01} = \xi_1(z)_{0,01} + \xi_2(z)_{0,01} = 1,1707 + 0,8542 = 2,0249.$$

Підсумкова матриця значень множинних коефіцієнтів кореляцій, результати перевірки їх значущості за критеріями r_{kp} , t-Стюдента та F-перетворення Фішера (Z) та ступенів лінійності і нелінійності приведена у табл. 4.

Таблиця 4

Підсумкова матриця значень множинних коефіцієнтів лінійних кореляцій та ступенів лінійності і нелінійності за критеріями r_{kp} , t-Стюдента та F-перетворення Фішера (Z)

За критерієм	До корекції		Після корекції	
	$r_1(2,3,4) = 0,62758$		$\bar{r}_1(2,3,4) = 0,55785$	
критичним коефіцієнтом кореляції (r_{kp})	$\alpha=0,05$	$\alpha=0,01$	$\alpha=0,05$	$\alpha=0,01$
ступінь лінійності $\xi_1(r)$	1,6166	1,2658	1,4372	1,1253
ступінь нелінійності $\xi_2(r)$	0,6186	0,7900	0,6958	0,8887
загальна ступінь щільності зв'язку $\xi_{12}(r)$	2,2352	2,0558	2,1330	2,0140
Стюдента (t)	$\alpha=0,05$	$\alpha=0,01$	$\alpha=0,05$	$\alpha=0,01$
ступінь лінійності $\xi_1(t)$	1,9133	1,4119	1,5954	1,1773
ступінь нелінійності $\xi_2(t)$	0,5227	0,7083	0,6268	0,8494

загальна ступінь щільності зв'язку $\xi_{12}(t)$	2,4360	2,1202	2,2222	2,0267
F-перетворенням Фішера (Z)	$\alpha=0,05$	$\alpha=0,01$	$\alpha=0,05$	$\alpha=0,01$
Ступінь лінійності $\xi_1(z)$	1,8043	1,3709	1,5407	1,1707
Ступінь нелінійності $\xi_2(z)$	0,5542	0,7295	0,6491	0,8542
загальна ступінь щільності зв'язку $\xi_{12}(z)$	2,3585	2,1004	2,1898	2,0249

Висновки

1. Розроблено методики застосування методів множинної лінійної кореляції та регресії в хемії та хемічній технології.

2. На числовому прикладі, взятого з хемічного матеріалознавства [24], показано процедуру розрахунків коефіцієнтів лінійного рівняння множинної регресії та його кореляції за кількістю параметрів математичної моделі та числа спостережень (числа експериментів).

3. На числовому прикладі отримано: лінійне рівняння множинної регресії залежності границі міцності σ_b від концентрацій основних компонентів C(Mg), C(Fe) та C(Si) в алюмінієвому стопі АМг6.

4. Розрахований коефіцієнт множинної кореляції та доведена його значущість та ступені лінійності та нелінійності лінійного зв'язку (для ступеня значущості 1 і 5 %) за критичним коефіцієнтом кореляції, критерієм Стюдента та z-перетворенням Фішера.

