МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ХАРКІВСЬКА НАЦІОНАЛЬНА АКАДЕМІЯ МІСЬКОГО ГОСПОДАРСТВА

МЕТОДИЧНІ ВКАЗІВКИ
до науково-дослідницької практики з дисципліни «ОРГАНІЗАЦІЯ НАУКОВИХ ДОСЛІДЖЕНЬ» (Статистичні методи. Аналіз та оформлення наукових досліджень)

(для студентів 5 курсу всіх форм навчання спец. 8.092108 – «Теплогазопостачання і вентиляція»)

Харків – ХНАМГ – 2009

Методичні вказівки до науково-дослідницької практики з дисципліни «Організація наукових досліджень» (Статистичні методи. Аналіз та оформлення наукових досліджень) (для студентів 5 курсу всіх форм навчання спец. 8.092108 – «Теплогазопостачання і вентиляція»). Укл.: Капцов І.І., Ромашко О.В., Гапонова Л.В., Гранкіна В.В. – Харків: ХНАМГ, 2009. – 59 с.

Укладачі: Капцов І.І.,
Ромашко О.В.,
Гапонова Л.В.,
Гранкіна В.В.

Рекомендовано кафедрою експлуатації газових і теплових систем, протокол № 5 від 28.05.2009р.

СТАТИСТИЧНІ МЕТОДИ.
АНАЛІЗ ТА ОФОРМЛЕННЯ НАУКОВИХ ДОСЛІДЖЕНЬ

Практичне заняття № 1.
Основні статистичні характеристики. Обробка результатів наукових досліджень методами кореляційного та регресійного аналізів. Методи графічного зображення результатів експериментів

Ціль роботи: Вивчити основні статистичні характеристики. Освоїти обробку результатів наукових досліджень методами кореляційного і регресійного аналізу. Розглянути методи графічного зображення результатів експериментів.

Теоретичні пояснення:

Теорія статистичного оцінювання розглядає два основні види оцінок: точкові та інтервальні.

Точковою оцінкою називають деяку функцію результатів спостереження $S_n^*(x_1, x_2, \ldots, x_n)$, значення якої в даних умовах береться за найбільше наближення до значення параметра S генеральної сукупності.

Проте при вибірці невеликого об'єму точкова оцінка S_n^* може істотно відрізнятися від дійсного значення параметру, тобто приводити до грубих помилок. Тому у разі малої вибірки часто використовують інтервальні оцінки.

Інтервальною оцінкою називають числовий інтервал (S_1^*, S_2^*), визначуваний за результатами вибірки, відносно якого можна стверджувати з визначеною, близькою до одиниці, ймовірністю, що він містить значення оцінюваного параметра генеральної сукупності.

Середні значення і їх оцінки

Випадкова величина x повністю задається ***щільністю імовірності*** $\rho(x)$ (інші назви – розподіл імовірності, розподіл величини x).

Середнє значення $\langle x \rangle$ вимірюваної величини x вказує центр розподілу, біля якого групуються результати окремих вимірювань:

$$\langle x \rangle = \frac{1}{n}\sum_{i=1}^{n} x_i \ .$$

Методи обчислення середніх

Приклад 1.1. Змішуються дві рідини об'ємом 2 і 5 літрів з концентраціями якоїсь речовини 500 і 200 мг/л відповідно. Визначити концентрацію речовини в суміші.

Рішення. За формулою [3] маємо

$$\bar{x} = \sum p_i x_i / \sum p_i = (2 \cdot 500 + 5 \cdot 200)/(2+5) = 285{,}7 \text{ мг/л}, \qquad (1.1)$$

а за формулою (5.4.7) [3]

$$S_* = \sqrt{\sum p_i (x_i - \bar{x})^2 / \sum p_i}\ , \qquad (1.2)$$

$$S_* = \sqrt{\frac{2(500-285{,}7)^2 + 5(200-285{,}7)^2}{2+5}} = 135{,}5 \text{ мг/л}.$$

Зважене стандартне відхилення визначається по формулі [3]

$$S = \sqrt{\frac{S_1^2(n_1-1) + S_2^2(n_2-1) + ... + S_k^2(n_k-1)}{\sum n_k - k}}\ . \qquad (1.3)$$

Приклад 1.2. Нехай серія з N дослідів дала наступні значення:

$n_1 = 10$, $\bar{x}_1 = 10$, $(S_1 = 2)$, $S_1^2 = 4$,

$n_2 = 12$, $\bar{x}_2 = 9$, $(S_2 = 3)$, $S_2^2 = 9$,

$n_3 = 6$, $\bar{x}_3 = 7$, $(S_3 = 1)$, $S_3^2 = 1$.

Потрібно визначити вибіркове зважене середнє $\bar{x}$, зважене стандартне відхилення і відношення стандартного відхилення до середнього значення (коефіцієнт варіації *V*).

Рішення:

$$\bar{x} = \frac{(10 \cdot 10 + 12 \cdot 9 + 6 \cdot 7)}{28} = 8{,}93\ ;$$

$$S = \sqrt{\frac{4 \cdot (10-1) + 9 \cdot (12-1) + 1 \cdot (6-1)}{28-3}} = 2{,}37\ ;$$

$$S^2 = 5{,}6\ .$$

Коефіцієнт варіації, визначуваний по формулі [3], рівний:

$$V = S/\bar{x}, \quad \bar{x} \neq 0, \quad V = 2{,}37/8{,}93 = 0{,}265.$$

Приклад 1.3. Дана вибірка, приведена табл. 1.1. Визначити значення медіани *Me* і моди *Mo*.

Рішення: Оскільки медіана лежить між 20-м і 21-м членами ряду, а 2 + 3 + 5 + 6 = 16 і 2 + 3 + 5 + 6 + 10 = 26, то *Me* належить до

п'ятого класу (інтервалу). Тоді при x_{Me} = 9, b = 2, n = 40, $\sum m_{Me-1}$ = 16, m_{Me} = 10 за формулою [3] маємо:

$$Me = \tilde{x} = x_{Me} + b \cdot \left(\frac{0{,}5 \cdot n - \sum m_{Me-1}}{m_{Me}} \right), \qquad (1.4)$$

де x_{Me} – нижня межа класу (інтервалу), що містить медіану; b – ширина медіанного інтервалу; $\sum m_{Me-1}$ – сума накопичених частот, передуючих медіанному інтервалу; m_{Me} – частота медіанного інтервалу.

$$Me = \tilde{x} = 9 + 2 \cdot \left(\frac{0{,}5 \cdot 40 - 16}{10} \right) = 9{,}8 .$$

Таблиця 1.1 – Розрахункова виборка

Інтервал	**Середина інтервалу**	**Частота m_i**	**Сума накопичених частот**
1<x<3	2	2	2
3<x<5	4	3	5(2+3)
5<x<7	6	5	10(5+5)
7<x<9	8	6	16(10+6)
9<x<11	10	10	26(16+10)
11<x<13	12	9	35(26+9)
13<x<15	14	5	40(35+5)
		n = 40	

Для нашого випадку середнє значення вибірки за формулою (1.1)

$$\bar{x} = \frac{(2 \cdot 2 + 4 \cdot 3 + 6 \cdot 5 + 8 \cdot 6 + 10 \cdot 10 + 12 \cdot 9 + 14 \cdot 5)}{40} = 9{,}3 .$$

Величину Mo підрахуємо за формулою [3]:

$$Mo = x_{Mo} + b \cdot \frac{m_{Mo} - m_{Mo-1}}{2 \cdot m_{Mo} - m_{Mo-1} - m_{Mo+1}}, \qquad (2.5)$$

де x_{Mo} – нижня межа інтервалу, що містить моду; b - величина модального інтервалу; m_{Mo} – частота модального інтервалу; $m_{Mo\text{-}1}$ – частота інтервалу, передуючого модальному; m_{Mo+1} – частота інтервалу, наступного за модальним.

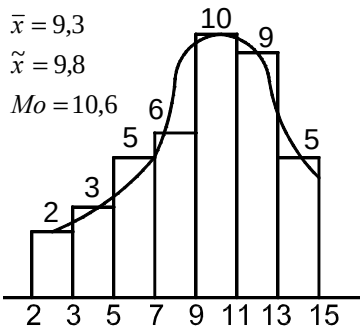

Рис. 1.1

$$Mo = 9 + \frac{2 \cdot (10-6)}{2 \cdot 10 - 6 - 9} = 10{,}6 .$$

Тут *Mo* є максимумом апроксимуючої параболи, що проходить через точки: (x_{Me-1}, m_{Me-1}); (x_{Me}, m_{Me}); (x_{Me+1}, m_{Me+1}). Для негативно-асиметричного унімодального розподілу справедливо нерівність $\bar{x} < \tilde{x} < Mo$ (рис. 2.1). Для симетричного унімодального безперервного розподілу значення *Mo*, *Me* і $\bar{x}$ *співпадають*.

Теоретичні середні (моменти розподілу)

Третій центральний момент служить для характеристики *асиметрії* (або *скошеності)* розподілу (рис. 1.2) і визначається за формулою

$$\mu_3 = \sum_1^n (x_i - \bar{x})^3 / (n-1) . \qquad (1.6)$$

Четвертий центральний момент служить для характеристики «крутизни» розподілу (рис. 1.3)

$$\mu_4 = \sum_1^n (x_i - \bar{x})^4 / (n-1) . \qquad (1.7)$$

На практиці зручніше користуватися безрозмірними величинами, наприклад, в даному випадку – *коефіцієнтом асиметрії* A_x

$$A_x = \mu_3 / S_x^3 \qquad (2.8)$$

та *ексцесом*

$$C_x = (\mu_4 / S_x^4) - 3 . \qquad (2.9)$$

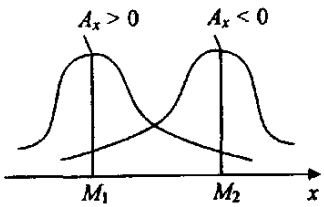

Рис. 2.2

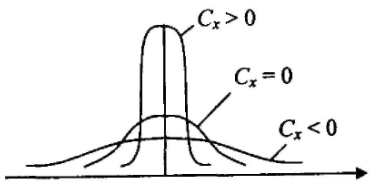

Рис. 2.3

Приклад 1.4. Для умов прикладу 2.3 в нашому випадку маємо:

$$\sum_{1}^{n}(x_i-\overline{x})^3 = 2\cdot(2-9{,}3)^3+3\cdot(4-9{,}3)^3+5\cdot(6-9{,}3)^3+6\cdot(8-9{,}3)^3+$$

$$+10\cdot(10-9{,}3)^3+10\cdot(10-9{,}3)^3+9\cdot(12-9{,}3)^3+5\cdot(14-9{,}3)^3=-717{,}8.$$

Тоді μ_3 = -717,8/(40-1)=-18,4,

Відкіля за формулою (1.8) при

S = 3,34; A_x = -18,4/3,34^3 = -0,494.

Тоді μ_4 = 11577,4 / 39 = 296,8, відкіля за формулою (1.9) маємо

C_x = (296,8/3,34^4) -3 = -0,615.

Помилка в оцінці показника асиметрії:

$$S_A=\sqrt{\frac{6\cdot(n-1)}{(n+1)\cdot(n+3)}}. \qquad S_A=\sqrt{\frac{6\cdot(40-1)}{(40+1)\cdot(40+3)}}=0{,}364, \qquad (1.10)$$

а розсіяння A_x і S_A визначаються дисперсіями:

$$S_{Ax}=\sqrt{6/(n+3)}, \qquad (1.11) \qquad S_{SA}=2\cdot\sqrt{6/(n+5)}, \qquad (1.12)$$

$$S_{Ax}=\sqrt{6/(40+3)}=0{,}374, \qquad S_{SA}=2\cdot\sqrt{6/(40+5)}=5{,}480.$$

Якщо в формулах (1.8) та (12.9) замість 5 допустити

S_* = 3,30, то A_{x*} = -0,512, а C_{x*} = -0,497.

Оцінки довірчих меж для дійсного значення виміряної величини

Приклад 1.5. Нехай для сорока вимірювань, результати яких приведені в прикладі 1.3, відомо, що S_* = 3,30, $\overline{x}$ = 9,3. Потрібно оцінити дійсне значення вимірюваної величини X з надійністю P = 0,95.

Рішення: За додатком В для $P = 2\Phi(t)$ = 0,95 знаходимо t = 1,960. Отже, з надійністю 0,95 по залежності [3] можна вважати, що довірча оцінка має вигляд

$$\left|X-\overline{x}\right|<t(P)\cdot\sigma/\sqrt{n}\ , \qquad (1.13)$$

$$\left|X-\overline{x}\right|=\left|X-9{,}3\right|<1{,}960\cdot(3{,}3/\sqrt{40})=1{,}02,$$

тобто значення X лежить в інтервалі (8,28; 10,32) або $X = \overline{x} \pm 1{,}02$.

Приклад 1.6. Для умов попереднього прикладу за додатком Е маємо t(0,95; 40 – 1) = 2,023. Тоді за формулою [3]

$$\left|X-\overline{x}\right|<t(P;k)\cdot S/\sqrt{n-1}\ , \qquad (1.14)$$

$$\left|X-\overline{x}\right|<2{,}023\cdot(3{,}3/\sqrt{40-1})=1{,}07,$$

тобто, значення X лежить в інтервалі X = 9,3 ± 1,07.

За правилом трьох сигм [3] при невідомій σ

$$\left|X-\bar{x}\right|<3S/\sqrt{n}\ , \qquad (1.15)$$

а для нашого випадку

$$X = 9{,}3 \pm 3 \cdot 3{,}3/\sqrt{40} \approx 9{,}3 \pm 1{,}6.$$

Приклад 2.7. У циліндрі з висотою стовпа води 0,7-1,0 м визначався час осадження окремих частинок піску. Розрахункові значення швидкості осадження (v) приведені в табл. 1.2. Потрібно знайти розподіл статистичної імовірності швидкості осадження (гідравлічної крупності) частинок піску фільтруючого матеріалу. Провести оцінку математичного очікування і середньоквадратичного відхилення. Оцінити найбільш вірогідну швидкість осадження частинок піску.

Рішення:

1. Діапазон зміни швидкостей осадження від $V_{\max}$ = 36,1 до $V_{\min}$ =15,1 ділимо на n інтервалів. Кількість інтервалів можна вибирати по напівемпіричній формулі

$$n = 1 + 3{,}32 \cdot \lg N = 1 + 3{,}32 \cdot \lg 100 \approx 7{,}64\,, \qquad (1.16)$$

де N - число частинок, для яких проводилося визначення швидкостей осадження (у нашому випадку N = 100).

Число інтервалів n, одержане за формулою, округляється до найближчого цілого (приймаємо n = 7).

2. Визначається довжина інтервалу варіаційного ряду

$$\Delta = (x_{\min} - x_{\max})/n = (36{,}1 - 15{,}1)/7 = 3{,}0\,.$$

Таблиця 1.2 – Результати визначення швидкостей осадження частинок піску

№ п.п.	***V*, см/с**	**№ п.п.**	***V*, см/с**	**№ п.п.**	***V*, см/с**	**№ п.п.**	***V*, см/с**	**№ п.п.**	***V*, см/с**
1	22,4	21	22,4	41	23,2	61	25,0	81	24,1
2	21,0	22	21,0	42	27,1	62	24,1	82	21,7
3	25,0	23	26,0	43	30,9	63	31,0	83	21,7
4	22,4	24	21,0	44	24,1	64	25,0	84	28,3
5	15,1	25	34,2	45	20,3	65	23,2	85	26,0
6	23,2	26	32,5	46	20,1	66	24,1	86	22,4
7	16,7	27	31,0	47	19,7	67	24,1	87	22,4
8	19,1	28	31,0	48	21,7	68	29,6	88	26,0

№ п.п.	*V*, см/с	№ п.п.	*V*, см/с	№ п.п.	*V*, см/с	№ п.п.	*V*, см/с	№ п.п.	*V*, см/с
9	26,0	29	26,0	49	25,0	69	28,3	89	21,7
10	24,1	30	28,3	50	24,1	70	27,1	90	21,7
11	23,2	31	28,3	51	26,0	71	27,1	91	21,7
12	27,1	32	21,0	52	26,0	72	28,3	92	26,0
13	19,7	33	29,5	53	24,1	73	23,2	93	27,1
14	19,7	34	28,3	54	21,7	74	26,0	94	21,7
15	21,0	35	26,0	55	25,0	75	26,0	95	22,4
16	26,0	36	25,0	56	24,1	76	25,0	96	22,4
17	32,5	37	23,2	57	29,5	77	36,1	97	30,9
18	28,3	38	34,2	58	25,0	78	30,9	98	28,3
19	24,1	39	23,2	59	31,0	79	28,3	99	27,1
20	26,0	40	20,3	60	27,1	80	27,1	100	26,0

3. Визначається кількість m числа частинок піску, швидкості яких потрапили у відповідні інтервали (x_{m-1}, x_m).

