Методичні вказівки до проведення практичних занять та виконання контрольних робіт з дисципліни «Спецглави математики». Ч. 2 для студентів напряму підготовки «Метрологія та інформаційно-вимірювальні технології» всіх форм навчання

Міністерство освіти і науки України
Вінницький національний технічний університет

Методичні вказівки
до проведення практичних занять
та виконання контрольних робіт
з дисципліни «Спецглави математики». Ч. 2
для студентів напряму підготовки
«Метрологія та інформаційно-вимірювальні технології» всіх форм навчання

Вінниця
ВНТУ
2016

Рекомендовано до друку Методичною радою Вінницького національного технічного університету Міністерства освіти і науки України (протокол № 10 від 22.05.2014 р.)

Рецензенти:
С. М. Довгалець, кандидат технічних наук, доцент
В. П. Штовба, доктор технічних наук, професор

Методичні вказівки до проведення практичних занять та виконання контрольних робіт з дисципліни «Спецглави математики». Ч. 2 для студентів напряму підготовки «Метрологія та інформаційно-вимірювальні технології» всіх форм навчання / Уклад. П. І. Кулаков, В. В. Присяжнюк − Вінниця : ВНТУ, 2016. – 49 с.

У другій частині методичних вказівок подаються короткі теоретичні відомості та практичні завдання з основ регресійного аналізу, дисперсійного аналізу та основи теорії випадкових процесів, наведені програма дисципліни, варіанти завдань, приклади розв'язування типових задач та рекомендації до виконання контрольної роботи. Для полегшення самостійного виконання завдань наводиться список рекомендованої літератури.

Призначений для студентів всіх форм навчання нематематичних спеціальностей, які вивчають дисципліни «Спецглави математики», «Теорія ймовірностей та математична статистика» та інші.

ЗМІСТ

Вступ………………………………………..………..…......…… 4
1. Програма дисципліни…...………………………………….......... 5
2. Короткі теоретичні відомості…..………………………............... 7
2.1 Основи регресійного аналізу...…………………………............ 7
2.2 Метод найменших квадратів ………………….…….………... 7
2.3 Парна регресія…....……………………………….................. 10
2.4 Множинна лінійна регресія…………………..………….............. 11
2.5. Нелінійна регресія. Нелінійна регресія за параметрами............... 13
2.6 Основи дисперсійного аналізу...……………………………..... 15
2.6.1 Однофакторний дисперсійний аналіз.....……........................ 16
2.6.2 Двофакторний дисперсійний аналіз..……………….......... 21
3. Завдання для виконання контрольної роботи…………………….. 25
3.1 Основи регресійного аналізу...…………………...……… 25
3.2 Елементи дисперсійного аналізу………………….............. 34
3.3 Перевірка статистичної гіпотези…………………............ 43
Література……………………………………………...……... 48

ВСТУП

Методичні вказівки до практичних занять та контрольних робіт з дисципліни «Спецглави математики» призначені для студентів всіх форм навчання, які навчаються за напрямом 6.051001 – «Метрологія та інформаційно-вимірювальні технології».

Дана методично-інструктивна література складається з: програми дисципліни, коротких теоретичних відомостей, після кожного розділу наведені контрольні питання та задачі для перевірки самопідготовки студента, а також завдання для виконання контрольних робіт. Всі завдання для виконання контрольних робіт складаються з 25 варіантів і наведені приклади розв'язування всіх завдань.

Методичні вказівки можна використовувати студентам денної форми навчання для виконання практичних занять, для підготовки до колоквіумів, для виконання типових розрахунків, контрольних домашніх робіт, а також для виконання контрольних робіт студентами заочної форми навчання.

Контрольні роботи виконуються з метою закріплення теоретичних і практичних знань студента.

Студенти денної форми навчання виконують контрольні роботи з таких тем, а саме: закони розподілу випадкових величин; закон розподілу функції одного випадкового аргументу; наближені способи визначення моментів функції випадкового аргументу; основи регресійного аналізу; елементи дисперсійного аналізу; перевірка статистичної гіпотези.

Студенти заочної форми навчання вивчають дисципліну «Спецглави математики» протягом двох семестрів і виконують в кожному семестрі по одній контрольній роботі. В першому семестрі з тем: закони розподілу випадкових величин; закон розподілу функції одного випадкового аргументу; наближені способи визначення моментів функції випадкового аргументу; і складають диференціальний залік, в другому семестрі з тем: основи регресійного аналізу; елементи дисперсійного аналізу; перевірка статистичної гіпотези і завершують вивчення дисципліни іспитом в другому семестрі. Студент заочної форми навчання повинен виконати контрольну роботу до початку екзаменаційної сесії.

Перед виконанням контрольної роботи студенту рекомендується опрацювати перелік теоретичних питань, наведених для кожного завдання. У процесі виконання студент може користуватися не тільки рекомендованою, але і будь-якою іншою доступною йому навчальною та технічною літературою, а також сайтом дистанційної форми навчання, який розташовано за посиланням http://elearn.lan (повна функціональність, доступ тільки з внутрішньої мережі ВНТУ) або http://elearn.vntu.edu.ua (обмежена функціональність, відкритий доступ з Інтернет).

Вибір варіанта завдання здійснюється за порядковим номером студента у журналі обліку академічної групи.

1 ПРОГРАМА ДИСЦИПЛІНИ

Вступ. Основні означення теорії ймовірності. Предмет та методи теорії ймовірності. Випадкова величина. Характеристика випадкової величини. Закон розподілу, функція та щільність розподілу ймовірностей, визначення та властивості. Характеристична функція випадкової величини. Числові характеристики випадкових величин та їх властивості. Початкові та центральні моменти, математичне сподівання, дисперсії, мода, медіана, асиметрія, ексцес.

Основні закони розподілу випадкової величини. Закони розподілу дискретної та неперервної випадкових величин. Розподіл Пуассона. Рівномірний розподіл. Показовий розподіл. Нормальний розподіл. Розподіл Коші.

Двовимірні розподіли. Означення та властивості двовимірних законів. Система двох випадкових величин. Умовні розподіли двох випадкових величин. Числові характеристики двовимірних законів.

Закони розподілу систем багатьох випадкових величин. Багатовимірні випадкові величини. Багатовимірне нормальне розподілення. Закони розподілення підсистеми неперервних випадкових величин та умовні закони розподілу.

Закони розподілу функцій випадкових величин. Закон розподілу функціонально перетворених випадкових величин, функції одного випадкового аргументу. Числові характеристики функціонально перетворених випадкових величин функції одного випадкового аргументу. Закон розподілу функції двох випадкових аргументів. Виведення формул закону розподілу функції основних математичних залежностей двох випадкових величин.

Функціональні перетворення двовимірних законів. Перетворення Якобіана. Закон розподілу функції багатьох випадкових аргументів. Композиція двовимірних і тривимірних нормальних законів розподілу з використанням поняття векторних відхилень. Характеристичні функції систем і функцій випадкових величин.

Лінеаризація функцій випадкових величин. Лінеаризація функцій одного випадкового аргументу. Лінеаризація функцій двох та багатьох випадкових аргументів. Наближене визначення математичного сподівання та дисперсії.

Основи кореляційного, регресійного та дисперсійного аналізу. Елементи кореляційного аналізу. Лінійна кореляція. Коефіцієнт кореляції. Види регресійних моделей. Лінійна парна регресія, множинна лінійна регресія. Нелінійна регресія, множинна нелінійна регресія. Нелінійна за параметрами. Коефіцієнт множинної регресії. Метод найменших квадратів. Метод максимуму правдоподібності. Загальна теорія дисперсійного аналізу. Одно-та багатофакторний дисперсійний аналіз. Критерій Фішера.

Предмет та задачі математичної статистики. Предмет та методи математичної статистики. Статистичний матеріал. Статистичний розподіл. Статистичні розподіли вибірок та їх числові характеристики. Варіаційний ряд. Знаходження моментів випадкової величини за результатами дослідів. Статистичні оцінки. Точкові оцінки. Інтервальні оцінки. Надійний інтервал і надійна ймовірність. Статистичні гіпотези. Параметричні і непараметричні критерії.