Література

1. **Адлер Ю.П.** Планирование эксперимента при поиске оптимальных условий / Ю.П. Адлер, Е.В. Маркова, Ю.В. Грановский. – 2-е изд., перераб. и доп. – Москва: Наука, 1976. – 280 с.: ил., табл. – Библиогр. в конце гл.
2. **Ахназарова С.Л.** Оптимизация эксперимента в химии и химической технологии / С.Л. Ахназарова, В.В. Кафаров. – Москва: Высш. шк., 1978. – 320 с.: ил., табл. – Библиогр.: с. 302 – 303 (53 наименов.). – Приложения: с. 304 – 317 (14 табл.).
3. **Бендат Дж.С.** Измерение и анализ случайных процессов / Дж.С. Бендат, А.Г. Пирсол; пер. с англ. Г.В. Матушевского, В.Е. Привальского; под ред. И.Н. Коваленко. – Москва: Мир, 1971. – 408 с. – Перевод за изд.: Measurement and analysis of random data / Julius S. Bendat, Allan G. Piersol. – John Wiley and Sons, Inc. – New York-London-Sydney, 1967.: ил., табл. – Предмет. указатель: с. 403-408. – Библиогр.: с. 400-402 (59 наименов.).
4. **Биометрия** / Н.В. Глотов, Л.А. Животовский, Н.В. Хованов, Н.Н. Хромов-Борисов. – Ленинград, 1982.
5. **Бондар А.Г., Статюха Г.А.** Планирование эксперимента в химической технологии. – Киев: Вища шк., 1976. – 220с.
6. **Волощенко А.Б.** Теорія ймовірностей та математична статистика: навч.-метод. посібник для самост. вивчення дисц. [для студ. економ. спеціал. вищ. навч. заклад.] / А.Б. Волощенко, І.А. Джалладова; [Мін-во освіти і науки України; гриф: лист № 14 / 18.2-613 від 22.03.2002 р.]. – Київ: Київ. Нац. економ. ун-т, 2003. – 256 с.: іл., табл. – Приклади розв. завдань і вправи для самост. розв'язання в кінці розд. – Блочно-модул. контроль: с. 183 – 203 (9 варіантів). – Відповіді: с. 204 – 216. – Бібліогр.: с. 217 (18 назв). – Додатки: с. 218 – 254 (8 табл.). – ISBN 966 – 574 – 459 – 3.
7. **Езекиэл М., Фокс К.** Методы анализа корреляций и регрессий. – Москва: Статистика, 1966. – 470с.
8. **Жлуктенко В.І.** Теорія ймовірностей і математична статистика: навч.-метод. посібник [для студ. економ. вищ. навч. заклад.]: У 2-х ч. – Ч. II. Математична статистика / В.І. Жлуктенко, С.І. Наконечний, С.С. Савіна; [Мін-во освіти і науки України; гриф: лист № 14 / 18.2-183 від 27.02.2001 р.]. – Київ: Київ. нац. економ. ун-т, 2001. – 336 с.: іл., табл. – Теор. запит. та завдання до теми в кінці теми. – Лаб. роб. після тем 14, 15. – Додатки: с. 242 – 246, 292 – 331. – Бібліогр.: с. 246 (4 назви). – ISBN 966–574–265 – 5.
9. **Зажигаев Л.С.** Методы планирования и обработки результатов физического эксперимента / Л.С. Зажи́гаев, А.А. Ки́шьян, Ю.И. Романи́ков. – Москва: Атомиздат, 1978. – 232 с.: ил., табл. – Приложение: с. 144-229 (16 табл.). – Библиогр.: с. 230-231.
10. **Іванюта І.Д.** Елементи теорії ймовірностей та математичної статистики: навч. посібник [для студ. економ. спеціал. вищ. навч. заклад.] / І.Д. Іванюта, В.І. Рибалка, І.А. Рудоміно-Дусяцька; [Мін-во освіти і науки України; гриф: лист № 14 / 18.2-271 від 11.02.2003 р.]. – Київ: Слово, 2003. – 271 с.: іл.,