4. Визначається статистична імовірність (частість) попадання швидкості осадження у відповідний інтервал

$$P_i = m_i / N . \qquad (1.17)$$

Результати розрахунків представлені в табл. 1.3.

Таблиця 2.3 – Визначення частості попадання

№ інтер-вала	Інтервали варіаційних рядів	Число попадань в інтервал m_i	Частість $P_i = m_i / N$	Середина інтервалу $\overline{x}_i$	Сумарна частість
1	15,1-18,1	2	0,02	16,6	0,02
2	18,1-21,1	12	0,12	19,6	0,14
3	21,1-24,1	23	0,23	22,6	0,37
4	24,1-27,1	31	0,31	25,6	0,68
5	27,1-30,1	20	0,20	28,6	0,88
6	30,1-33,1	9	0,09	31,6	0,97
7	33,1-36,1	3	0,03	34,6	1,00

За результатами розрахунків побудуємо гістограму розподілу швидкості осадження частинок піску (рис. 1.4). З рисунка видно, що найбільш вірогідне значення швидкості осадження – це 25,0-26,0 см/с (25,6 см/с). Статистична імовірність появи даної випадкової величини складає 31,0%.

Визначимо параметри статистичного розподілу:

- математичне очікування $\bar{x}$ і стандартне відхилення S по формулах (1.1) и (1.5)

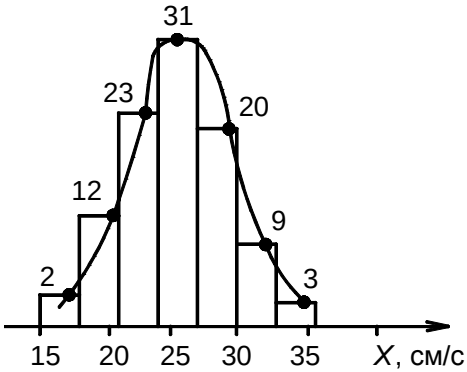

Рис. 1.4 – Гістограма розподілу швидкості осідання часток піску

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

$$S = \sqrt{\frac{1}{N-1}\sum_{i=1}\left(x_i - \bar{x}\right)^2} = \sqrt{\sum_{i=1}\left(x_i - \bar{x}\right)^2 P_i}\,.$$

$$\bar{x} = \sum_{i=1}^{n} \bar{x}_i P_i = \ 0{,}02\cdot 16{,}6 + 0{,}12\cdot 19{,}6 + 0{,}23\cdot 22{,}6 +$$

$$+0{,}31\cdot 25{,}6 + 0{,}20\cdot 28{,}6 + 0{,}09\cdot 31{,}6 + 0{,}03\cdot 34{,}6 = 25{,}4.$$

$$\bar{x} = 25{,}4 \ \text{ см/с.}$$

$$S = \sqrt{\begin{array}{l} 0{,}02\cdot\left(16{,}6-25{,}4\right)^2 + 0{,}12\cdot\left(19{,}6-25{,}4\right)^2 + \\ +0{,}23\cdot\left(22{,}6-25{,}4\right)^2 + 0{,}31\cdot\left(25{,}6-25{,}4\right)^2 + \\ +0{,}20\cdot\left(28{,}6-25{,}4\right)^2 + 0{,}09\cdot\left(31{,}6-25{,}4\right)^2 + \\ +0{,}03\cdot\left(34{,}6-25{,}4\right)^2 \end{array}} = \sqrt{15{,}715} = 3{,}96$$

$$S = 3{,}96 \ \text{ см/с} \approx 4{,}0 \text{ см/с.}$$

Згідно правила трьох сигм (3σ), математичне очікування швидкості осідання лежить в межах:

$\bar{x} - \frac{3S}{\sqrt{n}} \le \bar{X} \le \bar{x} + \frac{3S}{\sqrt{n}}$, тобто в нашому випадку 24,2 $\le \bar{X} \le$ 26,6.

Величину математичного очікування можна оцінити і за формулою (1.15).

Порівняння дисперсій і середніх значень

Порівняння дисперсій

Порівняння двох дисперсій нормальних генеральних сукупностей

Приклад 2.8. *По двох незалежних вибірках* n_1 = 10 і n_2 = 15, витягнутим з нормальних сукупностей X_1 и X_2, знайдені виправлені вибіркові дисперсії $S^2_{x_1}$ = 0,86 и $S^2_{x_2}$ = *0,43. При* q = 0,05 перевірити H_o: $D(X_1) = D(X_2)$ *при конкуруючій гіпотезі* H_1: $D(X_1) > D(X_2)$.

Рішення: $F_{спостер.}$ = 0,86 / 0,43 = 2,0, оскільки H_1: $D(X_1) > D(X_2)$ – критична область правостороння. За Додатком F $F_{кр.}$ = 2,65 при q = 0,05; k_1 = 10 – 1 = 9; k_2 = 15 – 1 = 14.

Оскільки $F_{спостер.} < F_{кр.}$, то вибіркові генеральні дисперсії розрізняються незначущо.

Приклад 2.9. Щільність зерен в шматку шихти, г/м^3, узятої з двох незалежних вибірок X_1 і X_2, приведена нижче:

x_1	2,41	2,38	2,43	2,50	2,44	2,41	2,52
x_2	2,65	2,70	2,73	2,66	2,63	2,71	2,73

При q = 0,1 перевірити H_o: $D(X_1) = D(X_2)$ при конкуруючій гіпотезі H_1: $D(X_1) \ne D(X_2)$.

Рішення: За формулою (1.1) визначаємо $\bar{x}_i$, а за формулою (1.5) – S_i. Тоді $\bar{x}_1$ = 2,44; $\bar{x}_2$ = 2,69; S_1^2 = 0,05; S_2^2 = 0,04. При цьому $F_{набл.}$ = 0,05 / 0,04 = 1,25. Критична область двостороння, оскільки конкуруюча гіпотеза має вигляд $D(X_1) - D(X_2)$. За цих умов за правилом № 2 [3, с. 104] при k_1 = 7 – 1 = 6 і k_2 = 7 – 1 = 6 і q_1= q / 2 = = 0,1 / 2 = 0,05 та за додатком F маємо $F_{кр}$ = 4,28 > $F_{спостер.}$, тобто гіпотезу H_1 відкидаємо, оскільки обидва вимірювання були виконані з однаковою точністю. Можна переходити до порівняння середніх (див. приклади 1.12 и 1.14).

Порівняння виправленої вибіркової дисперсії з гіпотетичною генеральною дисперсією нормальної сукупності

Приклад 2.10. По трьох незалежних вибірках однакового об'єму n = 11 з нормальних генеральних сукупностей визначені вибіркові дисперсії: S_1^2 = 0,25; S_2^2 = 0,36; S_3^2 = 0,49.

Потрібно: а) при рівні значущості q = 0,05 перевірити нульову гіпотезу про однорідність дисперсій; б) оцінити генеральну дисперсію; в) розрахувати інтервальні оцінки дисперсій.

Рішення: А. Розрахункове значення критерію Кохрена:

$$G_{набл.} = S_{\max}^2 / \sum_{i=1}^{l} S_i^2 = \frac{0{,}49}{(0{,}25+0{,}36+0{,}49)} = 0{,}4454\,. \qquad (1.18)$$

За додатком Н при рівні значущості q = 0,05, числі мір свободи k = 11 – 1 = 10 і числі вибірок l = 3 критична точка $G_{кр.}$(0,05; 10; 3) = 0,6025. Оскільки ($G_{набл.}$ < $G_{кр}$), вважаємо, що виправлені вибіркові дисперсії розрізняються незначуще (тобто дисперсії однорідні).

Б. Враховуючи однорідність дисперсій S_1^2, S_2^2 и як оцінку генеральної дисперсії приймаємо середню арифметичну виправлену дисперсію:

$$D_r = (0{,}25+0{,}36+0{,}49)/3 \approx 0{,}37\,. \qquad (1.19)$$

В. За додатком G $\chi^2_{q/2}$ = 20,5, а $\chi^2_{1-q/2}$ = $\chi^2_{0{,}975}$ = 3,25. Тоді інтервальну оцінку дисперсії можна визначити з умови

$$\frac{S_x^2 \cdot k}{\chi^2_{q/2}} \le \sigma_x^2 \le \frac{S_x^2 \cdot k}{\chi^2_{1-q/2}}\,, \qquad (1.20)$$

де σ_x^2 – дисперсія змінної *X;* $\chi^2_{q/2}$ – значення χ^2 – розподілу для рівня значущості *q*/2; $\chi^2_{1-q/2}$ – значення χ^2 – розподілу при рівні значущості (1 – *q*/2); $k = n - 1$ – число ступенів свободи; S_x^2 – виправлена вибіркова дисперсія.

$$\frac{0{,}25\cdot 10}{20{,}5} \le \sigma_{x1}^2 \le \frac{0{,}25\cdot 10}{3{,}25};\quad 0{,}12 \le \sigma_{x1}^2 \le 0{,}77;\ 0{,}18 \le \sigma_{x2}^2 \le 1{,}11;\ 0{,}24 \le \sigma_{x3}^2 \le 1{,}51.$$

Порівняння середніх

Порівняння двох середніх генеральних сукупностей, дисперсії яких відомі (великі незалежні вибірки n_i > 30)

Приклад 1.11. Є дві незалежні вибірки n_1 = 50 і n_2 = 60, витягнуті з нормальних генеральних сукупностей концентрації зважених речовин в річці до і після населеного пункту. Знайдені вибіркові середні: $\bar{x}_1$ =10 мг/л і $\bar{x}_2$ = 13 мг/л. Генеральні дисперсії відомі: $\sigma^2_{1(x_1)}$ = 8, $\sigma^2_{1(x_2)}$ = 10. Потрібно при рівні значущості q = 0,01 перевірити нульову гіпотезу H_0: $M(x_1) = M(x_2)$ проти гіпотези:

а) H_1: $M(x_1) \neq M(x_2)$;

б) H_1: $M(x_1) < M(x_2)$.

Рішення: Знайдемо спостережуване значення критерію:

$$z_{набл.} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma^2_{(x_1)}}{n_1} + \frac{\sigma^2_{(x_2)}}{n_2}}} = \frac{10-13}{\sqrt{\frac{8}{50} + \frac{10}{60}}} = -5{,}25\,. \qquad (1.21)$$

За умовою (*а*) конкуруюча гіпотеза має вигляд H_1: $M(x_1) \neq M(x_2)$, тому критична область двостороння. Знайдемо праву критичну точку з рівності

$$\Phi(z_{кр}) = (1-q)/2 = (1-0{,}01)/2 = 0{,}495\,. \qquad (1.22)$$

За таблицею функції Лапласа (додаток В) знаходимо $z_{кр}$ = 2,58. Оскільки $|z_{спостер.}| > z_{кр}$, то, відповідно до правила 1, нульову гіпотезу відкидаємо, тобто вибіркові середні розрізняються значущо.

По умові (*б*) конкуруюча гіпотеза має вид H_1: $M(x_1) < M(x_2)$. Знайдемо допоміжну точку з рівності

$$\Phi(z_{кр}) = (1-2q)/2 = (1-2\cdot 0{,}01)/2 = 0{,}490\,. \qquad (1.23)$$

За таблицею Лапласа (додаток А) знаходимо $z_{кр}$ = 2,33. Оскільки $z_{спостер.}$= –5.25 < –2,33, то за правилом 3 нульову гіпотезу відкидаємо і вважаємо, що $M(x_1) < M(x_2)$.

Після знаходження $z_{спостер.}$ за формулою (1.21) задаємо бажану імовірність виведення *P* і за нею знаходимо в Додатку В відповідне значення *t*(*P*) (наприклад, при *P* = 0,99 знаходимо *t* = 2,576, а при *P* = 0,95 визначаємо *t* = 1,960). Тоді, якщо $|z_{спостер.}|$ > *t*(*P*), то

розбіжність середніх значень можна вважати значущою з надійністю виведення *P*.

Порівняння двох середніх нормальних генеральних сукупностей, дисперсії яких невідомі і однакові (малі незалежні вибірки n_i < 30)

Приклад 2.12. Дані дві незалежні малі вибірки n_1 = 16 і n_2 = 20 з нормальної генеральної сукупності X_1 і X_2 з вибірковими середніми $\bar{x}_1$ = 42 і $\bar{x}_2$ = 40 та виправленими дисперсіями $S_{x_1}^2$ = 0,70 і $S_{x_2}^2$ = = 0,37. Потрібно при рівні значущості q = 0,05 перевірити гіпотезу H_0: $M(x_1) = M(x_2)$ проти альтернативної гіпотези:

а) H_1: $M(x_1) \neq M(x_2)$; б) H_1: $M(x_1) > M(x_2)$.

Рішення: Виправлені дисперсії різні, тому заздалегідь перевіримо гіпотезу про рівність генеральних дисперсій, використовуючи критерій Фішера (див. приклади 1.8, 1.9).

Знайдемо відношення більшої дисперсії до меншої:

$$F_{розр.} = 0{,}70 / 0{,}37 = 1{,}89.$$

Враховуючи, що дисперсія $S_{x_1}^2$ значно більша $S_{x_2}^2$, приймаємо як конкуруючу гіпотезу H_1: $D(x_1) > D(x_2)$. Тут критична область - правостороння. За додатком F, при рівні значущості q = 0,05 і мірах свободи $k_1 = n_1 - 1 = 16 - 1 = 15$ і $k_2 = n_2 - 1 = 20 - 1 = 19$ знаходимо критичну точку $F_{кр}$ (0,05; 15; 19) = 2,23.

Оскільки $F_{розр.} < F_{кр}$, то приймається гіпотеза H_0 про рівність генеральних дисперсій. Якщо гіпотеза про рівність дисперсій підтверджується, то можна порівняти середні.

Визначимо розрахункове значення критерію Ст'юдента по формулі

$$T = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(n_1 - 1)\cdot S_{x1}^2 + (n_2 - 1)\cdot S_{x2}^2}} \cdot \sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 - 2)}{n_1 + n_2}}. \qquad (1.24)$$

Якщо $$|T_{спост.}| < t_{двост.кр.}(q; k = n_1 + n_2 - 2), \qquad (1.25)$$

то гіпотеза H_0 приймається, якщо не виконується, то приймається конкуруюча гіпотеза H_1.

$$T_{розр.} = \frac{42 - 40}{\sqrt{(16-1)\cdot 0{,}70 + (20-1)\cdot 0{,}37}} \cdot \sqrt{\frac{16 \cdot 20 \cdot (16 + 20 - 2)}{16 + 20}} = 8{,}30.$$

За умовою (а) конкуруюча гіпотеза має вид H_1: $M(x_1) \neq M(x_2)$, тому при рівні значущості 0,05 і числі мір свободи

$k = n_1 + n_2 - 2)$ = 16 + 20 – 2 = 34 знаходимо за Додатком критичну точку $t_{двост.кр.}$ (0,05; 34) = 2,032. Оскільки $T_{розр.}$ = 8,3 > $t_{двост.кр.}$. = = 2,032, тобто умова (1.25) не виконується, то вважаємо, що значення $\bar{x}_1$ і $\bar{x}_2$ відрізняються значущо.