Основи теорії випадкових процесів. Характеристика та види випадкових процесів. Випадкові процеси, часові та кореляційні характеристики. Властивості кореляційної, нормованої та взаємної кореляційної функцій. Класифікація випадкових процесів. Стаціонарні випадкові процеси в широкому і вузькому розумінні. Властивості кореляційної та нормованої функцій стаціонарних випадкових процесів. Визначення інтервалу кореляції та максимального інтервалу кореляції. Ергодичні стаціонарні випадкові процеси. Зв’язок між кореляційними і спектральними характеристиками. Перетворення Хінчіна-Вінера. Властивості спектральної характеристики випадкового процесу. Загальна теорія перетворення випадкових процесів. Загальна характеристики проходження стаціонарних випадкових сигналів через лінійні та нелінійні кола. Визначення кореляційної функції вихідного сигналу при вхідному стаціонарному випадковому процесі.

2 КОРОТКІ ТЕОРЕТИЧНІ ВІДОМОСТІ

2.1 Основи регресійного аналізу

Регресійний аналіз знайшов широке застосування для побудови математичних моделей, які описують різноманітні явища, процеси, досліди, залежність функції від одного або декількох аргументів.

Види регресійної моделі:

- парна регресія – $y_i = a_0 + a_1 x_i$;
- поліноміальна регресія – $y_i = a_0 + a_1 x_i + a_2 x_i^2 + ... + a_n x_i^n$;
- нелінійна регресія – $y = a_0 + \frac{a_1}{x}$, $y = a_0 + a_1 \ln(x)$;
- нелінійна за параметрами – $y = \frac{x}{a_0 + a_1 x}$, $y = a_0 \ell^{a_1 x}$;
- множинна лінійна регресія – $y_i = a_0 + a_1 x_1 + a_2 x_2 + ... + a_n x_n$.

2.2 Метод найменших квадратів

Найбільш розповсюдженим методом для визначення оцінок параметрів регресійних моделей є метод найменших квадратів (МНК). В даному методі оцінки параметрів регресійної моделі визначають з умови, що сума квадратів відхилень експериментальних значень y_i від розрахункових (теоретичних) значень $\hat{y}_i = \varphi(x_i, a_0, ..., a_n)$ мінімальна, тобто:

$$\sum_{i=1}^{m} \left[y_i - \varphi(x_i, a_0, ..., a_n) \right]^2 = \sum_{i=1}^{m} \delta_i^2 = Q \to \min, \qquad (2.1)$$

де δ_i – відхилення.

При розгляді МНК обмежимося випадком, коли залежність задана поліномом, тобто:

$$\hat{y}_i = a_0 + a_1 x_i + a_2 x_i^2 + ... + a_n x_i^n. \qquad (2.2)$$

Відповідно до (2.1), найкращими значеннями коефіцієнтів a_0, a_1, a_n будуть ті, для яких сума квадратів відхилення:

$$Q = \sum_{i=1}^{m} \delta_k^2 = \sum_{i=1}^{m} \left(y_i - \left(a_0 + a_1 x_i + ... + a_n x_i^n \right) \right)^2, \qquad (2.3)$$

буде мінімальна. Мінімум функції багатьох змінних, як відомо, досягається тоді, коли всі її частинні похідні дорівнюють нулю. Тому, диференціюючи (2.3) за змінною a_n, одержуємо:

$$\begin{cases} \dfrac{\partial Q}{\partial a_0} = -2\sum_{i=1}^{m}\left(y_i - \left(a_0 + a_1x_i + a_2x_i^2 + \ldots + a_nx_i^n\right)\right) = 0; \\ \dfrac{\partial Q}{\partial a_1} = -2\sum_{i=1}^{m}x_i\left(y_i - \left(a_0 + a_1x_i + a_2x_i^2 + \ldots + a_nx_i^n\right)\right) = 0; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \dfrac{\partial Q}{\partial a_n} = -2\sum_{i=1}^{m}x_i^n\left(y_i - \left(a_0 + a_1x_i + a_2x_i^2 + \ldots + a_nx_i^n\right)\right) = 0. \end{cases} \tag{2.4}$$

Після спрощення одержимо:

$$\begin{cases} ma_0 + a_1\sum x_i + a_2\sum x_i^2 + \ldots + a_n\sum x_i^n = \sum y_i; \\ a_0\sum x_i + a_1\sum x_i^2 + a_2\sum x_i^3 \ldots + a_n\sum x^{n+1} = \sum y_1x_i; \\ a_0\sum x_i^{\ 2} + a_1\sum x_i^3 + a_2\sum x_i^4 \ldots + a_n\sum x^{n+2} = \sum y_ix_i^{\ 2}; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0\sum x_i^{\ n} + a_1\sum x_i^{n+1} + a_2\sum x_i^{n+2} \ldots + a_n\sum x^{2n} = \sum y_ix_i^{\ n}. \end{cases} \tag{2.5}$$

Розв’язок системи лінійних рівнянь знайдемо за допомогою визначників:

$$a_0 = \frac{D_0}{D}; a_1 = \frac{D_1}{D}; \ldots; a_n = \frac{D_n}{D}, \tag{2.6}$$

де D – головний визначник системи (2.5),

D_i – визначник, який отримується із головного визначника шляхом заміни стовпців коефіцієнтів при невідомих a_i на стовпці з коефіцієнтами з вільними членами.

Виправлене середнє квадратичне коефіцієнтів a_i виражається такою формулою:

$$S(a_i) = S(\delta)\sqrt{\frac{D_{(i+1)(i+1)}}{D}}, \tag{2.7}$$

де $D_{(i+1)(i+1)}$ – алгебраїчне доповнення елементів головного визначника D, отримане шляхом видалення з матриці визначника стовпця $(i+1)$ і рядка $(i+1)$.

Стандартна похибка $S(\delta)$:

$$S(\delta) = \sqrt{\frac{\sum_{i=1}^{m}\delta_i^2}{m-n-1}}, \tag{2.8}$$

причому δ_i – обчислюються на підставі системи (2.1).

Надійний інтервал існування a_n обчислюється за формулою:

$$\Delta(a_i) = t_{(P,k)} S(a_i), \tag{2.9}$$

де $t_{(P,k)}$ – коефіцієнт Стьюдента, знаходиться за спеціальною таблицею розподілу Стьюдента, при кількості ступенів вільності k=m-n-1 і при заданій надійній ймовірності p.

Виправлена дисперсія функції Y, визначається за формулою:

$$S_y^2 = \sum_{i=0}^{n} \left(\frac{\partial \varphi}{\partial a_i} \right)^2 S_{x_i}^2 + 2\sum_{i=0}^{n} \sum_{j=1}^{k} \left(\frac{\partial \varphi}{\partial a_i} \right) \left(\frac{\partial \varphi}{\partial a_j} \right) k_{ij}, \tag{2.10}$$

де k_{ij}, або $k(a_i, a_j)$ – кореляційний момент, $k(a_i, a_j) = S^2(\delta) \dfrac{D_{(i+1)(j+1)}}{D}$,

$D_{(i+1)(j+1)}$ – алгебраїчне доповнення елементів головного визначника D, одержуване шляхом видалення з матриці визначника стовпця (i+1) і рядка (j+1) із множенням отриманого визначника на $(-1)^{i+j+2}$.

Наприклад, для полінома другого степеня $\hat{y}_i = a_0 + a_1 x_i + a_2 x_i^2$, рівняння (2.10) має вигляд:

$$\begin{aligned} S^2(y) &= S^2(a_0) + x^2 S^2(a_1) + x^4 S^2(a_2) + 2xk(a_0 a_1) + \\ &+ 2x^2 k(a_0 a_2) + 2x^3 k(a_1 a_2). \end{aligned} \tag{2.11}$$

Параметри поліноміальної регресії, коли степінь полінома більше трьох, бажана визначати за допомогою матричного аналізу.

Наприклад: якщо визначаються параметри поліноміальної регресії $y_i = a_0 + a_1 x + a_2 x^2 + ... + a_n x^n$, то коефіцієнти a_i зручніше визначати за формулою:

$$A = \left(X^T X \right)^{-1} X^T Y, \tag{2.12}$$

де – $A = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}$, $X = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots x_1^n \\ 1 & x_2 & x_2^2 & \cdots x_2^n \\ \cdots & \cdots & \cdots & \cdots\cdots \\ 1 & x_m & x_m^2 & \cdots x_m^n \end{pmatrix}$, $Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \cdots \\ y_n \end{pmatrix}$.