- табл. – Завдання до самостійн. роботи: с. 235 – 261 (15 завд.). – Додатки: с. 262 – 267 (6 табл.). – Бібліогр.: с. 268 (6 назв.). – ISBN 966 – 8407 – 01 – 6.
11. **Корн Г.** Справочник по математике для научных работников и инженеров / Г. Корн, Т. Корн. – 4-е изд.; пер. с англ. И.Г. Арамановича, А.М. Березмана, И.А. Вайнштейна и др.; под общ. ред. И.Г. Арамановича. – Москва: Наука, 1978. – 832 с. – Перевод за изд.: *Mathematical Handbook for Scientists and Engineers Definitions, Theorems and Formulas for Reference and Review.* – Second, Enlarged and Revised Edition / Granino A. Korn, Ph. D., Theresa M. Korn, M.S. – McGraw-Hill Book Company: New York-San Francisco-Toronto-London-Sydney, 1968. – ил., табл. – Библиогр.: с. 796 – 800 (183 наим.). – Указ. важн. обозн.: с. 801 – 803. – Предмет. указ.: с. 804 – 831. – Перечень табл. по гл.: с. 20 – 22.
 12. **Кузишин О.В.** Критерії оцінки розподілу мікроступів на поверхні твердого тіла / О.В. Кузишин, О.Г. Сіренко, Л.Я. Мідак, Г.О. Сіренко // *Фізика і хімія твердого тіла.* – 2008. – Т. 9. – № 2. – С.407-414: ил. 1, табл. 2. – Бібліогр.: с. 412 (52 назви).
 13. **Лакин Г.Ф.** Биометрия: Учеб. пособие [для биол. спец. вузов] / Георгий Филиппович Лакин. – 4-е изд., пераб. и доп. – Москва: Высш. шк., 1990. – 352 с.: ил., табл. – Прилож.: с. 319 – 345 (26 мат. табл.). – Библиогр.: с. 346 – 347 (58 наим.). – Предмет. указ.: с. 348 – 350.
 14. **Лукомский Я.И.** Теория корреляции и ее применение к анализу производства. – Москва: Госстатиздат, 1961.
 15. **Математичні методи в хімії та хімічній технології:** Навч. посіб. / Ю.К. Рудавський, Е.М. Мокрий, З.Г. Піх, І.Й. Куриляк, М.М. Чип. – Львів: Світ, 1993. – 208 с. – ISBN 5 – 7773 – 0136 – 3.
 16. **Мюллер П., Нойман П., Шторм Р.** Таблицы по математической статистике / Пер. с нем. и предисловие В.М. Ивановой. – Москва: Финансы и статистика, 1982. – 272 с.: ил.
 17. **Налимов В.В.** Логические основания планирования эксперимента / В.В. Налимов, Т.И. Голикова. – Москва: Металлургия, 1976. – 128 с.: ил., табл. – Библиогр.: с. 126 – 128 (81 наим.).
 18. **Налимов В.В.** Статистические методы планирования экстремальных экспериментов / В.В. Налимов, Н.А. Чернова. – Москва: Наука, 1965. – 340 с.: ил., табл. – Библиогр.: с. 328 – 338 (204 наим.). – Предмет. указ.: с. 339 – 340. – Приложения: с. 309 – 327 (I. Элементы матричной алгебры. Симплексы. II. Планы дробных реплик).
 19. **Налимов В.В.** Теория эксперимента. – Москва: Наука, 1971. – 207с.
 20. **Неділько С.А.** Математичні методи в хімії: підручник [для студ. хім. спеціал. вищ. навч. закладів] / Сергій Неділько; [Мін-во освіти і науки України; гриф: лист № 1 / 11-1536 від 13.04.2004 р.]. – Київ: Либідь, 2005. – 256 с.: ил. – Завдання для самостійн. роботи та бібліогр. в кінці розд. – ISBN 966 – 06–03843.
 21. **Нелинейная корреляция и регрессия** / С.Н. Воловельская, А.И. Жилин, С.А. Кулиш, В.Б. Сивый. – Київ: Техніка, 1971. – 130 с.
 22. **Сигорский В.П.** Математический аппарат инженера / Виталий Петрович Сигорский. – 2-е изд., стереотип. – Київ: Техніка, 1977. – 768 с.: – ил., табл. – Библиогр. в конце гл. – Предмет. указ.: с. 752 – 764.
 23. **Сіренко Г.О.** Методи оцінок впливу факторів на функції відгуку та процедури відсіювання параметрів оптимізації при вирішенні багатопараметричних завдань у матеріалознавстві / Г.О. Сіренко, О.В. Кузишин, О.Г. Сіренко, Л.Я. Мідак, Л.М. Солтис // *Фізика і хімія твердого тіла.* – 2009. – Т. 10. – № 2. – С.423-439: ил. 2, табл. 10. – Бібліогр.: с. 437-438 (26 назв).
 24. **Степнов М.Н.** Статистическая обработка результатов механических испытаний / Михаил Никитович Степнов. – Москва: Машиностроение, 1972. – 232 с.: ил., табл. – Библиогр.: с. 229-230 (36 назв.).
 25. **Тихомиров В.Б.** Планирование и анализ эксперимента / Владислав Борисович Тихомиров. – Москва: Легкая индустрия, 1974. – 264 с.: ил., табл. – Приложение: с. 255-257 (4 табл.). – Библиогр.: с. 258-261 (99 наименов.).
 26. **Федоров В.В.** Теория оптимального эксперимента (планирование регрессионных экспериментов): монография / Валерий Вадимович Федоров. – Москва: Наука, 1971. – 312 с.: ил., табл. – Библиогр.: с. 309 – 312 (79 наим.).

Сіренко Г.О. – доктор технічних наук, професор, завідувач катедри неорганічної та фізичної хемії.

Мідак Л.Я. – кандидат хімічних наук, доцент катедри неорганічної та фізичної хемії.

Сіренко О.Г. – кандидат біологічних наук, науковий співробітник відділу ландшафтного будівництва.

Рецензент

Фрейк Д.М. – доктор хімічних наук, професор, завідувач катедри фізики і хемії твердого тіла Прикарпатського національного університету імені Василя Стефаника.