За умовою (б) конкуруюча гіпотеза має вид H_1: $M(x_1) > M(x_2)$, тому при q =- 0,05 і k = 34 знаходимо за Додатком Е критичну точку $t_{правост.кр.}$ = 1,69. Оскільки умова (1.25) не виконується, вважаємо, що $\bar{x}_1$ дійсно більше $\bar{x}_2$.

Перевірка гіпотези про рівність середніх

Методи порівняння середніх значень, викладені в прикладах 1.11 і 1.12, застосовні і для вирішення задачі перевірки гіпотези про рівність середніх.

Порівняння вибіркової середньої з гіпотетичної генеральної середньої нормальної сукупності (дисперсія генеральної сукупності невідома)

Приклад 2.13. По вибірці об'ємом n = 6, витягнутої з нормальної генеральної сукупності, знайдене вибіркове середнє межі міцності при стискуванні зразків-балочок $\bar{x}$ = 431,1 кгс/см^2 і виправлене середнє квадратичне відхилення S = 4,71. Потрібний при рівні значущості q = 0,05 перевірити гіпотезу H_0: $a = a_0$ = 435 при конкуруючій гіпотезі H_1: $a \neq a_0$ = 435.

Рішення: Находимо критерій перевірки T за формулою

$$T = (\bar{x} - a_0)\sqrt{n} / S , \qquad (1.26)$$

де S – виправлене середнє квадратичне відхилення, що визначається за формулою (1.5), має розподілення Ст'юдента з $k = n - 1$ ступенями свободи.

$$T = \frac{(431{,}1 - 435{,}0)\sqrt{6}}{4{,}71} = -2{,}03 .$$

За додатком Е при $k = n - 1$= 6 – 1 = 5 находимо критичну точку $t_{двост.кр.}$(0,05; 5) = 2,57. Оскільки виконується умова

$$T_{спост.} < t_{двост.кр.}(q;\, k = n - 1) , \qquad (1.27)$$

то гіпотеза H_0: $a = a_0$, приймається, тобто приходимо до висновку, що вибіркове среднє $\bar{x}$ = 431,1 трохи відрізняється від гіпотетичної генеральної середньої a_0 = 435.

Порівняння двох середніх нормальних генеральних сукупностей з невідомими дисперсіями (залежні вибірки)

Приклад 2.14. Знайдені наступні межі міцності при розколюванні двох складів шихти, МПа:

x_1	20,1	19,9	20,3	18,3	21,0	20,7	20,3
x_2	22,7	23,0	22,9	22,9	21,3	22,4	21,0

При рівні значущості q = 0,05 встановити, значущо або незначущо розрізняються середні значення межі міцності. Передбачається, що розподіл нормально.

Рішення: Знайдемо різницю, віднімаючи з чисел першого рядка числа другої. Одержимо:

d_i: - 2,6; - 3,1; - 2,6; - 4,5; - 0,3; - 1,7; - 0,7.

Знайдемо вибіркову середню, враховуючи, що $\sum d_i$ = -15,5; $\bar{d}$ = -15,5 / 7 = -2,2. Тоді виправлене середнє квадратичне відхилення буде рівне

$$S_d = \sqrt{\frac{\sum d_i^2 - \left[\sum d_i\right]^2 / n}{n-1}} = 1{,}44\,, \tag{1.28}$$

а спостережуване значення критерію $T_{спост}$

$$T_{спост} = \bar{d}\sqrt{n} / S_d = -2{,}2\sqrt{7} / 1{,}44 = -4{,}03\,. \tag{1.29}$$

По додатку Е знаходимо критичну точку

$$t_{двост.кр.}(0{,}05;\ 6) = 2{,}45 < \left|T_{спост.}\right| = -4{,}03\,. \tag{1.30}$$

Так як умова $\left|T_{спост.}\right| < t_{двост.кр.}$. не виконується, вважаємо, що середні значення вимірювань межі міцності відрізняються значущо.

Перевірка гіпотези нормальності закону розподілу випадкової величини

Приклад 2.15. З урахуванням вихідних даних, приведених в прикладі 2.3, і подальших розрахунках в прикладі 1.4, маємо: n = 40; l = 7; A_x = -0,494; C_x = *-0,615;* A_x^* = -0,512; C_x^* = -0,497; x = 9,3; S^* = 3,30; S = 3,25. Потрібно перевірити гіпотезу нормальності розподілу випадкової величини наближеним методом.

Рішення: Помилку оцінки показника асиметрії A_x визначимо за формулою

$$S_{Ax} = \sqrt{\frac{6\cdot(n-1)}{(n+1)\cdot(n+3)}} = \sqrt{\frac{6\cdot(40-1)}{(40+1)\cdot(40+3)}} = \sqrt{\frac{6{,}39}{41{,}43}} = 0{,}133\,, \qquad (1.31)$$

а помилку оцінювання C_x – за формулою

$$S_{Cx} = \sqrt{\frac{24\cdot(n-2)\cdot(n-3)}{(n-1)^2\cdot(n+3)\cdot(n+5)}} = \sqrt{\frac{24\cdot 38\cdot 37}{39^2\cdot 43\cdot 45}} = 0{,}107\,. \qquad (1.32)$$

Оскільки

$$\frac{|A_x = -0{,}494|}{(2...3)} \approx (0{,}165...0{,}267) > S_{Ax} = 0{,}133\,,$$

а

$$\frac{|C_x = -0{,}615|}{(2...3)} \approx (0{,}205...0{,}308) > S_{Cx} = 0{,}107\,,$$

то можна поставити під сумнів нормальність закону розподілу досліджуваної величини.

Застосовний критерій χ^2, для чого вихідні дані і частину розрахунків представимо в табличній формі.

Для застосування критерію χ^2 об'єднаємо крайні інтервали, щоб число даних в кожному інтервалі було не менше п'яти. Одержані дані представлені в перших двох стовпцях табл. 1.4. Крайні інтервали узяті нескінченними. У третьому стовпці підраховані відношення

$$t_i = \frac{x_i - x}{S} = \frac{x_i - 9{,}3}{3{,}25}\,. \qquad (1.33)$$

Таблиця 1.4 – Перевірка нормальності закону розподілення

Інтервали	m_i	t_i	$\Phi(t_i)$	p_i	$m_i - np_i$	$(m_i - np_i)2 / np_i$
-∞ ... 5	5	-1,323	-0,4071	0,0929	1,284	0,444
5 ... 7	5	-0,708	-0,2605	0,1466	-0,864	0,127
7 ... 9	6	-0,092	-0,0367	0,2238	-2,952	0,973
9 ... 11	10	0,523	0,1991	0,2358	0,568	0,034
11 ... 13	9	1,138	0,3725	0,1734	2,064	0,614
13 ... +∞	5	∞	0,5000	0,1275	-0,100	0,002
	Σ 40			1,0000		2,194 = χ_p^2

Так, для правого кінця, наприклад, першого інтервалу, це буде t_1 = (5 – 9,3) / 3,25 = - 1,323. У четвертому стовпці приведені відповідні значення інтеграла імовірності $\Phi(t_i)$ з додатку А. По значенням $\Phi(t_i)$ в п'ятому стовпці обчислена імовірність p_i, як різниці відповідних значень, наприклад $\Phi(t): p_i = \Phi(t_i) - \Phi(t_{i-1})$, p_2 = -0,2605 – (-0,4071) = = 0,1466. При обчисленні імовірності p_1 враховано, що $\Phi(-\infty)$ = - 0,5. Порівняння χ^2_p = 2,194 с $\chi^2_{кр.}$, одержаних при числі мір свободи k = 6 – 3 = 3 і рівні значущості q = 0,05, показує, що немає підстави сумніватися в нормальності розподілу, оскільки χ^2_p = 2,194< $\chi^2_{кр.}$ = 7,8.

Визначення теоретичного закону розподілу

Приклад 1.16. Для умов прикладу 1.6 визначити теоретичний закон розподілу швидкості осадження зерен піску фільтруючого матеріалу.

Попередній аналіз одержаної гістограми показує, що теоретичним законом розподілу може бути нормальний закон (закон Гауса).

Для перевірки гіпотези про те, що швидкість осадження підпорядкована нормальному закону розподілу, використовуємо один з найбільш поширених критеріїв – критерій Пірсону (χ^2), який визначається за формулою:

$$\chi^2 = \sum_{n=1}^{k} \frac{(m_i - nP_i)^2}{nP_i}, \qquad (1.34)$$

де m_i – число попадань у відповідний інтервал; n – об'єм вибірки (у нашому випадку n = 100); P_i – імовірність попадання в *і-й* інтервал; k – кількість інтервалів.

Критерій χ^2 передбачає порівняння кількості елементів, що потрапили в і-й інтервал з вибірки об'ємом n з кількістю елементів, які в середньому потрапляли б в цей інтервал з вибірки того ж об'єму, у разі вірності висунутої нами гіпотези про закон розподілу. Іншими словами, порівнюється відносна частота попадання в інтервал з гіпотетичною ймовірністю попадання в нього.

Критерій Пірсону має число мір свободи, рівне $n = k - l - 1$, де l – кількість оцінюваних параметрів в законі розподілу.

При нормальному законі розподілу оцінюється дисперсія і математичне очікування, тобто l = 2.

При використанні критерію χ^2 необхідно враховувати, що інтервали з кількістю елементів менше 10 необхідно об'єднувати з сусідніми. При цьому кількість мір свободи, природно, зменшується.

Раніше одержаний варіаційний ряд (табл. 1.3) перетворюється в новий (табл. 1.5).

Таблиця 1.5 – Перетворений варіаційний ряд

№ інтервалу	**Інтервали варіаційних рядів**	**Число попадань в інтервал, *m***	**Частість, $P_i = m_i / n_i$**
1	15,1-21,1	14	0,14
2	21,1-24,1	23	0,23
3	24,1-27,1	31	0,31
4	27,1-30,1	20	0,20
5	30,1-36,1	12	0,12

Розрахунки, пов'язані з перевіркою гіпотези про нормальний розподіл, проведемо в табл. 1.6.

Таблиця 2.6 – Перевірка гіпотези нормальності розподілу

m_i	t_{i-1}	t_i	$\Phi(t_{i-1})$	$\Phi(t_i)$	$P_i = \Phi(t_i) - \Phi(t_{i-1})$	$n \cdot \overline{P}$	$m_i - n\overline{P}$	$\left(m_i - n\overline{P}\right)^2$	$\left(m_i - n\overline{P}\right)^2 / nP$
14	-2,60	-1,09	-0,495	-0,362	0,133	13,3	0,7	0,49	0,037
23	-1,09	-0,33	-0,362	-0,133	0,229	22,9	0,1	0,01	0,00
31	-0,33	0,43	-0,133	0,166	0,299	29,9	1,1	1,21	0,04
20	0,43	1,19	0,166	0,383	0,217	21,7	-1,7	2,89	0,133
12	1,19	2,70	0,383	0,496	0,113	11,3	0,7	0,49	0,043
100						Σ=99,1			Σ = 0,253

Тут t_i і t_{i-1} – нормовані випадкові величини, визначувані за формулою:

$$t = (x_i - \overline{x}) / \sigma ,$$

Наприклад, (15,1 – 25,4) / 3,96 = -10,3 / 3,96 = -2,6;
(21,1 – 25,4) / 3,96 = -4,3 / 3,96 = -1,09,

x_i; x_{i-1} – нормовані межі інтервалів варіаційних рядів; $\Phi(t_i)$; $\Phi(t_{i-1})$ – значення нормованих функцій Лапласа від t_i і t_{i-1}, визначаються по таблицях (додатки А і В); $P_i = \Phi(t_i) - \Phi(t_{i-1})$ – гіпотетична імовірність попадання випадкової величини в інтервал (t_{i-1} – t_i); nP – гіпотетична частота попадання випадкової величини у відповідний інтервал.

Порівняння розрахункової величини критерію χ^2 = 0,253 з критичним його значенням $\chi^2_{кр.}$ = 5,99, одержаним для заданого рівня значущості q = 5% і числа мір свободи n = 5 – 2 – 1 =2 говорить про те, що $\chi^2 < \chi^2_{кр.}$ і що швидкість осадження зерен фільтруючого матеріалу має нормальний розподіл імовірності генеральної сукупності, який можна виразити наступною формулою:

$$y = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x_i - \overline{x})^2}{2\sigma^2}\right]. \qquad (1.35)$$

Після підстановки $\overline{x}$ = 25,4 і σ = 3,96 отримаємо

$$y = \frac{1}{3{,}96\sqrt{2 \cdot 3{,}14}} \exp\left[-\frac{(x_i - 25{,}4)^2}{2 \cdot 3{,}96^2}\right] = 0{,}1 \cdot \exp\left[-0{,}032 \cdot (x - 25{,}4)^2\right],$$

де y – щільність імовірності розподілу випадкової величини; x – випадкова величина (швидкість осадження).

Обробка результатів наукових досліджень методами кореляційного і регресійного аналізів

Обробка результатів експерименту

Експеримент проводиться відповідно до вибраної матриці планування, або плану. План є множина з N точок *К-мірного* факторного простору.

На цьому просторі діє декілька функцій, що описують *поведінку відгуків* плану. Значення відгуків за допомогою кодування за шкалою бажаності зводяться до *одного* значення узагальненого відгуку.

Таким чином, внаслідок проведення експерименту N точкам факторного простору, співставляється N експериментальних точок узагальненої функції відгуку.

Сама функція невідома і представлена внаслідок проведення одного кроку планування своїми N значеннями. З цих N значень вибирається найбільш оптимальне. Вибір відбувається відповідно до перевірки гіпотез *про відмінність* двох середніх і *однорідності дисперсій*.

Отже, вибрана одна точка факторного простору, що доставляє *локальний оптимум* на безлічі N точок факторного простору. Постулювавши *гладкість і безперервність* функції відгуку, ми можемо підібрати таку околицю цієї оптимальної точки, що функція узагальненого відгуку наближатиметься гіперплощиною $\Psi(x_1, x_2,\ldots, x_K)$, заданою на факторному просторі. Різниці

$$\xi_i = y_i - \Psi_i, \quad i = 1,2,\ldots,N$$

називаються *нев'язністю*.

Величина нев'язності визначається як <u>помилкою</u> *експерименту*, так і *придатністю* лінійної моделі, пов'язаної величиною області експерименту. І обидві ці причини змішані.

Ступінь ефективності наближення функції узагальненого відгуку гіперплощиною визначається за допомогою *оцінної функції*, заданої на безлічі нев'язності.

Як оцінні функції U використовують: суму квадратів або суму модулів нев'язності, або мінімальний критерій

$$U = \sum_{i=1}^{N} \xi_i^2, \quad U = \sum_{i=1}^{N} |\xi_i|, \quad U = \min \max |\xi_i|$$

В принципі можуть бути використані і інші оцінки, наприклад, сума модулів кубів нев'язності. Оцінки порівнюють по величині *погрішності коефіцієнтів* регресії і *простоті обчислень*. В цьому відношенні метод найменших квадратів найбільш оптимальний.

Метод найменших квадратів

Метод найменших квадратів приводить до мінімально можливої залишкової суми квадратів відгуку, або нев'язності, і в цьому сенсі є оптимальною оцінкою коефіцієнтів моделі.

Функцію відгуку у області експерименту ми представляємо у вигляді лінійного рівняння за факторами:

$$y = b_0 + \sum_{j=1}^{M} b_j x_j$$

Дане рівняння називається *рівнянням регресії*.