Визначити дисперсії параметрів рівняння регресії і кореляційних моментів можна за допомогою кореляційної матриці:

$$K_{ij} = (X^T X)^{-1} S^2(\delta) = \begin{vmatrix} S^2(a_0) & k(a_0, a_1) & k(a_0, a_2) \\ k(a_1, a_0) & S^2(a_1) & k(a_1, a_2) \\ k(a_2, a_0) & k(a_2, a_1) & S^2(a_2) \end{vmatrix}. \qquad (2.13)$$

2.3 Парна регресія

Відповідно до МНК необхідно знайти такі параметри рівняння лінійної парної регресії $y_i = a_0 + a_1 x_i$, щоб сума квадратів відхилень $\sum_{i=1}^{m} \delta_k^2$ експериментальних значень від теоретичних була б мінімальною (рисунок 2.1):

$$Q = \sum_{i=1}^{m} \delta_k^2 = \sum_{i=1}^{m} \left(y_i - (a_0 + a_1 x_i)\right)^2 \to \min . \qquad (2.14)$$

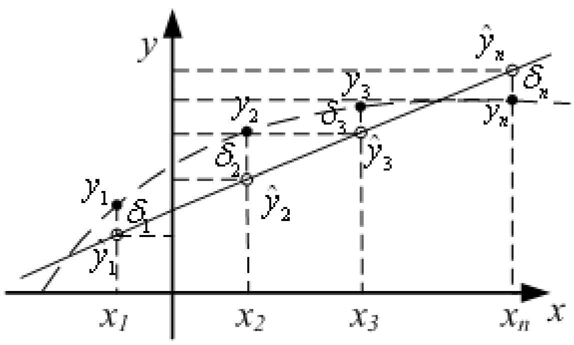

Рисунок 2.1

Для визначення коефіцієнти парної регресії запишемо систему (2.5) при умові, що $n = 1$:

$$\begin{cases} m a_0 + a_1 \sum x_i = \sum y_i; \\ a_0 \sum x_i + a_1 \sum x_i^2 = \sum y_1 x_i. \end{cases}$$

Розв'язком даної системи лінійних рівнянь будуть оцінки параметрів a_0, a_1:

$$a_1 = \frac{m\sum_{i=1}^{m} x_i y_i - \sum_{i=1}^{m} x_i \cdot \sum_{i=1}^{m} y_i}{m\sum_{i=1}^{m} x_i^2 - \left(\sum_{i=1}^{m} x_i\right)^2} = \frac{\frac{\sum_{i=1}^{m} x_i y_i}{m} - \overline{x} \cdot \overline{y}}{\frac{\sum_{i=1}^{m} x_i^2}{m} - \left(\overline{x}\right)^2};$$

$$a_0 = \overline{y} - a_1 \overline{x}$$

або

$$a_1 = \frac{K_{xy}}{\sigma_x^2} = r_{xy}\,\frac{\sigma_y}{\sigma_x},$$

де $\bar{x} = \dfrac{\sum x_i}{m}$ – середнє експериментальне змінної x;

$\bar{y} = \dfrac{\sum y_i}{m}$ – середнє експериментальне змінної y;

$\sigma_x = \sqrt{\dfrac{\sum x_i^2}{m} - (\bar{x})^2}$ – середнє квадратичне відхилення змінної x:

$\sigma_y = \sqrt{\dfrac{\sum y_i^2}{m} - (\bar{y})^2}$ – середнє квадратичне відхилення змінної y:

$K_{xy} = \dfrac{1}{m}\sum_{i=1}^{m} x_i y_i - \bar{x}\cdot\bar{y}$ – кореляційний момент;

$r_{xy} = \dfrac{K_{xy}}{\sigma_x \sigma_y}$ – коефіцієнт кореляції.

Коефіцієнти рівняння регресії, отримані на підставі МНК, називаються оцінками параметрів регресійного рівняння.

Підставивши отримані оцінки a_0 та a_1 в рівняння (2.14), отримаємо мінімальне значення Q_{MIN}:

$$Q_{\min} = m\sigma_y^2\left(1 - r_{xy}^2\right).$$

Із останнього рівняння видно, чим ближче наближається модуль коефіцієнта кореляції до одиниці $\left|r_{xy}\right| \to 1$, тим менше Q_{MIN} і тим менша розбіжність між парною лінійною регресією та теоретичною лінійною залежністю.

2.4 Множинна лінійна регресія

Характеризує залежність функції Y від декількох величин X:

$$y = a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_n x_n\,. \qquad (2.15)$$

Коефіцієнти регресії визначаються МНК. При кількості випадкових аргументів $n \geq 2$, коефіцієнти можна визначати за допомогою системи лінійних рівнянь (2.5) при виконанні заміни $x_2 = x_i^2$, $x_k = x_i^k$, але параметри множинної лінійної регресії доцільніше визначати, застосовуючи матричний спосіб визначення коефіцієнтів рівняння регресії, тоді:

$$A = \left(X^T X\right)^{-1} X^T Y ,$$

де – $A = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}$, $X = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots x_{1n} \\ 1 & x_{21} & x_{22} & \cdots x_{2n} \\ \cdots & \cdots & \cdots & \cdots \cdots \\ 1 & x_{m1} & x_{m2} & \cdots x_{mn} \end{pmatrix}$, $Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \cdots \\ y_n \end{pmatrix}$.

Мірою залежності між випадковими величинами *Y* та *X*, де $Y = (x_1, x_2, ..., x_n)$ є коефіцієнт множинної кореляції *R*, який є узагальненням парного коефіцієнта кореляції r_{ij} і обчислюється за формулами:

$$R = \sqrt{\frac{\sum\left(\hat{y}_i - \bar{y}\right)^2}{\sum(y_i - \bar{y})^2}} = \sqrt{1 - \frac{\sum\left(y_i - \hat{y}_i\right)^2}{\sum(y_i - \bar{y})^2}} , \tag{2.16}$$

де $y_i, \hat{y}_i, \bar{y}$ – відповідно експериментальне, теоретичне і середнє експериментальне значення *Y* або

$$R = \sqrt{\frac{A^T X^T Y - m\bar{y}^2}{Y^T Y - m\bar{y}^2}} = \sqrt{1 - \frac{Y^T Y - A^T X^T Y}{Y^T Y - m\bar{y}^2}} . \tag{2.17}$$

Чим ближче значення *R* до одиниці, тим краще вибрано функцію регресії $y = f(x_1, x_2, ..., x_n)$. Але коефіцієнт множинної кореляції *R* визначений за формулами (2.16, 2.17), не завжди дає правильну відповідь про вид вибраної регресії, тому рекомендується для більш точного визначення *R* і особливо при малих об'ємах вибірки *m*, застосовувати уточнений (скорегований) коефіцієнт множинної кореляції:

$$R^* = \sqrt{1 - \frac{m-1}{m-n-1}\left(1 - R^2\right)} \tag{2.18}$$

або

$$R^* = \sqrt{1 - \frac{\bar{S}^2_{ЗАЛ}}{\bar{S}^2_{Y}}} ,$$

де $\bar{S}^2_{ЗАЛ} = \dfrac{\sum\limits_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{m-n-1}$ – залишкова дисперсія,

$$\overline{S}_Y^2 = \frac{\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2}{m-1}$$ – загальна дисперсія.

Визначити дисперсії параметрів рівняння регресії та кореляційних моментів можна за допомогою кореляційної матриці (2.14).

Для множинної лінійної регресії характерне нормування коефіцієнтів регресії. Множинна лінійна регресія дає можливість порівняти вплив на випадкову величину Y різних величин X. У загальному випадку змінні x_i можуть мати різні одиниці вимірювання (кілограми, гривні, метри тощо). Отже, для того щоб порівняти і з'ясувати відносну вагомість кожного з чинників, використовують так звані нормовані коефіцієнти регресії, які визначають за формулою:

$$\alpha_i^* = a_i^* \frac{S(x_i)}{S(y)}, \qquad (2.19)$$

де α_i^* – коефіцієнт регресії після нормування;

$S(x_i)$ — виправлене середнє квадратичне відхилення змінної x_i;

$S(y)$ – виправлене середнє квадратичне відхилення ознаки y:

$$S(x_i) = \sqrt{\frac{\sum_{i=1}^{m}(x_i - \overline{x})^2}{m-1}} = \sqrt{\frac{\sum_{i=1}^{m}(x_i)^2 - m(\overline{x})^2}{m-1}}. \qquad (2.20)$$

2.5 Нелінійна регресія. Нелінійна регресія за параметрами

У випадках, коли при збільшенні степеня полінома сума квадратів відхилень залишається достатньо великою і не зменшується або фізичні, хімічні і т. д. процеси, які досліджуються, можна описати іншою, неполіноміальною залежністю, виникає необхідність подати рівняння регресії у вигляді нелінійної функції $Y = f(a_0, a_1 \ldots, a_n, X)$.

При визначенні коефіцієнтів нелінійного рівняння регресії методом найменших квадратів, система (2.1) відповідно буде нелінійною. Але як відомо розв'язок системи нелінійних рівнянь при неправильному виборі початкових значень можна і не одержати. Тому бажано за допомогою різноманітних математичних функцій перетворити (подати) нелінійну функцію $Y = f(A_0, A_1 \ldots, A_n, X)$ в лінійну $y^* = a_0 + a_1 x^*$, шляхом заміни змінної.

Даний метод добре реалізується для функцій такого вигляду:

а) гіперболічна функція $Y = A_0 + A_1 X^{-1}$, при заміні $x^* = X^{-1}$ одержуємо лінійну функцію $y^* = a_0 + a_1 x^*$;

б) дрібно-лінійна функція першого типу $Y = (A_0 + A_1 X)^{-1}$, для якої в результаті заміни $y^* = 1/Y$ одержуємо $y^* = a_0 + a_1 x^*$;

в) дрібно-лінійна функція другого типу $Y = \dfrac{X}{A_0 + A_1 X}$, для якої в результаті заміни $y^* = 1/Y$, $x^* = 1/X$, $a_0 = A_1$, $a_1 = A_0$ одержуємо $y^* = a_0 + a_1 x^*$;

г) експоненціальна $Y = A_0 \ell^{A_1 X}$, виконавши перетворення $\ln(Y) = \ln(A_0) + A_1 X$, і в результаті заміни $y^* = \ln(Y)$, $a_0 = \ln(A_1)$, одержуємо $y^* = a_0 + a_1 x^*$;

д) степенева $Y = A_0 X^{A_1}$, для якої $\ln(Y) = \ln(A_0) + A_1 \ln(X)$, і в результаті замін $y^* = \ln(Y)$, $x^* = \ln(x)$, $a_0 = \ln(A_1)$, одержуємо $y^* = a_0 + a_1 x^*$;

е) логарифмічна $Y = A_0 + A_1 \ln(X)$, для якої в результаті заміни змінної $x^* = \ln(X)$ одержуємо $y^* = a_0 + a_1 x^*$.

Після виконання перетворення послідовність визначення параметрів нелінійного рівняння, аналогічна при визначенні параметрів парної регресії.

В нелінійній регресії мірою залежності між випадковими величинами Y та X, є кореляційне відношення R', яке визначається за формулою (2.16).

У випадку, коли рівняння нелінійної регресії є лінійним відносно параметрів, наприклад $Y = A_0 + A_1 X^{-1}$, то всі розрахунки можна виконувати за допомогою матричного аналізу, при умові, що матриця X має такий вигляд:

$$X = \begin{pmatrix} 1 & 1/x_1 \\ 1 & 1/x_2 \\ \vdots & \vdots \\ 1 & 1/x_m \end{pmatrix}.$$

Контрольні питання

1. Дати означення статистичної залежності між ознаками X та $У$.
2. Що означає кореляційна залежність між ознаками X та $У$?
3. Записати модель парної лінійної регресії.
4. Чому дорівнює a_0^*?
5. Чому дорівнює a_1^*?
6. Який закон розподілу ймовірностей мають випадкові величини a_0^*, a_1^* для парної лінійної регресії?
7. Чому дорівнює $K(a_0^*, a_1^*)$?
8. Чому дорівнює S_ε?
9. Чому дорівнює надійний інтервал для a_0^*?
10. Чому дорівнює надійний інтервал для a_1^*?

11. Чому дорівнює надійний інтервал для $y_i = a_0^* + a_1^* x_i$?
12. Чому дорівнює a^* інтервал для множинної регресії?
13. Чому дорівнює надійний інтервал для множинної регресії?
14. Чому дорівнює надійний інтервал S_ε для множинної регресії?
15. Чому дорівнює коефіцієнт множинної регресії?
16. Що розуміють під нормуванням коефіцієнта регресії?
17. Чому дорівнює вектор a^* для нелінійної множинної регресії?
18. Чому дорівнює матриця X для лінійної множинної регресії?
19. Чому дорівнює матриця X для нелінійної регресії?
20. Чим вимірюється кореляційний зв'язок для нелінійних моделей регресії?

2.6 Основи дисперсійного аналізу

Дисперсійний аналіз (ДА) дозволяє оцінити наявність можливого впливу деяких факторів на досліджувану ознаку (функцію). Фактори існують якісні і кількісні, які не піддаються кількісному опису. Фактори позначаються літерами латинської абетки *A, B, C* …і т. д. Конкретні значення кожного фактора називаються рівнями цього фактора.

Ідея ДА полягає в розкладанні оцінки загальної дисперсії досліджуваної ознаки X на складові, які залежать від впливу рівнів кожного фактора, від сумісного впливу декількох факторів та від різноманітних випадкових факторів.

В залежності від кількості факторів ДА поділяється на однофакторний, двофакторний та багатофакторний дисперсійний аналіз.

В залежності від виду рівнів факторів існує три типу моделей ДА: модель I-го типу – фактори мають фіксовані рівні; модель II-го типу – фактори мають випадкові значення рівнів; модель III-го типу – деякі фактори мають фіксовані значення, а інші – випадкові рівні. Загальна теорія ДА розроблена тільки для моделі I-го типу.

Оцінка впливу факторів на досліджувану функцію визначається перевіркою нульової гіпотези рівності двох дисперсій *H0*: $S_A^2 = S_\varepsilon^2$, $M(S_A^2) = M(S_\varepsilon^2)$, при альтернативній $S_A^2 > S_\varepsilon^2$. Для перевірки гіпотези застосовується односторонній розподіл Фішера F-критерій.

Висновки про вплив факторів справедливі тільки для даного результату експерименту, при зміні діапазону, рівнів факторів, або умов проведення досліджень оцінка може змінитися.

2.6.1 Однофакторний дисперсійний аналіз

Математична модель однофакторного дисперсійного аналізу має такий вигляд:

$$x_{ij} = \bar{x} + \alpha_j + \varepsilon_{ij}, \tag{2.21}$$

де x_{ij} – значення ознаки X, одержане при i-му експерименті на j-му рівні фактора;

$\bar{x}$ – загальна середня величина ознаки X;

α_j – ефект впливу фактора на значення ознаки X на j-му рівні;

ε_{ij} – випадкова компонента, що впливає на значення ознаки X на i-му експерименті на j-му рівні.

При цьому $M(\varepsilon_{ij}) = 0$ і ε_{ij} як випадкові величини мають закон розподілу ймовірностей $N = (0;\sigma^2)$ і між собою незалежні ($\kappa_{ij} = 0$).