На кожному кроці планування експерименту відбувається обчислення коефіцієнтів b_j. Для кожного досліду рівняння регресії виконується приблизно, з точністю до нев'язності:

$$y_i - b_0 - \sum_{k=1}^{M} b_k x_{ki} = \xi_i$$

Метод найменших квадратів полягає в тому, що коефіцієнти моделі b_j вибирають з умови

$$U = \sum_{i=1}^{N} \xi_i^2 = \sum_{i=1}^{N} (y_i - b_0 - \sum_{k=1}^{M} b_k x_{ki})^2 = \min.$$

Для досягнення мінімуму функції $U(b_k)$ необхідно прирівняти всі приватні похідні нулю: $\partial U / \partial b_k = 0$. Після простих перетворень одержуємо

$$Nb_0 + \sum_{i=1}^{N}\sum_{k=1}^{M} b_k x_{ki} = \sum_{i=1}^{N} y_i \quad b_0 \sum_{i=1}^{N} x_{ki} + \sum_{i=1}^{N}\sum_{k=1}^{M} b_k x_{ki}^2 = \sum_{i=1}^{N} y_i x_{ki}.$$

Тепер прийшов час пригадати про чудові властивості матриці планування. Згідно властивості *симетричності* матриці планування сума, алгебри елементів будь-якого, вектор стовпця (будь-якого фактору) рівна нулю:

$$\sum_{i=1}^{N} x_{ki} = 0$$

А за умовою нормування

$$\sum_{i=1}^{N} x^2{}_{ki} = N$$

Крім того, ми скористаємося ортогональністю матриці планування, а саме

$$\sum_{i=1}^{N} x_{ki} x_{mi} = 0, \quad k \neq m, \quad k, m = 1,2,..,M.$$

З урахуванням цих властивостей одержуємо відомі вирази для коефіцієнтів моделі плану

$$b_0 = \frac{1}{N}\sum_{i=1}^{N} y_i, \quad b_k = \frac{1}{N}\sum_{i=1}^{N} y_i x_{ki}, \quad k = 1,2,..,M.$$

Формули для обчислення коефіцієнтів взаємодії мають аналогічний вигляд

$$b_{km} = \frac{1}{N}\sum_{i=1}^{N} y_i x_{ki} x_{mi}, \quad k \neq m, \quad k, m = 1,2,..,M.$$

Таким чином, прості формули обчислення коефіцієнтів регресії – є слідство застосування методу МНК до ортогональної, симетричної і нормованої матриці планування

Кореляційний аналіз

Під кореляцією розуміється всякий зв'язок між двома або декількома досліджуваними явищами. Кореляція може бути детерміністичною або випадковою (імовірнісною). Перший тип зв'язку визначається строгими закономірностями, які описуються фізико-хімічними формулами.

При кореляційному аналізі перевіряється лише самий факт зв'язку, тобто статистична гіпотеза про відсутність або наявність зв'язку. Сама природа величин, між якими такий випадковий зв'язок передбачається, дозволяє судити про нього як про імовірнісний. Результат кореляційного аналізу також носить статистичний характер, тому що висновок про наявність або відсутність зв'язку приймається за деякою наперед заданою довірчою імовірністю.

Прикладом задачі кореляційного аналізу може служити дослідження впливу температурного режиму на вихід якого-небудь хімічного продукту в складному технологічному процесі. При цьому зі збільшенням температури можливо не тільки підвищення швидкості досліджуваної реакції, але і протікання побічних реакцій, а також і зворотна реакція розкладання продукту. Тому зв'язок між температурою і виходом можна охарактеризувати як випадкову.

Лінійна кореляція

Як правило при кореляційному аналізі досліджуються тільки лінійні зв'язки між величинами, а статистичні критерії свідчать про наявність або відсутність передбачуваного лінійного зв'язку. Тому негативна відповідь при перевірці гіпотези про кореляцію може означати не тільки відсутність зв'язку, але і можливу наявність нелінійної залежності між досліджуваними величинами.

Для кількісної оцінки лінійної кореляції користуються вибірковим коефіцієнтом парної кореляції r_{xy} – безрозмірною величиною до значень середніх квадратичних відхилень досліджуваних величин:

$$r_{xy} = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\left(n\sum_{i=1}^{n} x_i^2 - \left[n\sum_{i=1}^{n} x_i\right]^2\right)\left(n\sum_{i=1}^{n} y_i^2 - \left[n\sum_{i=1}^{n} y_i\right]^2\right)}}.$$

Коефіцієнт кореляції за абсолютною величиною не перевершує одиниці ($|r_{xy}| \le 1$) і може приймати такі значення:

1) $r_{xy} = 0$ – цей випадок відповідає відсутності зв'язку між x і y (рис. 1.5, а);

2) $r_{xy} = +1$ – між x і y існує строгий позитивний лінійний зв'язок (рис. 1.5, б);

3) $r_{xy} = -1$ – між x і y існує строгий негативний зв'язок (рис. 1.5, в);

4) $-1 < r_{xy} < +1$ – це випадок, що найбільше часто зустрічається, і тут про кореляцію судять уже лише з точки зору більшої або меншої імовірності.

Розмір коефіцієнта кореляції $|r_{xy}|$ служить тільки для оцінки тісноти лінійного зв'язку між величинами x і y: чим ближче абсолютна величина коефіцієнта до 1, тим зв'язок сильніше; чим ближче $|r_{xy}|$ до нуля, тим зв'язок менше. Якщо випадкові величини x і y пов'язані точною лінійною функціональною залежністю

$$y = a \cdot x + b,$$

то $r_{xy} = \pm 1$. Знак «+» або «–» потрібно використати в залежності від знака коефіцієнту a ($a > 0$ або $a < 0$).

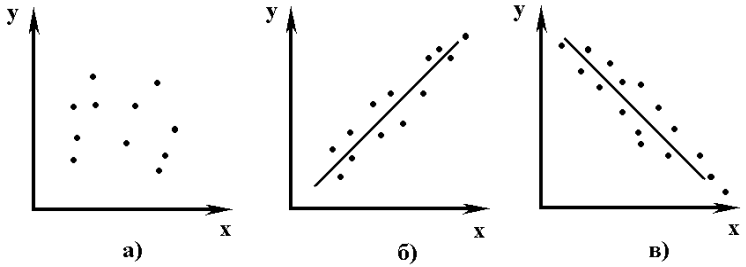

Рис. 1.5 – Кореляційна залежність між випадковими величинами x і y

Залежність коефіцієнта кореляції перевіряється шляхом порівняння абсолютної величини емпіричного коефіцієнта кореляції, помноженої на $\sqrt{n-1}$, із його критичними значеннями при заданому ступені надійності (рівня довірчості) P. Критичні значення $H = |r_{xy}|\sqrt{n-1}$ при числі вимірів n до 10 для різноманітних значень надійності складають:

при $P = 0{,}9$ $H = 1{,}65$;
при $P = 0{,}95$ $H = 1{,}90$;
при $P = 0{,}99$ $H = 2{,}29$.

Якщо для емпіричного коефіцієнта кореляції r_{xy} $H = |r_{xy}|\sqrt{n-1}$ виявиться більше критичного значення H, то з надійністю P слід відкинути гіпотезу про некорельованість аналізованих величин.

При інженерних розрахунках рівень довірчості P = 0,95 достатній.

Регресійний аналіз

Дослідження й оптимізація складних, неорганізованих систем можливі лише за допомогою статистичних, імовірнісних методів. Вихідною точкою для таких досліджень є аналог фізичної формули – *математичної моделі* системи, що носить назву *моделі експерименту* або *рівняння регресії*. Проте не завжди експериментальний матеріал дає можливість знайти зручний і точний вид моделі. У більш загальному випадку математична модель створюється на підставі статистичного методу – регресійного аналізу.

Регресійний аналіз – це дослідження регресійного рівняння – лінійної моделі плану. Дослідження проводиться в рамках методу найменших квадратів (МНК). Регресійний аналіз базується на трьох постулатах.

Перший постулат.

Параметр оптимізації y є *випадкова величина* з нормальним законом розподілу.

Для перевірки нормальності необхідно провести десятки паралельних дослідів. Далі користуються одним з критеріїв згоди.

Найбільшого поширення в практиці набув *критерій Пірсону*. Ідея цього методу полягає в контролі відхилень гістограми експериментальних даних від гістограми з таким же числом інтервалів, побудованої на основі розподілу, збіг з яким визначається, в даному випадку з нормальним розподілом.

При великому числі паралельних дослідів розмах відгуку наближається до інтервалу трьох сигм.

Хай, наприклад, в результаті n > 50-ти паралельних дослідів середнє значення відгуку, мінімальне і максимальні значення $\bar{y}$, $y_{\min}$, $y_{\max}$, обчислили оцінку дисперсії s^2. Вибрали m інтервалів, обчислили величину інтервалу Δ, кванта, і підрахували частотності попадання k_i, i = 1, 2,…, m значень відгуку в i-тий інтервал, тим самим підготували дані для побудови гістограми.

Для визначення теоретичних частотностей, наприклад, в MathCad

1) задаємо ранжирувану змінну i: = 0…m – 1;

2) визначаємо значення середин інтервалів $z_i = z_{min} + (1/2 + i)\Delta$;

3) обчислюємо вектор теоретичних частотностей за формулою $K_i = \mathrm{dnorm}(z_i, \bar{y}, s^2) \cdot n$.

Використання критерію Пірсону полягає в підрахунку величини хі-квадрат

$$\chi^2 = \sum_{i=1}^{m} \frac{(k_i - K_i)^2}{K_i} = \sum_{i=1}^{m} \frac{(k_i - nP_i)^2}{nP_i},$$

де k_i, K_i – експериментальні і теоретичні значення частот в і-му інтервалі розбиття, m – число інтервалів розбиття, P_i – значення імовірності в тому ж інтервалі розбиття, $n = \sum_{i=1}^{m} k_i$.

Число мір свободи змінної хі-квадрат в критерії Пірсону обчислюється в даному випадку за формулою $\upsilon = m - 3$, а не за формулою $n - 1$, як в критерії Бартлета. Дві міри свободи задіяні на обчисленні двох сум в чисельнику і знаменнику відношення критерію Пірсону.

Отже, задавшись рівнем значущості, частіше всього 0,05, звертаємося до функції квантилів хі-квадрат розподілу, що має в MathCad ім'я dchisq, з параметрами dshisq(1-0.05, m-3). Якщо обчислене за дослідними даними значення критерію Пірсону χ^2 менше ніж обчислене за функцією квантилів, то гіпотеза нормальності приймається, інакше відкидається.

При порушенні нормальності ми позбавляємося можливості встановлення імовірності, з якою справедливі ті або інші вислови щодо результатів експерименту.

Другий постулат. Дисперсія відгуку не залежить від його абсолютної величини. Перевірка здійснюється за допомогою перевірки дисперсії на однорідність із застосуванням критеріїв Фішера або Бартлета. Порушення цього постулату неприпустимо. При порушенні однорідності дисперсії вдаються до функціонального перетворення відгуку. Перетворення шукається методом проб і помилок. Звичайно починають з логарифма.

Сенс нелінійного перетворення в тому, що він підкреслює, або виділяє один інтервал зміни змінної перед іншим. Наприклад, квадратичне перетворення підкреслює великі величини, а малі (менші 1) ще більше зменшує. Логарифм, навпаки, «притримує» швидкість росту монотонно зростаючих величин, і тому логарифми дисперсій можуть вже виявитися однорідними, тоді як самі дисперсії неоднорідні.

Третій постулат. Значення факторів суть “невипадкові” величини. Це несподіване твердження означає, що встановлення кожного фактору на заданий рівень і його підтримку на цьому рівні істотно точніше, ніж помилка відтворюваності.

Четвертий постулат. Фактори не корельовано. Для матриці планування ця властивість виконується автоматично через властивість ортогональності.

Перевірка адекватності моделі

Перевірка адекватності моделі – це перевірка *придатності* обчислених після реалізації плану *коефіцієнтів* лінійної моделі.

На рис. 2.6 приведені два графіки з однаковими лініями регресії і однаковим розташуванням експериментальних точок (середніх значень) відносно лінії регресії, але з різною дисперсією. Розкид відносно середніх значень показується відрізками, що становлять довірчий інтервал $\pm t \cdot s_{\{y\}}$. Коефіцієнт перед стандартом визначається величиною довірчої імовірності.

Модель можна вважати адекватною тільки у випадку А. В даному випадку розкид в точках такого ж порядку, що і розкид відносно лінії.

У випадку Б досліди “дуже” точні. Потрібна складніша модель, нелінійна, щоб точність її прогнозу була порівнянна з точністю експерименту.

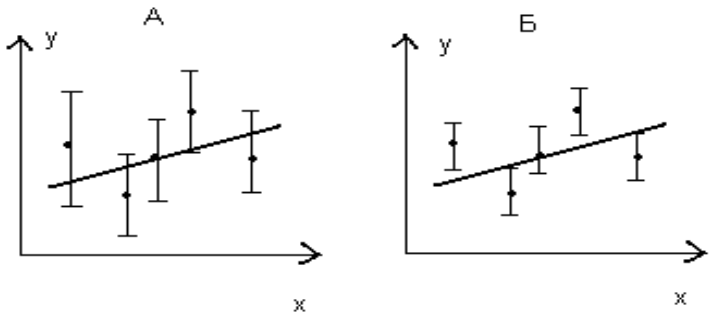

Рис. 1.6 – Перевірка адекватності

Для перевірки адекватності моделі обчислюють *залишкову дисперсію, або дисперсію адекватності* за формулами:

1) якщо немає паралельних дослідів то

$$s_{АД}^2 = \frac{\sum_{i=1}^{N} \Delta y_i^2}{f},$$

де $\Delta y_i = y_i - \sum_{k=0}^{M} b_k x_{ki}$ залишкова різниця є *i-та* нев'язність, M – число факторів, f – число мір свободи дисперсії адекватності, яке обчислюється за правилом: число різних дослідів мінус число визначуваних коефіцієнтів, рівне числу факторів +1:

$$f = N - (k + 1).$$

2) За наявності числа повторних дослідів n, рівного для всіх рядків плану, при розрахунку дисперсії адекватності обчислюють залишкову різницю відгуків за іншою формулою:

$$\Delta y_i = \overline{y}_i - \sum_{k=0}^{M} b_k x_{ki} = \overline{y}_i - \widehat{y}_i \,,$$

тобто обчислюють відхилення регресійної прямої від середнього значення паралельних дослідів.

Але найголовніше сама дисперсія адекватності множиться на число паралельних дослідів:

$$s_{АД}^2 = \frac{n\sum_{i=1}^{N} \Delta y_i^2}{f}$$

При n = 1 дану формулу переходить в першу формулу дисперсії адекватності.

Повторні досліди накладають жорсткіші умови на перевірку адекватності, і для ухвалення гіпотези адекватності потрібна більша відповідність експериментальних і розрахункових точок.

3) Найбільш загальний випадок є нерівномірне дублювання дослідів. В цьому випадку

$$s_{АД}^2 = \frac{\sum_{i=1}^{N} n_i (\overline{y}_i - \widehat{y}_i)^2}{f}$$

Для перевірки гіпотези про адекватність моделі використовують критерій Фішера (F-критерій) для відношення двох дисперсій: залишкової дисперсії до дисперсії відтворюваності

$$F = \frac{s_{АД}^2}{s_{\{y\}}^2}$$

Більше значення – в чисельнику, менше в знаменнику, тобто експериментальне значення критерію Фішера завжди більше 1.

Якщо експериментальне значення F не перевищує табличного значення, то гіпотеза адекватності моделі приймається, інакше її відкидають.

Якщо модель адекватна, тобто відношення дисперсій адекватності і відтворюваності не перевищує табличного, дисперсії не помітні по критерію Фішера, то можемо побудувати лінійну регресійну функцію на *к*-мірному факторному просторі, і переходити до наступного кроку планування, але цікаво перевірити значущість окремих коефіцієнтів регресії.

Перевірка значущості коефіцієнтів

Перевірка здійснюється або побудовою довірчого інтервалу, або за допомогою критерію Ст'юдента.

А) *побудова довірчого інтервалу*

Приймемо без доказу такий факт, що при використанні повного експерименту фактору або регулярних дробових реплік довірчі інтервали для всіх коефіцієнтів (у тому числі і ефектів взаємодії) рівні один одному.

Дисперсія будь-якого коефіцієнта регресії обчислюється за формулою

$$s_{\{b\}}^2 = \frac{s_{\{y\}}^2}{f}, \quad f = \sum_{i=1}^{N}(n_i - 1),$$

де $s_{\{y\}}^2$ – дисперсії відтворюваності, n_i – число паралельних дослідів в *i-тому* експерименті, N – число рядків матриці планування, f_i – число мір свободи при обчисленні дисперсії відтворюваності в і-му рядку плану.