Відповідно до суті ДА подамо загальну суму квадратів відхилення результатів досліджень x_{ij} від загального середнього $\bar{x}$ як суму двох складових, одна із яких зумовлена впливом випадкових факторів, а інша характеризує вплив фактора на досліджену величину X:

$$\sum_{i=1}^{n_j}\sum_{j=1}^{p}(x_{ij} - \bar{x})^2 = \sum_{i=1}^{n_p}\sum_{j=1}^{p}(x_{ij} - \bar{x}_j)^2 + \sum_{j=1}^{p} n_j(\bar{x}_j - \bar{x})^2$$

або:

$$Q_0 = Q_\varepsilon + Q_X, \tag{2.23}$$

де $Q_0 = \sum_{i=1}^{n_j}\sum_{j=1}^{p}(x_{ij} - \bar{x})^2$ – загальна сума квадратів відхилень, загального середнього від кожного значення величину X, характеризує розсіювання результатів досліджень під впливом двох факторів, випадкового і фактора A;

$Q_\varepsilon = \sum_{i=1}^{n_p}\sum_{j=1}^{p}(x_{ij} - \bar{x}_j)^2$ – сума квадратів відхилень всередині групи на всіх рівнях фактора A, зумовлена впливом випадкових факторів;

$Q_X = \sum_{j=1}^{p} n_j(\bar{x}_j - \bar{x})^2$ – сума квадратів відхилень загального середнього від середніх групових значень, відхилення між рівнями фактора A, характеризує вплив фактора на досліджену величину X.

Для перевірки гіпотези про вплив факторів визначаємо виправлені дисперсії. Виправлена загальна дисперсія дорівнює:

$$S_0^2 = \frac{\sum_{i=1}^{n_j}\sum_{j=1}^{p}(x_{ij} - \bar{x})^2}{N - 1}. \tag{2.24}$$

Внутрішньогрупова дисперсія – виправлена дисперсія S_ε^2, що характеризує розсіювання всередині групи, зумовлене впливом випадкових факторів, обчислюється за формулою:

$$S_\varepsilon^2 = \frac{\sum_{i=1}^{n_j}\sum_{j=1}^{p}(x_{ij} - \overline{x}_j)^2}{N - p}, \tag{2.25}$$

де $N - p = k_1$ є кількість ступенів вільності для S_ε^2, оскільки при цьому використовується p співвідношень при обчисленні групових середніх $\overline{x}_j, j = \overline{1, p}$.

Міжгрупова дисперсія – виправлена дисперсія S_A^2, що характеризує розсіювання групових середніх $\overline{x}_j$ відносно загальної середньої $\overline{x}$, яке викликане впливом фактора A на результат експерименту ознаки X, обчислюється за формулою:

$$S_A^2 = \frac{\sum_{j=1}^{p} n_j(\overline{x}_j - \overline{x})^2}{p - 1}, \tag{2.26}$$

де p-1 = k_2 – це кількість ступенів вільності для S_A^2, оскільки групові середні варіюють відносно однієї загальної середньої $\overline{x}$.

Визначається розрахунковий критерій Фішера (статистичний критерій):

$$F^* = \frac{S_A^2}{S_\varepsilon^2}, \tag{2.27}$$

За значеннями ступенів вільності k_1 та k_2 та при заданому рівні значущості α знаходимо критичну точку (значення знаходяться за допомогою таблиці «Критичні точки розподілу Фішера (F-розподілу)».

Якщо $F^* \le F_{kp}$ то нульова гіпотеза про вплив фактора на результати досліджень відхиляється, а коли $F^* > F_{kp}$, то цим самим підтверджується вплив фактора на ознаку X.

Результати спостережень та обчислення статистичних оцінок зручно подати в упорядкованому вигляді за допомогою табл. 2.1

Таблиця 2.1

Ступінь впливу фактора А (групи)	Спостережуване значення ознаки X	Групові середні	Загальна середня

1	$x_{11}, x_{12}, x_{13}, \ldots, x_{1n}$	$\bar{x}_1 = \dfrac{\sum_{i=1}^{m} x_{1i}}{n_1}$	$\bar{x} = \dfrac{\sum_{i=1}^{n_j} \sum_{j=1}^{p} x_{ij}}{N}$, $N = \sum_{j=1}^{p} n_j$
2	$x_{21}, x_{22}, x_{23}, \ldots, x_{2n}$	$\bar{x}_2 = \dfrac{\sum_{i=1}^{m} x_{2i}}{n_2}$	
3	$x_{31}, x_{32}, x_{33}, \ldots, x_{3n}$	$\bar{x}_3 = \dfrac{\sum_{i=1}^{m} x_{3i}}{n_3}$	
. . .	....	....	
p	$x_{p1}, x_{p2}, x_{p3}, \ldots, x_{pn}$	$\bar{x}_p = \dfrac{\sum_{i=1}^{m} x_{1p}}{n_p}$	
Вид варіацій ознаки	Сума квадратів відхилень	Кількість ступенів вільності	Статистичні оцінки дисперсії
внутрішньогрупова	$\sum_{i=1}^{n_j} \sum_{j=1}^{p} (x_{ij} - \bar{x}_j)^2$	$N - p$	$S_{\varepsilon}^2 = \dfrac{\sum_{i=1}^{n_j} \sum_{j=1}^{p} (x_{ij} - \bar{x}_j)^2}{N - p}$
міжгрупова	$\sum_{j=1}^{p} n_j (\bar{x}_j - \bar{x})^2$	$p - 1$	$S_A^2 = \dfrac{\sum_{j=1}^{p} n_j (\bar{x}_j - \bar{x})^2}{p - 1}$
загальна	$\sum_{i=1}^{n_j} \sum_{j=1}^{p} (x_{ij} - \bar{x})^2$	$N - 1$	$S_0^2 = \dfrac{\sum_{i=1}^{n_j} \sum_{j=1}^{p} (x_{ij} - \bar{x})^2}{N - 1}$

Задачі

1. З’ясувати, чи істотно впливає каталізатор на кінцевий продукт хімічної реакції при $\alpha = 0{,}001$. Ступінь впливу каталізатора на кінцевий продукт заданої хімічної реакції наведено в таблиці:

Ступінь впливу каталізатора	Кінцевий продукт хімічної реакції
1	3,2; 3,1; 3,1; 2,8; 3,3; 3,0
2	2,6; 3,1; 2,7; 2,9; 2,7; 2,8
3	2,9; 2,6; 3,0; 3,1; 3,0; 2,8

4	3,7; 3,4; 3,2; 3,3; 3,5; 3,3
5	3,0; 3,4; 3,2; 3,5; 2,9; 3,1

2. В результаті проведення досліду з метою з'ясування впливу чорного пару (при рівні значущості α = 0,01) на врожайність пшениці з ділянки в 9 га (3 га були під чорним паром; 3 га – під картоплею; 3 га – під кормовими травами) дістали такі результати:

Фактор	Врожайність, ц/га
Чорний пар	26,6; 26,6; 30,6
Площа під картоплею	24,3; 25,2; 25,2
Площа під кормовими травами	26,6; 28,0; 31,0

3. Експериментально досліджувався вплив на зносостійкість колінчастих валів технології їх виготовлення – вплив фактора *A*, який має чотири рівні, тобто застосовувалися чотири технології виготовлення валів. Одержані результати наведено в таблиці:

Ступінь впливу фактора *A*	Кількість відпрацьованих місяців
A_1	9, 8, 10, 12
A_2	10, 12, 11, 8
A_3	8, 16, 10, 18
A_4	9, 18, 10, 8

При рівні значущості α = 0,01 перевірити вплив технології на зносостійкість валів.

4. Для перевірки методики навчання виробничим навикам на якість підготовки із випускників виробничо-технічного училища навмання вибирають чотири групи учнів, які після закінчення навчання за різними методиками тестувалися на кількість виготовлених однотипних деталей протягом робочої зміни. Результати тестування наведено в таблиці:

Ступінь впливу фактора *A* (методики)	Кількість виготовлених деталей за робочу зміну
A_1	60, 80, 75, 80
A_2	75, 66, 85, 80
A_3	60, 80, 65, 60
A_4	95, 85, 100, 80

При рівні значущості α = 0,05 з'ясувати вплив методики навчання на якість підготовки учнів.

5. Досліджується залежність врожайності пшениці від сорту пшениці, яких чотири. Результати досліджень наведено в таблиці:

Ступінь впливу фактора *A* (сорт пшениці)	Врожайність, ц/га
A_1	28,7; 26,7; 21,6; 25,0; 28,2
A_2	24,5; 28,5; 27,7; 28,7; 32,5

A_3	23,2; 24,7; 20,0; 24,0; 24,0
A_4	29,0; 28,7; 20,5; 28,0; 27,0

При рівні значущості α = 0,01 з'ясувати вплив сортності пшениці на її врожайність.