Хай перший постулат регресійного аналізу про те, що відгуки суть випадкові величини з нормальним розподілом, виконаний в результаті перевірки по критерію Пірсону. Тоді коефіцієнти регресії також випадкові величини з нормальним розподілом. Розглянемо величину

$$\tau = \frac{1}{f}\frac{\sum_{i}^{N}\sum_{j}^{ni}(b_{i,j} - \beta_{i,j})}{S_{\{b\}}} = \frac{\Delta b}{S_{\{b\}}}$$

Величина τ має розподіл Ст'юдента.

Нагадаємо, що середнє арифметичне центрованих відносно середнього і нормованих на дисперсію випадкових величин з нормальної генеральної сукупності має розподіл Ст'юдента.

У цьому виразі відома, тільки одна величина – СКО коефіцієнта регресії. Другу величину – значення випадкової величини визначимо, задавшись рівнем значущості α. Для цього ми звертаємося до функції квантилів зворотного інтегрального розподілу Ст'юдента з аргументами $P = 1 - \alpha/2$, оскільки розподіл Ст'юдента симетрично відносно нуля, і f – числом мір свободи при обчисленні випадкової величини:

$$t = qt(1 - \alpha / 2, f).$$

Квантиль t – є величина напівінтервалу. Імовірність знаходження в інтервалі $[-t, t]$ для випадкової величини τ рівна $P = 1 - \alpha$. Тому довірчий інтервал Δb_j визначається як

$$\Delta b = \pm t \cdot s_{\{b\}},$$

P – довірча імовірність.

Формулу для довірчого інтервалу можна записати в наступній еквівалентній формі:

$$\Delta b = \pm \frac{t \cdot s_{\{y\}}}{\sqrt{f}}$$

Коефіцієнт значимо, якщо його абсолютна величина більше довірчого інтервалу.

Б) *використання критерію Ст'юдента*

Коефіцієнт Ст'юдента можна використовувати і безпосередньо. Розділимо останню нерівність на стандарт коефіцієнта регресії:

$$t = \frac{|\Delta b|}{s_{\{b\}}} < \frac{|b_j|}{s_{\{b\}}} \quad \Rightarrow \quad t < \frac{|b_j|}{s_{\{b\}}}$$

У лівій частині нерівності коштує табличне значення коефіцієнта Ст'юдента. У правій – експериментально обчислювані значення.

Якщо критерій Ст'юдента опинився менше відношення коефіцієнт СКО погрішності коефіцієнтів, то гіпотеза про те, що коефіцієнт значимо приймається з ймовірністю P.

Приклад. Хай в результаті повного експерименту фактору 2^3 помилка з двома паралельними дослідами в кожному рядку плану погрішність відтворюваності $s_{\{y\}} = 1$. Обчислені коефіцієнти: $b_0 = 10$, $b_1 = 2$, $b_2 = 1$, $b_3 = -0.5$. Побудувати довірчий інтервал і перевірити значущість коефіцієнтів.

Дисперсія коефіцієнтів регресії $s^2_{\{bj\}} = 1/8$, $s_{\{bj\}} = \sqrt{1/8} = 0{,}354$. Табличне значення критерію Ст'юдента при рівні значущості 0,05 і числі мір свободи $f = 8qt$ (0,975 · 8) = 2,3.

Таким чином, довірчий інтервал коефіцієнтів регресії $\Delta b_j = \pm 2{,}3 \cdot 0{,}354 = \pm 0{,}814$.

Отримали, що всі коефіцієнти регресії окрім b_3 значущі, оскільки їх абсолютна величина більше довірчого інтервалу, а для b_3 – вона менше.

Дисперсія передбаченого значення відгуку

Значення дисперсії коефіцієнта регресії можна використовувати для визначення дисперсії передбаченого відгуку по рівнянню регресії в деякій точці факторного простору $\overline{x}^*$.

Значення відгуку обчислюється в результаті підстановки цієї точки в рівняння регресії:

$$y_{N+1} = \widehat{y} = \sum_{i=0}^{N} b_i x_i^*$$

Дисперсія передбаченого відгуку обчислюється за формулою

$$s^2_{\{\widehat{y}\}} = \sum_{i=0}^{N} x_i^{*2} s^2_{\{bi\}}.$$

ПРИКЛАДИ

Оцінювання коефіцієнта кореляції

Вважається, що коефіцієнт кореляції r_{xy} є мірою лінійної взаємозалежності між величинами x і y. У інженерній практиці рекомендується приймати наступну міру цієї взаємозалежності [3, с. 118]:

r_{xy}:	> 0,9	0,8-0,9	0,6-0,8	0,4--0,6	0,2-0,4
Залежність:	функціональна	сильна (суттєва)	помірна	слабка	несуттєва

Приклад 1.17. Потрібно провести регресійний аналіз залежності часу досягнення граничної втрати натиску у фільтрі ($t_н$) від розміру зерен завантаження (d). Вихідні результати і розрахунки приведені в табл. 1.7.

Рішення: Відомо, що мірою лінійної залежності двох змінних є кореляційне співвідношення і коефіцієнт кореляції. Парний коефіцієнт кореляції можна оцінити за допомогою виразів:

- для малих вибірок

$$r = \sum\left[(x_i - \bar{x})\cdot(\bar{y}_i - \bar{\bar{y}})\right] / \sqrt{\sum(x_i - \bar{x})^2 \cdot \sum(\bar{y}_i - \bar{\bar{y}})^2}\ ; \qquad (1.36)$$

- для великих вибірок

$$r = \frac{\sum xy - (\sum x)\cdot(\sum y)/n}{\sqrt{\left[\sum x^2 - (\sum x)^2/n\right]\cdot\left[\sum y^2 - (\sum y)^2/n\right]}}\ . \qquad (1.37)$$

Таблиця 1.7 – Вихідні дані і проміжні розрахунки

(d, см) (x_i)	τ_1 (y_1)	τ_2 (y_2)	$\bar{\tau}$ ($\bar{y}_i$)	$x_i - \bar{x}$	$\bar{y}_i - \bar{\bar{y}}$	$(x_i - \bar{x})^2 \cdot 10^{-4}$	$(\bar{y}_i - \bar{\bar{y}})^2$	$(x_i - \bar{x})\cdot(\bar{y}_i - \bar{\bar{y}})$
0,120	16,7	16,6	16,65	-0,015	-3,84	2,25	14,74	0,0576
0,125	17,8	18,1	17,95	-0,010	-2,54	1,00	6,45	0,0254
0,130	19,3	19,4	19,35	-0,005	-1,14	0,25	1,30	0,0057
0,135	20,4	20,5	20,45	0,000	-0,04	0,00	0,00	0,0000
0,140	21,6	21,7	21,65	0,005	1,16	0,25	1,34	0,0058
0,145	23,0	23,1	23,05	0,010	2,56	1,00	6,55	0,0256
0,150	24,0	24,3	24,15	0,015	3,86	2,25	14,90	0,0579
$\bar{x}$ = = 0,135			$\bar{\bar{y}}$ = =20,49			Σ = 7,00	Σ = = 45,28	Σ = = 0,178

Вважається [3, с. 121], що краща оцінка коефіцієнта кореляції p виходить за формулою

$$r^* = r\cdot\left[1 + \frac{1 - r^2}{2\cdot(n-3)}\right]. \qquad (1.38)$$

За формулою (1.36), з урахуванням даних з табл. 1.7, маємо:

$$r = 0{,}178 / \sqrt{0{,}0007 \cdot 45{,}28} = 0{,}9998$$

або

$$r^* = 0{,}9998 \cdot (1 + (1 - 0{,}9998^2) / [2 \cdot (7-3)]) = 0{,}99985 .$$

Коефіцієнт детермінації B буде рівний $B = r^2 =$ 0,9996, а $(r^*)^2 =$ 0,9997.

Таким чином, залежність між часом досягнення граничних втрат натиску у фільтрі і розміром зерен завантаження дуже добре корельовано.

Після знаходження довірчих меж для m_z (при довірчій імовірності p = 0,95)

$$z_1 = z - 1{,}96 / \sqrt{n-3} \quad \text{і} \quad z_2 = z + 1{,}96 / \sqrt{n-3} \tag{1.39}$$

можна знайти довірчі межі для генерального коефіцієнта кореляції, підставляючи z_1 і z_2 у формулу:

$$r = thz = (e^{2z} - 1) / (e^{2z} + 1) \tag{1.40}$$

Тоді для даного прикладу маємо

- стандарт розподілу

$$\sigma_r \approx (1 - r^2) / \sqrt{n} \approx (1 - 0{,}9998^2) / \sqrt{7} = 0{,}00015 ; \tag{1.41}$$

- довірчі межі p

$$p = r \pm \frac{1{,}96 \cdot (1 - r^2)}{\sqrt{n}} = 0{,}9998 \pm \frac{1{,}96 \cdot (1 - 0{,}9998^2)}{\sqrt{7}} = 0{,}9998 \pm 0{,}0003 ; \tag{1.42}$$

$$z = \frac{1}{2} \ln \frac{1+r}{1-r} = \frac{1}{2} \ln \frac{1 + 0{,}9998}{1 - 0{,}9998} \approx 4{,}6 ; \tag{1.43}$$

$$\sigma_z = 1 / \sqrt{n-3} = 1 / \sqrt{7-3} = 0{,}5 . \tag{1.44}$$

Тоді з довірчою ймовірністю p = 0,95 значення m_z знаходиться в межах:

$$m_z = z \pm 1{,}96 / \sqrt{n-3} = 4{,}6 \pm 1{,}96 / \sqrt{7-3} = 4{,}6 \pm 0{,}98 ;$$

z_1 = 4,6 – 0,98 = 3,62 и z_2 = 4,6 + 0,98 = 5,58;

$$r_{x_1} = thz_1 = (e^{2 \cdot 3{,}62} - 1) / (e^{2 \cdot 3{,}62} + 1) = 0{,}99857 ;$$

$$r_{x_2} = thz_2 = (e^{2 \cdot 5{,}58} - 1) / (e^{2 \cdot 5{,}58} + 1) = 0{,}99997 ,$$

тобто при рівні значущості q = 0,05 величина вибіркового коефіцієнта кореляції знаходиться в межах 0,99857 < p < 0,99997.

Оцінювання прямої регресії

Для умов нашого прикладу проведемо аналіз шуканої залежності $\tau = f(d)$ (рис. 1.7). Як слідує з цього малюнка, результати дослідів по досягненню граничних втрат тиску в першому наближенні можна описати рівнянням лінійним типу:

$$\hat{y} = a_{yx} + b_{yx} \cdot x$$

або

$$\hat{x} = a_{xy} + b_{xy} \cdot y = a' + b_{xy} \cdot y \, .$$

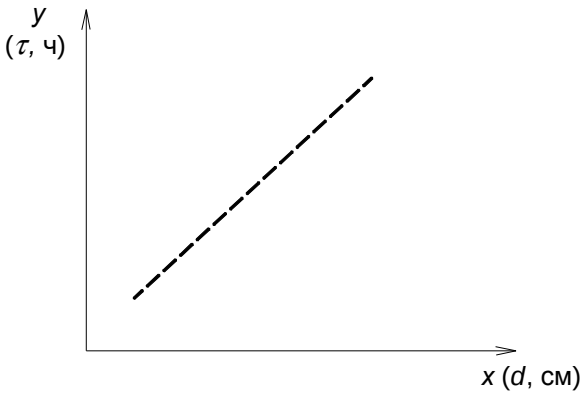

Рис. 1.7 – Графік залежності $\tau = f(d)$

Розрахунки, пов'язані з визначенням коефіцієнтів регресії, проведемо в табличній формі.

Таблиця 1.8 – Розрахункові дані

x	$\bar{y}$	$x\bar{y}$	x^2	$\bar{y}^2$
0,120	16,65	2,00	0,0144	277,2
0,125	17,95	2,24	0,0156	322,2
0,130	19,35	2,52	0,0169	374,4
0,135	20,45	2,76	0,0182	418,2
0,140	21,65	3,03	0,0196	468,7
0,145	23,05	3,34	0,0210	531,3
0,150	24,15	3,65	0,0225	592,9
$\sum x$ = 0,945	$\sum \bar{y}$ =143,45	$\sum x\bar{y}$ = 19,54	$\sum x^2$ =0,1282	$\sum \bar{y}^2$ =2984,5

Тоді:

$$b_{yx} = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2} = \frac{7 \cdot 19{,}54 - 0{,}945 \cdot 143{,}45}{7 \cdot 0{,}1282 - (0{,}945)^2} = 278{,}8\,; \qquad (1.45)$$

$$a_{yx} = \frac{\sum y - b_{yx}\sum x}{n} = \frac{143{,}45 - 278{,}8 \cdot 0{,}945}{7} = -17{,}14\,; \qquad (1.46)$$

$$b_{xy} = \frac{n\sum xy - \sum x \sum y}{n\sum y^2 - (\sum y)^2} = \frac{7 \cdot 19{,}54 - 0{,}945 \cdot 143{,}45}{7 \cdot 2948{,}5 - (143{,}45)^2} = 0{,}0039\,; \qquad (1.47)$$

$$a_{xy} = \frac{\sum x - b_{xy}\sum y}{n} = \frac{0{,}945 - 0{,}0039 \cdot 143{,}45}{7} = 0{,}0553\,. \qquad (1.48)$$

Оцінювання регресії $\hat{y}$ за x $\qquad \hat{y} = -17{,}14 + 278{,}8x$,

а оцінювання регресії $\hat{x}$ за y $\qquad \hat{x} = 0{,}0553 + 0{,}0039y$.

Для визначення координати точки перетину знайдених рівнянь регресії представимо друге рівняння у вигляді

$$y = (\hat{x} - 0{,}0553)/0{,}0039 = 256{,}4\hat{x} - 14{,}2 = b^*_{xy} \cdot x - a^*_{xy}\,,$$

тоді, прирівнюючи праві частини рівнянь, маємо:

$$256{,}4x - 14{,}2 = -17{,}4 + 278{,}8x;$$

$$x = x_0 = \frac{17{,}14 - 14{,}2}{278{,}8 - 256{,}4} = 0{,}131\,,$$

звідкіля $y_0 = 256{,}4 \cdot 0{,}131 - 14{,}2 \approx 19{,}39$.

У знайдених рівняннях регресії коефіцієнти b_{yx} і b^*_{xy} є кути нахилу відповідних прямих. Тоді кут між цими прямими

$$tg\alpha = tg(\alpha_y - \alpha_x) = \frac{b_{yx} - b^*_{xy}}{1 + b_{yx} \cdot b^*_{xy}} = \frac{b_{yx} - 1/b^*_{xy}}{1 + b_{yx}/b^*_{xy}} = \frac{b_{yx} \cdot b_{xy} - 1}{b_{xy} + b_{yx}} =$$

$$= \frac{278{,}8 - 256{,}4}{1 + 278{,}8 \cdot 256{,}4} \approx 0{,}000313\,,$$

тобто $\alpha < 12'$.

Перевіримо, наскільки адекватно одержане рівняння $y = -17{,}14 + 278{,}8x$ описує процес $\tau_н = f(d)$. Розрахунки, необхідні для визначення адекватності, проведемо в табличній формі.

Спершу перевіримо відтворюваність дослідів по критерію Кохрена [3, п. 5.4.10]. Розрахункове значення критерію складе

$$G_{набл.} = \frac{S^2_{\max}}{\sum S^2} = \frac{0{,}045}{0{,}075} = 0{,}6\,. \qquad (12.49)$$

Таблиця 1.9 – Розрахунки за визначенням адекватності моделі

x	Результати дослідів			S^2	y_p	$(y_p - \bar{y})^2$
	y_1	y_2	$\bar{y}$			
0,120	16,7	16,6	16,65	0,005	16,32	0,1089
0,125	17,8	18,1	17,95	0,045	17,71	0,0576
0,130	19,3	19,4	19,35	0,005	19,10	0,0625
0,135	20,4	20,5	20,45	0,005	20,50	0,0025
0,140	21,6	21,7	21,65	0,005	21,89	0,0576
0,145	23,0	23,1	23,05	0,005	23,29	0,0576
0,150	24,0	24,3	24,15	0,005	24,68	0,1089
				$\Sigma = 0{,}075$		0,4556

Табличне значення критерію Кохрена визначаємо за родатком Н. Прі рівні значущості q = 0,05 і числі мір свободи

$$k = m - 1 = 2 - 1 = 1.$$

$$G\ (0{,}05;\ l) = 0{,}7271.$$

Таким чином, умов $G_{спост.} \le G_{табл.}$ виконується, тобто дисперсії однорідні, і як генеральної дисперсії можна прийняти середню арифметичну виправлених дисперсій, або:

$$S^2_{відтвор.} = \sum_{i=1}^{N} S_i^2 / N = 0{,}075/7 \approx 0{,}011\,. \qquad (1.50)$$

Дисперсія адекватності визначається за формулою

$$S^2_{ад.} = \frac{m}{N-B} \cdot \sum_{i=1}^{N} \left(y_p - \bar{y}\right)^2 = \frac{2}{7-2} \cdot 0{,}4556 = 0{,}182\,, \qquad (1.51)$$

де B - число членів знайденого рівняння y_p = -17,14 + 278,8x.

Розрахункове значення крітерию Фішера

$$F_p = S^2_{ад.} / S^2_{відтвор.} = 0{,}182/0{,}011 = 16{,}56\,. \qquad (1.52)$$

При рівні значущості q = 0,05, числах мір свободи дисперсії адекватності $k_{ад.}$ = 7 – 2 = 5 і відтворюваності $k_{відтвор.}$ = $n(m - 1)$ = = 7 · (2 – 1) = 7 табличне значення критерію Фішера складе F(0,05; $к_{ад}$ = 5; $к_{відтвор}$ = 7) = 3,97 (додаток F).

Оскільки $F_p > F_m$, вважається, що рівняння y = - 17,14 + 278,8x неадекватно описує процес $\tau_н = f(d)$. Найбільш коректним рішенням в

ситуації, що склалася, ймовірно, можна рахувати перехід до рівняння іншого ступеня і підвищення точності вимірювання параметра y_i.

Приклад 1.18. Аналіз апріорно одержаної інформації показує, що при видобуванні нікелю із стічних вод електрохімічним способом існує залежність між ефектом видобування (кількістю витягуваного нікелю %) і щільністю струму. Необхідно визначити міру цієї залежності і її вигляд або математичний опис (рівняння регресії) виходу нікелю від щільності струму. Результати досліджень представлені в табл. 1.10.

Таблиця 1.10 – Вихідні та розрахункові дані

№ п/п	*Ефект, %*			**Щільність струму, *x***	$x_i - \bar{x}$	$\bar{y}_i - \bar{\bar{y}}$	$(x_i - \bar{x})^2 \times$ $\times 10^{-4}$	$(\bar{y}_i - \bar{\bar{y}})^2$	$(x_i - \bar{x}) \cdot$ $\cdot (\bar{y}_i - \bar{\bar{y}})$
	y_1	y_2	$\bar{y}$						
1	8,5	9,7	9,1	0,7	-0,56	-7,27	0,31	52,85	4,07
2	11,1	12,9	12,0	0,9	-0,36	-4,37	0,13	19,1	1,57
3	14,1	16,1	15,1	1,1	-0,16	-1,27	0,03	16,1	0,20
4	15,0	15,8	15,4	1,25	-0,01	-0,97	0,00	0,94	0,01
5	22,0	23,0	22,5	1,6	0,34	6,13	0,12	37,6	2,08
6	24,0	24,2	24,1	2,0	0,74	7,73	0,55	59,8	5,72
$\sum \bar{y} = 98{,}2$ $\bar{\bar{y}} = 16{,}37$				$\sum x = 7{,}55$ $\bar{x} = 1{,}26$			Σ = 1,11	Σ = = 186,39	Σ = = 13,65

Рішення: Коефіцієнт кореляції, визначуваний по вираженню (1.36), рівний:

$$r_{xy} = \frac{13{,}65}{\sqrt{1{,}11 \cdot 186{,}39}} = \frac{13{,}65}{14{,}38} = 0{,}95 .$$

У нашому випадку r_{xy} = 0,95, отже, залежність між щільністю струму x і ефектом видобування (виходом) нікелю по струму – *функціональна*.

Знайдемо математичний опис залежності виходу нікелю (ефекту електрохімічного видобування) від щільності струму. Спочатку проведемо графічний аналіз шуканої залежності. Для цього, за даними табл. 1.10, побудуємо графік $y = f(x)$ (рис. 1.8).

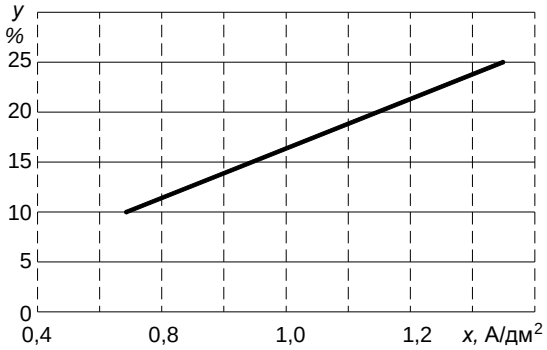

Рис. 1.8 – Залежність ефекту видобування нікелю (y) від щільності струму (x)

Таблиця 1.11 – Розрахунок коефіцієнтів регресії

№ п/п	y	x	x^2	xy	y_p
1	9,1	0,7	0,49	6,37	9,6
2	12,0	0,9	0,81	10,8	12,0
3	15,1	1,1	1,21	16,61	14,4
4	15,4	1,25	1,56	19,25	16,3
5	22,5	1,6	2,56	36,00	20,5
6	24,1	2,0	4,00	48,20	25,3
Σ	98,2	7,55	10,63	137,23	

Як видно з рис. 1.8, результати дослідів по видобуванню нікелю в першому наближенні можна описати рівнянням лінійним типу $y = a_0 + bx$. Для знаходження коефіцієнтів регресії a_0 і b скористаємося методом найменших квадратів.

Коефіцієнти рівняння регресії можна визначити за формулами (1.45) (1.46). Розрахунки, пов'язані з визначенням коефіцієнтів регресії, проведемо в табличній формі.

$$b = \frac{6 \cdot 137{,}23 - 98{,}2 \cdot 7{,}55}{6 \cdot 10{,}63 - 7{,}55^2} = 12{,}09\,;$$

$$a_0 = \frac{98{,}2 - 12{,}09 \cdot 7{,}55}{6} = 1{,}15\,.$$

Тоді рівняння регресії прийме вигляд: $y_p = 1{,}15 + 12{,}09x$.

Розрахунки, необхідні для визначення адекватності одержаного рівняння регресії, проведемо в табличній формі.

Таблиця 1.12 – Розрахунок адекватності рівняння регресії

№ п/п	x	Результати досліджень			S^2	y_p	$(y_p - \bar{y})^2$
		y_1	y_2	$\bar{y}$			
1	0,7	8,5	9,7	9,1	0,72	9,6	0,25
2	0,9	11,1	12,9	12,0	1,62	12,0	0,00
3	1,1	14,1	16,1	15,1	2,00	14,4	0,49
4	1,25	15,0	15,8	15,4	0,32	16,3	0,81
5	1,6	22,0	23,0	22,5	0,50	20,5	4,00
6	2,0	24,0	24,2	24,1	0,02	25,3	1,44
					$\Sigma = 5{,}18$		$\Sigma = 6{,}99$

Спочатку перевіримо відтворюваність наших дослідів за критерієм Кохрена

$$G_p = 2{,}00 / 5{,}18 = 0{,}386.$$

Табличне значення критерію G, визначене за Додатком Н при рівні значущості $q = 0{,}05$ і числі мір свободи $k = m - 1 = 2 - 1 = 1$, складає $G_m = 0{,}7808$. Оскільки $G_p < G_m$, досліди відтворні.

За формулою (1.50) дисперсія відтворюваності складе

$$S^2_{воспр.} = 5{,}18 / 6 = 0{,}86,$$

а дисперсія адекватності за формулою (1.51)

$$S^2_{ад.} = \frac{m}{N - B} \cdot \sum_{i=1}^{N} (y_p - \bar{y})^2 = \frac{2}{6-2} \cdot 6{,}99 = 3{,}50 .$$

Тоді розрахункове значення критерію Фішера буде рівне

$$F_p = 3{,}50 / 0{,}86 = 4{,}07.$$

Табличне значення критерію Фішера (додаток F), визначене при рівні значущості 0,05, числах мір свободи дисперсій адекватності $k_{ад} = 6 - 2 = 4$ і відтворюваності $k_{воспр} = 6 \cdot (2 - 1) = 6$, складає $F_m = 4{,}53$. Оскільки $F_p < F_m$, то рівняння регресії $y_p = 1{,}15 + 12{,}09x$ адекватно описує процес електрохімічного видобування нікелю з промстоків.

Приклад 1.19. Потрібно підібрати математичне вираження для опису результатів досліджень швидкостей осадження частинок піску фільтруючого матеріалу залежно від його крупності $v = f(d)$. При проведенні досліджень одержані наступні результати.

Таблиця 1.13 – Вихідні дані

№ п/п	**Математичне очікування швидкості осадження *v*, см/с**	**Крупність (розмір) часточок піску *d*, мм**
1	9,1	0,7
2	12,0	0,9
3	15,1	1,1
4	15,4	1,25
5	22,5	1,6
6	24,1	2,0

Для підбору математичного вираження рівняння залежності швидкості осадження від лінійних розмірів піску проведемо графічний аналіз вищезгаданої залежності. Нанесемо на координатну сітку точки з координатам и d_i и v_i.

Графічний аналіз (рис. 1.9) показує, що залежність $v=f(d)$ лінійна. Отже, порівняння має вигляд

$$v = a_0 + b \cdot d \qquad (y = a_0 + b \cdot x).$$

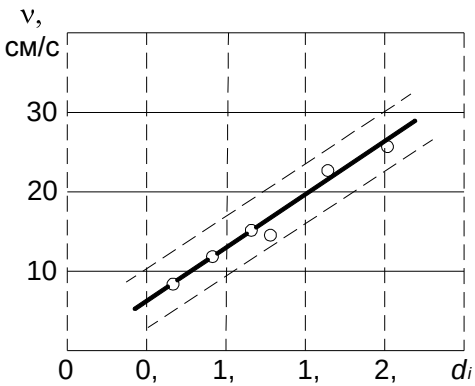

Рис. 1.9 – Залежність швидкості осадження від лінійних розмірів частинок піску

Визначення коефіцієнтів рівняння регресії a_0 і b проведемо за формулами (1.45) і (1.46).

Розрахунок за визначенням коефіцієнтів регресії проведемо в табл. 1.14.

$$b = \frac{6{,}0 \cdot 137{,}23 - 7{,}55 \cdot 98{,}2}{6 \cdot 10{,}63 - 7{,}55^2} = \frac{823{,}38 - 741{,}41}{63{,}78 - 57{,}00} = \frac{81{,}97}{6{,}78} = 12{,}09\ ;$$

$$a_0 = \frac{98{,}2 - 12{,}09 \cdot 7{,}55}{6} = \frac{6{,}92}{6{,}0} = 1{,}15\ .$$

$$y = 1{,}15 + 12{,}09 \cdot x \qquad \text{або} \qquad v = 1{,}15 + 12{,}09 \cdot d\ .$$

Таблиця 1.15 – Розрахунок коефіцієнтів регресії

№ п/п	y^2	x	x^2	xy	y_p	$(y_p - y_e)^2$
1	9,1	0,7	0,49	6,37	9,6	0,25
2	12,0	0,9	0,81	10,8	12,0	0,00
3	15,1	1,1	1,21	16,61	14,4	0,49
4	15,4	1,25	1,5625	19,25	16,3	0,81
5	22,5	1,6	2,56	36,00	20,5	4,00
6	24,1	2,0	4,00	48,20	25,3	1,44
	Σ=98,0	Σ = 7,55	Σ=10,6325	Σ=137,23		Σ = 6,99

Підставимо в одержане рівняння значення діаметрів і визначимо розрахункові значення швидкостей осадження:

v_1 = 1,15 + 12,09 – 0,7 = 9,6;
v_2 = 1,15 + 12,09 – 0,9 =12,0;
v_3 = 1,15 + 12,09 – 1,1 = 14,4;
v_4 = 1,15 + 12,09 – 1,25 = 16,3;
v_5= 1,15 + 12,09 – 1,6 = 20,5;
v_6= 1,15 + 12,09 – 2,0 = 25,3.

Середньоквадратичне відхилення результатів експерименту від значень рівняння регресії складає

$$S = \sqrt{\frac{\sum (y_p - y_e)^2}{n-1}} = \sqrt{\frac{6{,}99}{6-1}} = 1{,}18 \text{ см/с.}$$

Методи графічного зображення результатів експериментів

При обробці результатів вимірювань і спостережень широко використовують методи графічного зображення. Результати

вимірювань, представлені в табличній формі, не дозволяють достатньо наочно характеризувати закономірності процесів, що вивчаються. Графічне зображення дає найбільш наглядне уявлення про результати експериментів, дозволяє краще зрозуміти фізичну суть досліджуваного процесу, виявити загальний характер функціональної залежності змінних величин, що вивчаються, встановити наявність максимуму і мінімуму, функції.

Після обробки результатів вимірювань і оцінки ступеня точності необхідно їх звести в таблиці для аналізу. Дані таких таблиць обробляють графічними методами.

Для графічного зображення результатів вимірювань (спостережень), як правило, застосовують систему прямокутних координат. Якщо аналізується графічним методом функція $y = f(x)$, то наносять в системі прямокутних координат значення x_1y_1, x_2y_2, …, x_ny_n (рис. 1.10). Перш ніж будувати графік, необхідно знати хід (течія) досліджуваного явища. Як правило, якісні закономірності і форма графіка експериментатору орієнтовно відомі з теоретичних досліджень.

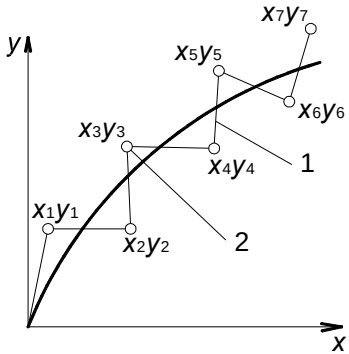

Рис. 1.10 – Графічне зображення функції $y = f(x)$; крива за результатами вимірювань; 2 – плавна

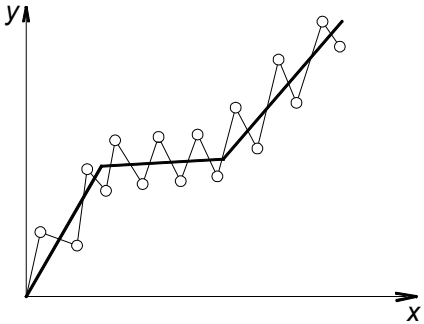

Рис. 1.11 – Графічне зображення функції $y = f(x)$ при наявності стрибка

Точки на графіку необхідно з’єднувати плавною лінією, так, щоб вона по можливості ближче проходила до всіх експериментальних точок. Якщо з'єднати точки прямими відрізками, то одержимо ламану криву. Вона характеризує зміну функції за даними експерименту. Звичайно функції мають плавний характер. Тому при графічному зображенні результатів вимірювань слід проводити між точками

плавні криві. Різке викривлення графіка пояснюється погрішностями вимірювань. Якби експеримент повторили із застосуванням засобів вимірювань вищої точності, то одержали б менші погрішності, а ламана крива більше б відповідала плавній кривій.

Проте можуть бути винятки. Так, іноді досліджуються явища, для яких в певних інтервалах спостерігається швидка стрибкоподібна зміна однієї з координат (рис. 1.11). Це пояснюється суттю фізико-хімічних процесів, наприклад фазовими перетвореннями вологи при дослідженні промерзаючих систем, радіоактивним розпадом атомів в процесі дослідження радіоактивності і т.д. У таких випадках необхідно особливо ретельно з'єднувати точки кривої. Загальне «усереднювання» всіх точок плавної кривої може привести до того, що стрибок функції підміняється погрішностями вимірювань.