6. Стальні болти із різною домішкою компоненти A в сталі, з якої вони виготовлялися, були піддані випробуванням на міцність. Результати цих випробувань наведено в таблиці:

Ступінь впливу фактора A (відсоткова домішка)	Міцність, кг/мм^2
A_1	25, 28, 20, 22
A_2	29, 22, 21, 18
A_3	19, 25, 30, 22
A_4	18, 30, 24, 20

При рівні значущості α = 0,01 з'ясувати вплив добавки компоненти на міцність болта.

7. Електролампочки напругою 220 В виготовлялися на трьох заводах із використанням різних технологій. З кожної партії, що надходили в науково-дослідний інститут від кожного заводу, навмання брали по п'ять електролампочок і піддавали їх випробуванням на тривалість горіння. Результати цього експерименту наведено в таблиці:

Ступінь впливу фактора A (відсоткова добавка)	Тривалість горіння, год
A_1	90, 85, 105, 110, 95
A_2	80, 110, 115, 90, 105
A_3	75, 120, 110, 90, 85

При рівні значущості α = 0,01 з'ясувати вплив технології виготовлення на тривалість горіння лампочок.

8. На дослідних ділянках, кожна з яких має площу 6 га, досліджувалась залежність врожайності пшениці від внесення в ґрунт добрив типу A_1, A_2, A_3, A_4. Результати експерименту наведено в таблиці:

Ступінь впливу фактора A (тип добрив)	Врожайність, ц/га
A_1	25,6; 36,2; 22,8; 30,2; 32,5; 28,4
A_2	28,5; 40,6; 42,8; 36,4; 22,4; 29,6
A_3	24,4; 38,6; 48,4; 50,2; 28,4; 22,8
A_4	29,5; 52,8; 24,2; 22,8; 56,2; 48,4

При рівні значущості α = 0,01 з'ясувати вплив типу добрив, що вноситься в ґрунт, на врожайність пшениці.

9. З кожної із восьми партій однотипних заготовок навмання бралися заготовки, які оброблялися на трьох верстатах-автоматах різної

модифікації. Кількість деталей, виготовлена верстатами, досліджувалася на стандартність. Результати досліджень подано в таблиці:

Фактор A (тип верстатів-автоматів)	Кількість деталей, виготовлених верстатами-автоматами, що відповідають стандарту
A_1	100, 86, 90, 89, 95, 22, 80, 79
A_2	99, 82, 98, 88, 100, 96, 98, 100
A_3	100, 88, 86, 98, 98, 100, 99, 99

Перевірити вплив модифікації верстатів-автоматів на якість виготовлених деталей при рівні значущості $\alpha = 0{,}01$.

10. Проводилось дослідження розподілу кількості кров'яних тілець у певній одиниці об'єму крові людей, що перебували певний час у трьох зонах на різній відстані від Чорнобильської АЕС та в зоні, вільній від радіації. Результати досліджень наведено в таблиці:

Фактор A	Кількість кров'яних тілець
A_1 (в зоні АЕС)	6; 8; 3; 2; 6; 9
A_2 (на відстані 50 км)	5; 4; 10; 11; 6; 8
A_3 (на відстані 100 км)	5; 4; 13; 12; 10; 15
A_4 (вільна від радіації зона)	18; 16; 21; 20; 22; 21

При рівні значущості $\alpha = 0{,}01$ з'ясувати вплив перебування людини в певній зоні на кількість кров'яних тілець.

2.6.2 Двофакторний дисперсійний аналіз

Математична модель двофакторного дисперсійного аналізу має вигляд:

$$x_{ijk} = \bar{x} + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}, \tag{2.28}$$

де x_{ijk} – значення ознаки X в i-му експерименті на j-му рівні впливу фактора A і на k-му рівні впливу фактора B;

$\bar{x}$ – загальна середня величина ознаки X;

α_i – ефект впливу фактора A на i-му рівні;

β_j – ефект впливу фактора B на j-му рівні;

γ_{ij} – ефект одночасного впливу факторів A і B;

ε_{ijk} – випадкова компонента.

При цьому $M(\varepsilon_{ij}) = 0$ і ε_{ij}, як випадкові величини мають закон розподілу ймовірностей $N = (0;\sigma^2)$ і між собою незалежні ($\kappa_{ij} = 0$).

У разі проведення дисперсійного аналізу досліджуваний масив даних, одержаних під час експерименту, поділяють на певні групи, які різняться дією на результати експерименту певних рівнів факторів.

Вважається, що досліджувана ознака має нормальний закон розподілу, а дисперсії в кожній окремій групі здобутих значень ознаки однакові. Ці припущення необхідно перевірити.

Для перевірки нульової гіпотези про відсутність впливу факторів або спільного впливу двох факторів на ознаку *X*, результати експерименту зручно подати у вигляді таблиці, яка поділена на групи. Виходячи з даних таблиці, маємо, що:

середнє значення ознаки *X* для кожної групи:

$$\bar{x}_{jk} = \frac{\sum_{i=1}^{n} x_{ijk}}{n}. \tag{2.29}$$

Середнє значення ознаки *X* по стовпцях:

$$\bar{z}_j = \frac{\sum_{i=1}^{n}\sum_{k=1}^{q} x_{ijk}}{nq}, j = \overline{1,p}. \tag{2.30}$$

Середнє значення ознаки *X* по рядках:

$$\bar{y}_k = \frac{\sum_{i=1}^{n}\sum_{j=1}^{p} x_{ijk}}{np}, k = \overline{1,q}. \tag{2.31}$$

Загальне середнє ознаки *X*:

$$\bar{x} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{p}\sum_{k=1}^{q} x_{ijk}}{npq}. \tag{2.32}$$

Виправлена дисперсія, яка зумовлена впливом фактора *A* на ознаку *X*:

$$S_A^2 = \frac{nq\sum_{j=1}^{p}(\bar{z}_j - \bar{x})^2}{p-1} = \frac{Q_A}{p-1}. \tag{2.33}$$

Виправлена дисперсія, яка зумовлена впливом фактора *B* на ознаку *X*:

$$S_B^2 = \frac{np\sum_{k=1}^{q}(\bar{y}_k - \bar{x})^2}{q-1} = \frac{Q_B}{q-1}. \tag{2.34}$$

Виправлена дисперсія, яка зумовлена одночасним впливом на ознаку X факторів A і B:

$$S_{AB}^2 = \frac{\sum_{j=1}^{p}\sum_{k=1}^{q}(\overline{x}_{jk} - \overline{z}_j - \overline{y}_k + \overline{x})^2}{(p-1)(q-1)} = \frac{Q_{AB}}{(p-1)(q-1)}. \qquad (2.35)$$

Виправлена дисперсія, яка зумовлена впливом на ознаку X випадкових факторів:

$$S_{\varepsilon}^2 = \frac{\sum_{i=1}^{n}\sum_{j=1}^{p}\sum_{k=1}^{q}(x_{ijk} - \overline{x}_{jk})^2}{N - pq} = \frac{Q_{\varepsilon}}{N - pq}. \qquad (2.36)$$

Виправлена загальна дисперсія дорівнює:

$$S_0^2 = \frac{\sum_{i=1}^{n}\sum_{j=1}^{p}\sum_{k=1}^{q}(x_{ijk} - \overline{x})^2}{N - 1} = \frac{Q_0}{N-1}. \qquad (2.37)$$

Для двофакторного дисперсійного аналізу повинно виконуватися рівняння:

$$Q_0 = Q_A + Q_B + Q_{AB} + Q_{\varepsilon}.. \qquad (2.38)$$

Обчислюється розрахункове значення критерію Фішера:

$$F_A^* = \frac{S_A^2}{S_{\varepsilon}^2}; F_B^* = \frac{S_B^2}{S_{\varepsilon}^2}; F_{AB}^* = \frac{S_{AB}^2}{S_{\varepsilon}^2}. \qquad (2.39)$$

При заданому рівні значущості α за таблицею критерію Фішера визначаються критичні (табличні) значення критерію Фішера (критичні точки): $F_{kpA}(\alpha, k_1, k_4)$, $F_{kpB}(\alpha, k_2, k_4)$, $F_{kpAB}(\alpha, k_3, k_4)$. Якщо розрахункове значення критерію Фішера більше значення критичної точки (наприклад $F_A^* > F_{kpA}$) то нульова гіпотеза про відсутність впливу фактора A відхиляється. Аналогічно перевірюються і інші нульові гіпотези.

Таблиця 2.2

Рівень фактора B	Рівень фактора A							Середня величина за рядками	Загальна середня величина
	A_1		A_2		..	Ap			
		середня групова величина		середня групова величина			середня групова величина		
B_1	$x_{111}, x_{112}, x_{113}, \ldots x_{11n}$	$\bar{x}_{11} = \dfrac{\sum_{i=1}^{n} x_{11n}}{n}$	$x_{121}, x_{122}, x_{123}, \ldots x_{12n}$	$\bar{x}_{12} = \dfrac{\sum_{i=1}^{n} x_{12n}}{n}$	..	$x_{1p1}, x_{1p2}, x_{1p3}, \ldots x_{1pn}$	$\bar{x}_{1p} = \dfrac{\sum_{i=1}^{n} x_{1pn}}{n}$	$\bar{y}_1 = \dfrac{\sum_{k=1}^{p}\sum_{i=1}^{n} x_{1ki}}{np}$	$\bar{x} = \dfrac{\sum_{i=1}^{n}\sum_{j=1}^{p}\sum_{k=1}^{q} x_{ijk}}{N} = \dfrac{\sum_{i=1}^{n}\sum_{j=1}^{q}\sum_{k=1}^{p} x_{ijk}}{npq}$
B_2	$x_{211}, x_{212}, x_{213}, \ldots x_{21n}$	$\bar{x}_{21} = \dfrac{\sum_{i=1}^{n} x_{i21}}{n}$	$x_{221}, x_{222}, x_{223}, \ldots x_{22n}$	$\bar{x}_{22} = \dfrac{\sum_{i=1}^{n} x_{22n}}{n}$	..	$x_{2p1}, x_{2p2}, x_{2p3}, \ldots x_{2pn}$	$\bar{x}_{2p} = \dfrac{\sum_{i=1}^{n} x_{2pn}}{n}$	$\bar{y}_2 = \dfrac{\sum_{k=1}^{p}\sum_{i=1}^{n} x_{2ki}}{np}$	
...	...	...	...	...	..	..	..	..	
B_q	$x_{q11}, x_{q12}, x_{q13}, \ldots x_{q1n}$	$\bar{x}_{g1} = \dfrac{\sum_{i=1}^{n} x_{g1i}}{n}$	$x_{q21}, x_{q22}, x_{q23}, \ldots x_{q2n}$	$\bar{x}_{g2} = \dfrac{\sum_{i=1}^{n} x_{g2i}}{n}$	..	$x_{qp1}, x_{qp2}, x_{qp3}, \ldots x_{qpn}$	$\bar{x}_{qp} = \dfrac{\sum_{i=1}^{n} x_{qpn}}{n}$	$\bar{y}_q = \dfrac{\sum_{k=1}^{p}\sum_{i=1}^{n} x_{qki}}{np}$	
Середня величина за стовпцями	$\bar{z}_1 = \dfrac{\sum_{i=1}^{n}\sum_{j=1}^{q} x_{1ij}}{nq}$		$\bar{z}_2 = \dfrac{\sum_{i=1}^{n}\sum_{j=1}^{q} x_{2ij}}{nq}$		..	$\bar{z}_p = \dfrac{\sum_{i=1}^{n}\sum_{j=1}^{q} x_{qpn}}{nq}$			

Джерело, що спонукає до розсіювання	Сума квадратів відхилень	Кількість ступенів вільності	Статистичні оцінки дисперсії
Фактор A	$Q_A = nq\sum_{j=1}^{p}(\bar{z}_j - \bar{x})^2$	$k_1=p-1$	$S_A^2 = \dfrac{Q_A}{p-1}$
Фактор B	$Q_B = np\sum_{j=1}^{q}(\bar{y}_j - \bar{x})^2$	$k_2=q-1$	$S_B^2 = \dfrac{Q_B}{q-1}$
Одночасна дія факторів A і B	$Q_{AB} = n\sum_{j=1}^{q}\sum_{k=1}^{p}(\bar{x}_{jk} - \bar{z}_j - \bar{y}_j + \bar{x})^2$	$k_3=(p-1)(q-1)$	$S_{AB}^2 = \dfrac{Q_{AB}}{(p-1)(q-1)}$
Дія випадкової компоненти ε_{ijk}	$Q_\varepsilon = \sum_{i=1}^{q}\sum_{j=1}^{p}\sum_{k=1}^{n}(x_{ijk} - \bar{x}_{jk})^2$	$k_4=N-pq$	$S_\varepsilon^2 = \dfrac{Q_\varepsilon}{N-pq}$
Загальне відхилення	$Q_0 = \sum_{i=1}^{q}\sum_{j=1}^{p}\sum_{k=1}^{n}(x_{ijk} - \bar{x})^2$	$k=N-1$	$S_0^2 = \dfrac{Q_0}{N-1}$

Контрольні питання

1. У чому сутність дисперсійного аналізу?
2. Математична модель для однофакторного дисперсійного аналізу.
3. Математична модель для двофакторного дисперсійного аналізу.
4. Що таке рівень впливу певного фактора на досліджувану ознаку X?
5. Розкрити сутність x_{ijk}.
6. Властивості випадкових компонент ε_{ij}, ε_{ijk}
7. Що називають внутрішньогруповою дисперсією?
8. Що називають міжгруповою дисперсією для однофакторного дисперсійного аналізу?
9. Групові та загальні середні і формули для їх обчислення.
10. Виправлена дисперсія, що характеризує розсіювання в середні групи, та формула для її обчислення.
11. Виправлена дисперсія, що характеризує вплив фактора, та формула для її обчислення.
12. Кількість ступенів вільності для S_A^2, S_B^2, S_{AB}^2, S_ε^2.
13. Статистичний критерій для перевірки істотності впливу фактора на досліджувану ознаку X.
14. Виправлена дисперсія, що вимірює розсіювання ознаки під впливом фактора A, та формула для її обчислення.
15. Виправлена дисперсія, що вимірює розсіювання ознаки під впливом фактора B, та формула для її обчислення.
16. Виправлена дисперсія, що вимірює розсіювання ознаки під впливом факторів A і B сумісно, та формула для її обчислення.
17. Виправлена дисперсія, що вимірює розсіювання під впливом інших випадкових факторів, та формула для її обчислення.
18. Формула для обчислення середніх групових.
19. Формула для обчислення середнього значення ознаки по стовпцях.
20. Формула для обчислення середнього значення ознаки по рядках.
21. Формула для обчислення загального середнього.
22. Обчислення спостережуваних значень критеріїв F_A^*, F_B^*, F_{AB}^*.

3 ЗАВДАННЯ ДЛЯ ВИКОНАННЯ КОНТРОЛЬНОЇ РОБОТИ

3.1 Основи регресійного аналізу

1. За даними результатів спостережень величин X і Y, які наведені у таблиці 3.1, визначити статистичні оцінки a_0^*, a_1^*, a_2^* для параметрів a_0, a_1, a_2 рівняння поліноміальної регресії $y = a_0 + a_1x + a_2x^2$, надійний інтервал існування коефіцієнтів при надійній ймовірності $p = 0{,}95$, або іншій за вказівкою викладача, та оцінку дисперсії заданої функції. Обчислити кореляційне відношення R'.