Іноді при побудові графіка одна-дві точки різко віддаляються від кривої. Спочатку потрібно проаналізувати фізичну суть явища, і якщо немає підстави вважати наявність стрибка функції, то таке різке відхилення можна пояснити грубою помилкою або промахом. Це може виникнути тоді, коли дані вимірювань заздалегідь не досліджувалися на наявність грубих помилок вимірювань. У таких випадках необхідно повторити вимірювання в діапазоні різкого відхилення точки. Якщо колишнє вимірювання виявилося помилковим, то на графік наносять нову точку. Якщо ж повторні вимірювання дадуть колишнє значення, необхідно до цього інтервалу кривої віднестися дуже уважно і особливо ретельно проаналізувати фізичну суть явища. Часто при графічному зображенні результатів експериментів доводиться мати справу з трьома змінними: $b = f(x, y, z)$.

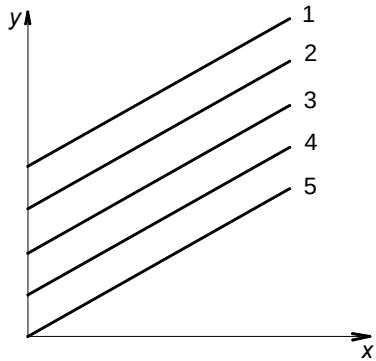

Рис. 1.12 – Графічне зображення функції $b = f(x, y, z)$:
1 – z_5; 2 – z_4; 3 – z_3; 4 – z_2; 4 – z_1.

В цьому випадку застосовують метод розділення змінних. Однієї з величин z в межах інтервалу вимірювань z_1 – z_n задають декілька послідовних значень. Для двох решти змінних x і y (при z_i = const) будують графіки $y = f_1(x)$. В результаті на одному графіку отримують сімейство кривих $y = f_1(x)$ для різних значень z (рис. 1.12).

Якщо необхідно графічно зобразити функцію з чотирма і більш змінним $a = f(b, x, y, z)$,

то будують серію графіків типу попередніх (рис. 1.12), але кожний з них при b_1, ..., b_n = const, або приймають з N змінних N – 1 постійними і будують графіки: спочатку $N - 1 = f_1(x)$, далі $N - 2 = f_2(x)$, $N - 3 = f_3(x)$ і т.д. Таким чином, можна прослідкувати зміну будь-якої змінної величини у функції від інших при постійних значеннях інших. Цей метод графічного аналізу вимагає ретельності, великої уваги до результатів вимірювань. Проте він в більшості випадків є найбільш простим і наглядним.

При графічному зображенні результатів експериментів велику роль грає вибір системи координат або координатної сітки. Координатні сітки бувають рівномірними і нерівномірними. У рівномірних координатних сіток ординати і абсциси мають рівномірну шкалу. Наприклад, в системі прямокутних координат довжина одиничних обрізків, що відкладаються, на обох осях однакова. З нерівномірних координатних сіток найбільш поширені напівлогарифмічні, логарифмічні, імовірнісні. Напівлогарифмічна сітка має рівномірну ординату і логарифмічну абсцису (рис. 1.13). Логарифмічна координатна сітка має обидві осі логарифмічні (див. рис. 1.13, імовірнісна – ординату, звичайно рівномірну, і абсцису – імовірнісну шкалу (рис. 1.14)).

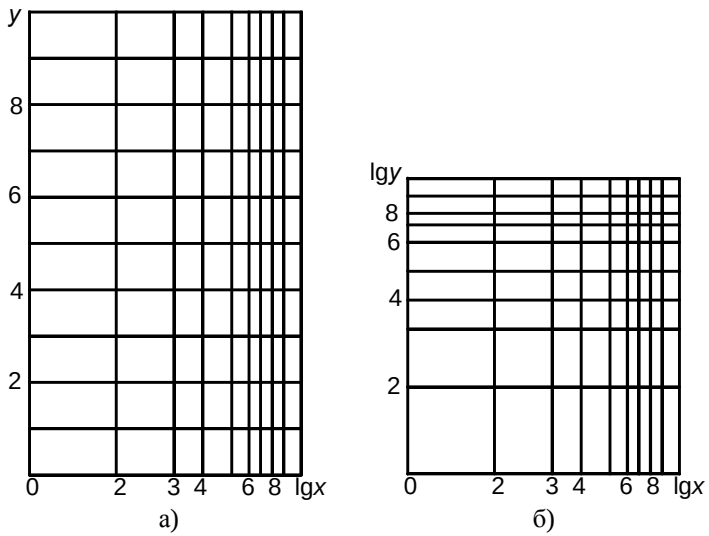

Рис. 1.13 – Координатна напівлогарифмічна (а) та логарифмічна (б) сітки

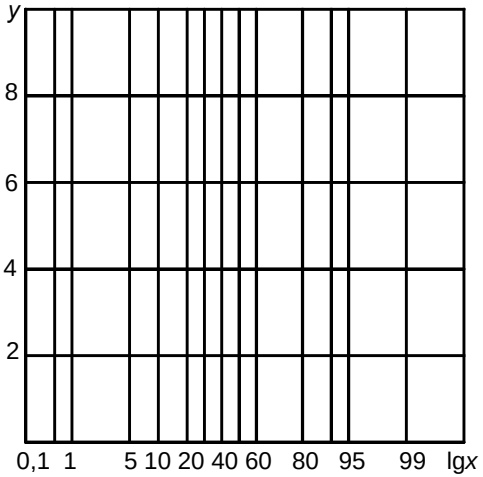

Рис. 1.14 – Координатна імовірнісна сітка

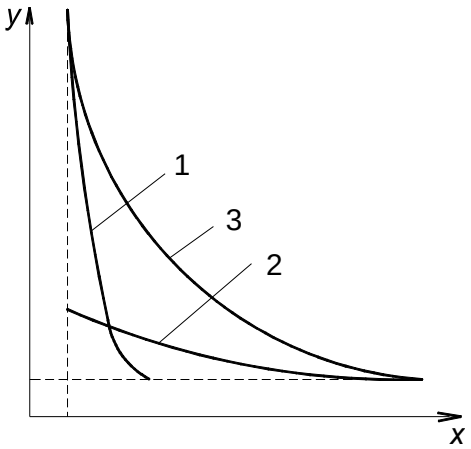

Рис. 1.14 – Форма графіка у залежності від масштабу:
1 – плоска; 2 – розширена; 3 – нормальна

Призначення нерівномірних сіток різне. В більшості випадків їх застосовують для наглядніщого зображення функцій. Функція $y = f(x)$ має різну форму на різних сітках.

Так, багато криволінійних функцій випрямляють на логарифмічних сітках. Велике значення в практиці графічного зображення експериментальних даних має імовірнісна сітка, вживана в

різних випадках: при обробці вимірювань для оцінки їх точності, при визначенні розрахункових характеристик (розрахункової вологості, розрахункових значень модуля пружності грунту і т.д.).

Іноді в процесі обробки експериментальних даних графічним способом необхідно скласти розрахункові графіки, прискорюючі знаходження по одній змінній інших. При цьому істотно підвищуються вимоги до точності викреслювання функції на графіку.

Викреслюючи розрахункові графіки, необхідно керуватися наступними практичними рекомендаціями. Залежно від числа змінних потрібно вибрати координатну сітку і визначити вид графіка - одна крива, сімейство кривих або серія сімейств. Великого значення набуває вибір масштабу графіка, що пов'язано з розмірами креслення і відповідно з точністю значень, що знімаються з нього. Відомо, що чим крупніший масштаб, тим вище точність значень, що знімаються. Проте, як правило, графіки не перевищують розмірів 20 х 15 см, що є зручним при складанні звітів. Лише в окремих випадках використовують графіки великих розмірів.

Досвід показує, що вживана для викреслювання графіків міліметрівка в межах розмірів 15-20 см дає погрішність +0,1-0,2 мм. Це слід мати на увазі при викреслюванні розрахункових графіків. Таким чином, абсолютна помилка значень, що знімаються з графіків, може досягати $\varepsilon = \pm 0{,}2\ M$, де M – прийнятий масштаб графіка. Очевидно, що точність вимірювань може бути вище за точність значень, що знімаються з графіка. Масштаб по координатних осях звичайно застосовують різний. Від вибору його залежить форма графіка - він може бути плоским (вузьким) або витягнутим (широким) уздовж осі (рис. 1.15). Вузькі графіки дають велику погрішність по осі y, широкі – по осі x. З рисунка видно, що правильно підібраний масштаб (нормальний графік) дозволяє суттєво підвищити точність відліків.

Розрахункові графіки, що мають максимум (мінімум) функції або який-небудь складний вигляд, особливо ретельно необхідно викреслювати в зонах вигину. На таких ділянках кількість точок для викреслювання графіка повинна бути значно більше, чим на плавних ділянках.

В деяких випадках будують номограми, що суттєво полегшують застосування для систематичних розрахунків складних теоретичних або емпіричних формул в певних межах вимірювання величин. Номограмміровани можуть бути будь-які вирази алгебри. В результаті складні математичні вирази можна вирішувати порівняно просто графічними методами. Побудова номограм – трудомістка

операція. Проте, будучи раз побудованою, номограма може бути використана для знаходження будь-якої із змінних, що входять в номограмоване рівняння. Застосування ЕОМ істотно знижує трудомісткість номограмування.

Існує декілька методів побудови номограм. Для цього застосовують рівномірні або нерівномірні координатні сітки. У системі прямокутних координат функції в більшості випадків на номограмах мають криволінійну форму. Це збільшує трудомісткість, оскільки потрібна велика кількість точок для нанесення однієї кривої.

У полу- або логарифмічних координатних сітках функції мають прямокутну форму і складання номограм спрощується.

Методика побудови номограм функцій однієї $y = f(x)$ або багатьох $y = f(x_1, x_2, ..., x_n)$ змінних описана раніше і зводиться до побудови кривої, сімейства або серії сімейств шляхом прийняття постійних і знаходження однієї змінної. Складні вирази алгебри доцільно зводити до простого добутку двох-трьох значень, наприклад: $d = abc$ де a, b, c – функції двох, трьох змінних. В цьому випадку необхідно спочатку, задавшись змінними, обчислити a, b, c. Далі, додаючи a, b, c постійні значення, знайти d: a, b, c необхідно варіювати в певних значеннях, наприклад від 0 до 100 через 5 або 10. Найбільш ефективним є такий спосіб побудови номограм, при якій a, b, c представляються як безрозмірні критерії.

Контрольні питання

1. Що таке середні значення?
2. Методи обчислення середніх.
3. Теоретичні середні (моменти розподілення).
4. Оцінки довірчих границь для істинного значення вимірюваної величини.
5. Порівняння дисперсій.
6. Порівняння середніх.
7. Перевірка гіпотези про рівність середніх.
8. Перевірка гіпотези нормальності закону розподілення випадкової величини.
9. Визначення теоретичного закону розподілення.
10. Що таке *кореляція*?
11. Що таке *регресія*?
12. Типи кореляції.
13. Суть кореляційного та регресійного аналізу.
14. Лінійна кореляція.
15. Оцінювання коефіцієнту кореляції.

16. Лінійний регресійний аналіз.
17. Оцінювання прямої регресії.
18. Що таке критерій Ст'юдента?
19. Що таке критерій Фішера?
20. Метод найменших квадратів.
21. Для чого потрібно графічне зображення результатів експерименту?
22. Вибір системи координат.

ДОДАТКИ

Додаток А

Інтеграл ймовірностей $\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_0^t \left[\exp(-t^2/2)\right] \cdot dt$, $\Phi(-t) = -\Phi(t)$

t	**Значення $\Phi(t)$ при сотих долях *t***									
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0,0	0,0000	0040	0080	0120	0160	0199	0239	0279	0319	0359
0,1	0398	0438	0478	0517	0557	0596	0636	1675	0714	0753
0,2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141
0,3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0,4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0,5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0,6	2257	2291	2324	2357	2389	2422	2454	2486	2517	2549
0,7	2580	2611	2642	2673	2703	2734	2764	2794	2823	2852
0,8	2881	2910	2939	2967	2995	3023	3051	3078	3106	3133
0,9	3159	3186	3212	3238	3264	3289	3315	3340	3365	3389
1,0	3413	3437	3461	3485	3508	3531	3554	3577	3599	3621
1,1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830
1,2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015
1,3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177
1,4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319
1,5	4332	4345	4357	4370	4382	4394	4406	4418	4429	4441
1,6	4452	4463	4474	4484	4495	4505	4515	4525	4535	4545
1,7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633
1,8	4641	4649	4656	4664	4671	4678	4686	4693	4699	4706
1,9	4713	4719	4726	4732	4738	4744	4750	4756	4761	4767
2,0	4772	4778	4783	4788	4793	4798	4803	4808	4812	4817
2,1	4821	4826	4830	4834	4838	4842	4846	4850	4854	4857
2,2	4861	4864	4868	4871	4875	4878	4881	4884	4887	4890
2,3	4893	4896	4898	4901	4904	4906	4909	4911	4913	4916
2,4	4918	4920	4922	4925	4927	4929	4931	4932	4934	4936

Таблиця допускає лінійну інтерполяцію з помилкою до 10.

Приклад. Розрахувати $\Phi(1{,}614)$.

Рішення: Беремо з таблиці значення $\Phi(1{,}61) = 0{,}1163$ та $\Phi(1{,}62) = 0{,}4474$ з різницею 0,0011 та вводимо поправку на відносне прирощення з аргументу (1,614-1,61)/0,01 = 0,4;

$$\Phi(1{,}614) = \Phi(1{,}61) + 0{,}0011 \cdot 0{,}4 = 0{,}4467.$$

Продовження таблиці для значень $t > 2{,}5$ см в Додатку В.

Додаток В

Величини, пов'язані з інтегралом ймовірностей $\Phi(t)$;

функція $t = t(P)$ є зворотньою для $P = 2\Phi(t)$

t	$\Phi(t)$	$1-2\Phi(t)$	$1-P$	$t=t(P)$	P
2,5	0,49379	0,01242	0,05	1,960	0,95
2,6	0,49534	0,00932	0,04	2,054	0,96
2,7	0,49653	0,00693	0,03	2,170	0,97
2,8	0,49744	0,00511	0,02	2,326	0,98
2,9	0,49813	0,00373	0,01	2,576	0,99
3,0	0,49865	0,00270	0,009	2,612	0,991
3,1	0,49903	0,00194	0,008	2,625	0,992
3,2	0,49931	0,00137	0,007	2,697	0,993
3,3	0,49952	0,00097	0,006	2,748	0,994
3,4	0,49966	0,00067	0,005	2,807	0,995
3,5	0,499767	0,000465	0,004	2,878	0,996
3,6	0,499841	0,000318	0,003	2,968	0,997
3,7	0,499892	0,000216	0,002	3,090	0,998
3,8	0,499927	0,000145	0,001	3,291	0,999
3,9	0,499952	0,000096	0,0009	3,320	0,9991
4,0	0,499968	0,000063	0,0008	3,353	0,9992
4,1	0,499979	0,000041	0,0007	3,390	0,9993
4,2	0,499987	0,000027	0,0006	3,432	0,9994
4,3	0,499991	0,000017	0,0005	3,481	0,9995
4,4	0,499995	0,000011	0,0004	3,540	0,9996
4,5	0,4999966	0,0000068	0,0003	3,615	0,9997
4,6	0,4999979	0,0000041	0,0002	3,720	0,9998
4,7	0,4999987	0,0000025	0,0001	3,891	0,9999
4,8	0,4999992	0,0000016	10^{-5}	4,417	$1 - 10^{-5}$
4,9	0,4999995	0,0000009	10^{-6}	4,892	$1 - 10^{-6}$
5,0	0,4999997	0,0000006	10^{-7}	5,327	$1 - 10^{-7}$

В таблиці значень $\Phi(t)$ помилка лінійної інтерполяції убуває зі збільшенням значень t ; вона не перебільшує:

10^{-4} – в інтервалі (2,5; 3,2); 10^{-5} – в інтервалі (3,2; 3,9);

10^{-6} – в інтервалі (3,9; 4,5); 10^{-7} – в інтервалі (4,5; 5,0).

В таблиці значень $t(P)$ інтерполяцію не роблять.

Додаток С

Критичні значення $t(P)$ відношення (1.1.3)
для вибраковування «вискакуючих» значень *x*
(*n* – число прийнятних результатів, *P* – надійність виводу)

n	Значення $t(P)$ при надійності				*n*	Значення $t(P)$ при надійності			
	0,95	**0,98**	**0,99**	**0,999**		**0,95**	**0,98**	**0,99**	**0,999**
5	3,04	4,11	5,04	9,43	20	2,145	2,602	2,932	3,979
6	2,78	3,64	4,36	7,41	25	2,105	2,541	2,952	3,819
7	2,62	3,36	3,96	6,37	30	2,079	2,503	2,802	3,719
8	2,51	3,18	3,71	5,73	35	2,061	2,476	2,768	3,652
9	2,43	3,05	3,54	5,31	40	2,048	2,456	2,742	3,602
10	2,37	2,96	3,41	5,01	45	2,038	2,441	2,722	3,565
11	2,33	2,89	3,31	4,79	50	2,030	2,429	2,707	3,532
12	2,29	2,83	3,23	4,62	60	2,018	2,411	2,683	3,492
13	2,26	2,78	3,17	4,48	70	2,009	2,399	2,667	3,462
14	2,24	2,74	3,12	4,37	80	2,003	2,389	2,655	3,439
15	2,22	2,71	3,08	4,28	90	1,998	2,382	2,646	3,423
16	2,20	2,68	3,04	4,20	100	1,994	2,377	2,639	3,409
17	2,18	2,66	3,01	4,13	∞	1,960	2,326	2,576	3,291
18	2,17	2,64	2,98	4,07					

Лінійна інтерполяція за аргументом *n* може дати помилку до 10^{-2} при $20 < n < 60$ та помилку до 10^{-3} при $60 < n < 100$.

При $n > 100$ критичні значення $t_n(P)$ з точністю до 10^{-3} можна розрахувати за формулою

$$t_n(P) = t_\infty(P) + \frac{t_{100}(P) - t_\infty(P)}{n} \cdot 100;$$

наприклад, при $P = 0{,}99$ та $n = 200$ маємо:

$$t_{200}(0{,}99) = 2{,}576 + \frac{2{,}639 - 2{,}576}{200} \cdot 100 = 2{,}576 + 0{,}031 = 2{,}607.$$

Додаток D

Критичні значення T_N Ф.Граббса

Кількість спостережень *n*	**Значення T_N при рівні значущості *q***			**Кількість спостережень *n***	**Значення T_N при рівні значущості *q***		
	0,1	**0,05**	**0,025**		**0,1**	**0,05**	**0,025**
3	1,406	1,412	1,414	14	2,297	2,461	2,602
4	1,645	1,689	1,710	15	2,326	2,493	2,638
5	1,791	1,869	1,917	16	2,354	2,523	2,670
6	1,894	1,996	2,067	17	2,380	2,551	2,701
7	1,974	2,093	2,182	18	2,404	2,577	2,728
8	2,041	2,172	2,273	19	2,426	2,600	2,754
9	2,097	2,237	2,349	20	2,447	2,623	2,778
10	2,146	2,294	2,414	21	2,467	2,644	2,801
11	2,190	2,343	2,470	22	2,486	2,664	2,823
12	2,229	2,387	2,519	23	2,504	2,683	2,843
13	2,264	2,426	2,562	24	2,520	2,701	2,862
				25	2,537	2,717	2,880

Додаток D, а

Відсоткові точки критерію Смірнова-Граббса $\left[i_{\max\left|x_{(i)}-x\right|} \right] / S$

n	Довірча імовірність *p*				*n*	Довірча імовірність *p*			
	0,8	**0,9**	**0,95**	**0,99**		**0,8**	**0,9**	**0,95**	**0,99**
3	1,406	1,412	1,414	1,414	27	2,568	2,749	2,913	3,239
4	1,645	1,689	1,710	1,728	28	2,582	2,764	2,929	3,258
5	1,791	1,869	1,917	1,972	29	2,596	2,778	2,944	3,275
6	1,894	1,996	2,067	2,161	30	2,609	2,792	2,958	3,291
7	1,974	2,093	2,182	2,310	31	2,622	2,805	2,972	3,307
8	2,041	2,172	2,273	2,431	32	2,634	2,818	2,985	3,322
9	2,097	2,237	2,349	2,532	33	2,646	2,830	2,998	3,337
10	2,146	2,294	2,414	2,616	34	2,657	2,842	3,010	3,351
11	2,190	2,343	2,470	2,689	35	2,668	2,853	3,022	3,364
12	2,229	2,387	2,519	2,753	36	2,678	2,864	3,033	3,377
13	2,264	2,426	2,563	2,809	37	2,689	2,874	3,044	3,389
14	2,297	2,461	2,602	2,859	38	2,699	2,885	3,055	3,401
15	2,327	2,494	2,638	2,905	39	2,709	2,894	3,065	3,413
16	2,354	2,523	2,670	2,946	40	2,718	2,904	3,075	3,424
17	2,386	2,551	2,701	2,983	41	2,727	2,913	3,084	3,435
18	2,404	2,577	2,728	3,017	42	2,736	2,922	3,094	3,445
19	2,426	2,601	2,754	3,049	43	2,745	2,931	3,103	3,455
20	2,447	2,623	2,779	3,079	44	2,753	2,940	3,112	3,465
21	2,467	2,644	2,801	3,106	45	2,762	2,948	3,120	3,474
22	2,486	2,664	2,823	3,132	46	2,770	2,956	3,129	3,483
23	2,504	2,683	2,843	3,156	47	2,778	2,964	3,137	3,492
24	2,521	2,701	2,862	3,179	48	2,785	2,972	3,145	3,501
25	2,537	2,718	2,880	3,200	49	2,793	2,980	3,152	3,510
26	2,553	2,734	2,897	3,220	50	2,800	2,987	3,160	3,518

Додаток Е

Критичні точки розподілу Ст'юдента

Число ступінів свободи *k*	**Рівень значущості *q* (двостороння критична область)**					
	0,10	**0,05**	**0,02**	**0,01**	**0,002**	**0,001**
1	6,31	12,7	31,82	63,7	318,3	637,0
2	2,92	4,30	6,97	9,92	22,33	31,6
3	2,35	3,18	4,54	5,84	10,22	12,9
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,01	2,57	3,37	4,03	5,89	6,86
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,40
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,03	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,96
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,73	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,51	3,79
23	1,71	2,07	2,50	2,81	3,49	3,77
24	1,71	2,06	2,49	2,80	3,47	3,74
25	1,71	2,06	2,49	2,79	3,45	3,72
26	1,71	2,06	2,48	2,78	3,44	3,71
27	1,71	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,46	2,76	3,40	3,66
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46
120	1,66	1,98	2,36	2,62	3,17	3,36
∞	1,64	1,96	2,33	2,58	3,09	3,29
	0,05	0,025	0,01	0,005	0,001	0,0005
	Рівень значущості *q* (одностороння критична область)					

Додаток F

Критичні точки *F*-розподілення Фішера-Снедекора при *q* = 0,05

(k_1 – число ступінів свободи більшої дисперсії,
k_2 – число ступінів свободи меншої дисперсії)

k_2	Значення $F_{табл.}$ при числі ступіней свободи k_1									
	1	**2**	**3**	**4**	**5**	**6**	**8**	**12**	**24**	∞
1	161,40	199,50	215,70	224,60	230,20	234,00	238,90	243,90	249,00	254,30
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,60	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
40	4,08	3,23	2,84	2,64	2,45	2,34	2,18	2,00	1,79	1,52
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,02	1,83	1,61	1,25
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1,00

Додаток G

Критичні точки розподілення χ^2

Число ступінів свободи *k*	**Значення χ^2 при рівні значущості *q***					
	0,01	**0,025**	**0,05**	**0,95**	**0,975**	**0,99**
1	6,6	5,0	3,8	0,0039	0,00098	0,00016
2	9,2	7,4	6,0	0,103	0,051	0,020
3	11,3	9,4	7,8	0,352	0,216	0,115
4	13,3	11,1	9,5	0,711	0,484	0,297
5	15,1	12,8	11,1	1,15	0,831	0,554
6	16,8	14,4	12,6	1,64	1,24	0,872
7	18,5	16,0	14,1	2,17	1,69	1,24
8	20,1	17,5	15,5	2,73	2,18	1,65
9	21,7	19,0	16,9	3,33	2,70	2,09
10	23,2	20,5	18,3	3,94	3,25	2,56
11	24,7	21,9	19,7	4,57	3,82	3,05
12	26,2	23,3	21,0	5,23	4,40	3,57
13	27,7	24,7	22,4	5,89	5,01	4,11
14	29,1	26,1	23,7	6,57	5,63	4,66
15	30,6	27,5	25,0	7,26	6,26	5,23
16	32,0	28,8	26,3	7,96	6,91	5,81
17	33,4	30,2	27,6	8,67	7,56	6,41
18	34,8	31,5	28,9	9,39	8,23	7,01
19	36,2	32,9	30,1	10,1	8,91	7,63
20	37,6	34,2	31,4	10,9	9,59	8,26
21	38,9	35,5	32,7	11,6	10,3	8,90
22	40,3	36,8	33,9	12,3	11,0	9,54
23	41,6	38,1	35,2	13,1	11,7	10,2
24	43,0	39,4	36,4	13,8	12,4	10,9
25	44,3	40,6	37,7	14,6	13,1	11,5
26	45,6	41,0	38,9	15,4	13,8	12,2
27	47,0	43,2	40,1	16,2	14,6	12,9
28	48,3	44,5	41,3	16,9	15,3	13,6
29	49,6	45,7	42,6	17,7	16,0	14,3
30	50,9	47,0	43,8	18,5	16,8	15,0

Додаток Н

Критичні точки $G_т$-розподілення Кохрена

(k – число ступінів свободи, l – кількість виборок)

l	Значення $G_т$ при q = 0,01 та числі ступінів свободи k						
	1	**2**	**3**	**4**	**5**	**6**	**7**
2	0,9999	0,9950	0,9794	0,9586	0,9373	0,9172	0,8988
3	9993	9423	8831	8335	7933	7606	7335
4	9676	8643	7814	7212	6761	6410	6129
5	0,9279	0,7885	0,6957	0,6329	0,5875	0,5531	0,5259
6	8828	7218	6258	5635	5195	4866	4608
7	8376	6644	5685	5080	4659	4347	4105
8	0,7945	0,6152	0,5209	0,4627	0,4226	0,3932	0,3704
9	7544	5727	4810	4251	3870	3592	3378
10	7175	5358	4469	3934	3572	3308	3106
12	0,6528	0,4751	0,3919	0,3428	0,3099	0,2861	0,2680
15	5747	4069	3317	2882	2593	2386	2228
20	4799	3297	2654	2288	2048	1877	1748
24	0,4247	0,2871	0,2295	0,1970	0,1759	0,1608	0,1495
30	3632	2412	1913	1635	1454	1327	1232
40	2940	1915	1508	1281	1135	1033	0957
60	0,2151	0,1371	0,1069	0,0902	0,0796	0,0722	0,0668
120	1225	0759	0585	0489	0429	0387	0357
∞	0000	0000	0000	0000	0000	0000	0000

l	Значення $G_т$ при q = 0,01 та числі ступінів свободи k						
	8	**9**	**10**	**16**	**36**	**144**	∞
2	0,8823	0,8674	0,8539	0,7949	0,7067	0,6062	0,5000
3	7107	6912	6743	6059	5153	4230	3333
4	5897	5702	5536	4884	4057	3251	2500
5	0,5037	0,4854	0,4697	0,4094	0,3351	0,2644	0,2000
6	4401	4229	4084	3529	2858	2229	1667
7	3911	3751	3616	3105	2494	1929	1429
8	0,3522	0,3373	0,3248	0,2779	0,2214	0,1700	0,1250
9	3207	3067	2950	2514	1992	1521	1111
10	2945	2913	2704	2297	1811	1376	1000
12	0,2535	0,2419	0,2320	0,1961	0,1535	0,1157	0,0833
15	2104	2002	1918	1612	1251	0934	0667
20	1646	1567	1501	1248	0960	0709	0500
24	0,1406	0,1338	0,1283	0,1060	0,0810	0,0595	0,0417
30	1157	1100	1054	0867	0658	0480	0333
40	0898	0853	0816	0668	0503	0363	0250
60	0,0625	0,0594	0,0567	0,0461	0,0344	0,0245	0,0167
120	0334	0316	0302	0242	0178	0125	0083
∞	0000	0000	0000	0000	0000	0000	0000

l	Значення $G_т$ при q = 0,05 та числі ступінів свободи *k*						
	1	2	3	4	5	6	7
2	0,9985	0,9750	0,9392	0,9057	0,8772	0,8534	0,8332
3	9669	8709	7977	7457	7071	6771	6530
4	9065	7679	6841	6287	5895	5598	5365
5	0,8412	0,6338	0,5981	0,5440	0,5063	0,4783	0,4534
6	7808	6161	5321	4803	4447	4184	3980
7	7271	5612	4800	4307	3974	3726	3535
8	0,6798	0,5157	0,4377	0,3910	0,3595	0,3362	0,3185
9	6385	4775	4027	3584	3286	3067	2901
10	6020	4450	3733	3311	3029	2823	2666
12	0,5410	0,3924	0,3624	0,2880	0,2624	0,2439	0,2299
15	4709	3346	2758	2419	2195	2034	1911
20	3894	2705	2205	1921	1735	1602	1501
24	0,3434	0,2354	0,1907	0,1656	0,1493	0,1374	0,1286
30	2929	1980	1593	1377	1237	1137	1061
40	2370	1576	1259	1082	0968	0887	0827
60	0,1737	0,1131	0,0895	0,0765	0,0682	0,0623	0,0583
120	0998	0632	0495	0419	0371	0337	0312
∞	0000	0000	0000	0000	0000	0000	0000

l	**Значення $G_т$ при q = 0,05 та числі ступінів свободи *k***						
	8	**9**	**10**	**16**	**36**	**144**	∞
2	0,8159	0,8010	0,7880	0,7341	0,6602	0,5813	0,5000
3	6333	6167	6025	5466	4748	4031	3333
4	5175	5017	4884	4366	3720	3093	2500
5	0,4387	0,4241	0,4118	0,3645	0,3066	0,2013	0,2000
6	3817	3682	3568	3135	2612	2119	1667
7	3384	3259	3154	2756	2278	1833	1429
8	0,3043	0,2926	0,2829	0,2462	0,2022	0,1616	0,1250
9	2768	2659	2568	2226	1820	1446	1111
10	2541	2439	2353	2032	1655	1308	1000
12	0,2187	0,2098	0,2020	0,1737	0,1403	0,1100	0,0833
15	1815	1736	1671	1429	1144	0889	0667
20	1422	1357	1303	1108	0879	0675	0500
24	0,1216	0,1160	0,1113	0,0942	0,0743	0,0567	0,0417
30	1002	0958	0921	0771	0604	0457	0333
40	0780	0745	0713	0595	0462	0347	0250
60	0,0552	0,0520	0,0497	0,0411	0,0316	0,0234	0,0167
120	0292	0279	0266	0218	0165	0120	0083
∞	0000	0000	0000	0000	0000	0000	0000

Навчальне видання

Методичні вказівки до науково-дослідницької практики студентів з дисципліни «Організація наукових досліджень» (Статистичні методи. Аналіз та оформлення наукових досліджень)
(для студентів 5 курсу всіх форм навчання
8.092108 – «Теплогазопостачання і вентиляція»)

Укладачі: Іван Іванович Капцов,
Олександр Васильович Ромашко,
Людмила Вікторівна Гапонова,
Вікторія Вікторівна Гранкіна

Редактор: М.З. Аляб'єв

План 2009, поз. 725М

Підп. до друку 21.09.09	Формат 60х84 1/16	Папір офісний
Друк на ризографі	Умовн.-друк. арк. 3,1	Обл.-вид. арк. 3,4
Замовл. №	Тираж 50 прим.	

61002, м. Харків, ХНАМГ, вул. Революції, 12
Сектор оперативної поліграфії ЦНІТ ХНАМГ
61002, м. Харків, ХНАМГ, вул. Революції, 12