Таблиця 3.1 – Дані результатів спостережень величин X і Y

Номер	x_i	Варіант №									
		1	2	3	4	5	6	7	8	9	10
		y_i	y_i	y_i	y_i	y_i	y_i	y_i	y_i	y_i	y_i
1	0,1	1,86	16,22	3,18	1,56	0,25	1,03	9,15	5,14	-2,08	0,13
2	0,2	1,62	11,05	3,57	0,86	0,11	1,45	4,93	7,01	-0,93	0,54
3	0,3	1,64	7,46	3,91	0,62	0,23	2,53	2,64	9,65	-0,07	0,77
4	0,4	1,42	5,46	4,1	0,26	0,12	4,02	1,19	11,8	0,3	1,05
5	0,5	1,4	4,4	4,07	0,18	0,22	6,81	0,56	13,11	0,76	1,03
6	0,6	1,26	3,77	4,4	0,21	0,34	11,48	0,33	13,55	1,22	1,15
7	0,7	1,16	2,09	4,26	-0,08	0,25	18,81	-0,08	14,86	1,25	1,2
8	0,8	1,33	2,81	4,31	0,1	0,59	31,6	0,03	14,97	1,52	1,15
9	0,9	1,23	2,45	4,36	0,03	0,69	52,42	-0,05	14,76	1,39	1,2
10	1,0	1,29	4,3	4,41	0,12	0,98	87,16	0,01	13,66	1,19	1,3
Номер	x_i	Варіант №									
		11	12	13	14	15	16	17	18	19	20
		y_i	y_i	y_i	y_i	y_i	y_i	y_i	y_i	y_i	y_i
1	0,1	7,75	2,73	1,94	6,89	0,81	8,22	1,01	6,17	8,21	3,14
2	0,2	4,45	3,49	3,11	4,1	0,37	3,42	0,52	2,97	4,43	2,31
3	0,3	3,22	4,04	3,83	2,58	0,23	2,54	0,39	1,67	2,05	2,05
4	0,4	2,33	4,26	4,03	1,45	0,12	0,13	0,09	0,86	1,05	1,67
5	0,5	1,91	4,64	4,3	0,9	0,03	0,09	0,06	0,56	0,48	1,55
6	0,6	1,7	4,64	4,57	0,65	0,02	-0,85	0,12	0,47	0,53	1,53
7	0,7	1,21	4,69	4,44	0,18	0,03	-1,11	-0,16	0,09	0,22	1,18
8	0,8	1,25	4,79	4,74	0,26	0,07	-0,28	0,02	0,23	0,07	1,32
9	0,9	1,06	4,8	4,76	0,13	0,02	0,02	-0,04	0,11	0,05	1,19
10	1,0	1,05	4,85	4,92	0,18	0,11	1,11	0,04	0,17	0,15	1,23
Номер	x_i	Варіант №									
		21	22	23	24	25	26	27	28	29	30
		y_i	y_i	y_i	y_i	y_i	y_i	y_i	y_i	y_i	y_i
1	0,1	5,05	17,85	2,36	3,24	0,21	14,6	11,25	3,84	2,49	4,98
2	0,2	3,02	9,42	3,99	1,9	0,32	10,21	6,25	2,46	1,74	3,76
3	0,3	0,52	5,36	5,25	1,47	0,64	8,43	2,92	1,81	1,54	3,2
4	0,4	-2,32	3,16	5,46	1,03	0,93	7,15	1,26	1,27	1,2	2,65
5	0,5	-4,4	2,11	5,0	0,89	1,52	6,43	0,65	1,08	1,11	2,45
6	0,6	-5,33	1,58	3,82	0,87	1,54	5,97	0,45	1,04	1,12	2,41
7	0,7	-5,29	0,92	1,69	0,52	3,15	5,28	0,03	0,68	0,79	2,07
8	0,8	-3,52	0,87	-,25	0,66	3,72	5,14	0,15	0,81	0,94	2,23
9	0,9	-1,14	0,63	-2,42	0,55	7,14	4,8	0,03	0,69	0,83	2,14
10	1,0	1,66	0,59	-4,05	0,6	9,91	4,65	0,09	1,24	0,89	2,22

2. За даними результатів спостережень величин X і Y, які наведені у таблиці 3.1, знайти статистичні оцінки a_0^*, a_1^* для параметрів a_0, a_1 рівняння нелінійної регресії, виконавши перетворення функції $y(x)$ до лінійної. Рівняння нелінійної регресії наведені в таблиці 3.2. Обчислити кореляційне

відношення R'. За визначеними коефіцієнтами R', зробити висновок про міру щільності зв'язку вибраних рівнянь регресії.

3. Побудувати в одній системі координат графіки: експериментальних точок, поліноміальної регресії та нелінійної регресії.

Таблиця 3.2 – Рівняння нелінійної регресії

Номер	Функція $y(x)$	Номер	Функція $y(x)$	Номер	Функція $y(x)$
1.	$\frac{1}{5b_0 + b_1 x}$	11.	$\frac{50b_0}{b_1 + 5x}$	21.	$b_0 + 25b_1 \cos 5x$
2.	$b_0 + \frac{10b_1}{x}$	12.	$\frac{50b_0 x}{0.5b_1 + x}$	22.	$\frac{50b_0}{b_1 + 9x^2}$
3.	$\frac{10x}{b_0 + 10b_1 x}$	13.	$50b_0 \ell^{-b_1 \frac{0.5}{x}}$	23.	$b_0 + 25b_1 \sin 4x$
4.	$b_0 + \frac{10b_1}{\ell^{10x}}$	14.	$b_0 + 50b_1 \ell^{-5x}$	24.	$\frac{1}{b_0 + 10b_1 \cdot \ln(x+1)}$
5.	$b_0 \ell^{(10b_1 x)}$	15.	$b_0 + \frac{b_1}{25x^2}$	25.	$\frac{x^2}{b_0 + b_1 (0.1x)^2}$
6.	$5b_0 \cdot 10^{(10x \cdot b_1)}$	16.	$b_0 + \frac{0.5b_1}{(10x)^3}$	26.	$b_0 + \frac{1}{b_1 \sqrt{x}}$
7.	$\frac{5}{b_0 + b_1 \ell^{7x}}$	17.	$\frac{10b_0}{1 + b_1 (10x)^2}$	27.	$\frac{25b_0}{b_1 + (3x)^3}$
8.	$b_0 + 10b_1 \sqrt{50x}$	18.	$\frac{1}{b_0 + b_1 25x^2}$	28.	$b_0 + \frac{2b_1}{\ln(x+1)}$
9.	$20(b_0 + b_1 \lg x)$	19.	$\frac{1}{b_0 + b_1 (4x)^3}$	29.	$\frac{20b_0}{b_1 + \sqrt{5x}}$
10.	$b_0 + 2b_1 \ln(10x)$	20.	$\frac{1}{b_0 + b_1 \sqrt{10x}}$	30.	$20(b_0 + b_1 \ell^{-5x})$

Приклад. 1. Визначити статистичні оцінки a_0^*, a_1^*, a_2^* для параметрів a_0, a_1, a_2 рівняння поліноміальної регресії $y = a_0 + a_1 x + a_2 x^2$, надійний інтервал існування коефіцієнтів та оцінку дисперсії заданої функції. Обчислити кореляційне відношення R'.

2. Знайти статистичні оцінки a_0^*, a_1^* для параметрів a_0, a_1 рівняння нелінійної регресії, виконавши перетворення функції $y = a_0 + {a_1}/{\sqrt{x}}$ до лінійної. Обчислити кореляційне відношення R'.

3. За визначеними коефіцієнтами R', зробити висновок про адекватність вибраних рівнянь регресії.

4. Побудувати в одній системі координат графіки: експериментальних точок, поліноміальної регресії та нелінійної регресії.

Результати спостережень величин X, Y подані в таблиці 3.3.

Таблиця 3.3

i	1	2	3	4	5	6	7	8	9	10
x_i	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
y_i	26,5	17,3	9,8	7,8	6,5	5,5	5,1	4,5	4,2	4,0

Розв'язування

1. Визначаємо коефіцієнти поліноміальної залежності.

1.1. Обчислимо коефіцієнти рівняння регресії методом найменших квадратів,

а) запишемо систему лінійних рівнянь для трьох невідомих, розв'язок системи – коефіцієнти полінома:

$$\begin{cases} ma_0 + a_1\sum x_i + a_2\sum x_i^2 = \sum y_i \\ a_0\sum x_i + a_1\sum x_i^2 + a_2\sum x_i^3 = \sum y_i x_i \\ a_0\sum x_i^2 + a_1\sum x_i^3 + a_2\sum x_i^4 = \sum y_i x_i^2, \end{cases}$$

$$\sum x_i = 5{,}5; \sum x_i^2 = 3{,}85; \sum x_i^3 = 3{,}025; \sum x_i^4 = 2{,}533;$$
$$\sum y_i = 91{,}2; \sum y_i x_i = 33{,}67; \sum y_i x_i^2 = 19{,}473.$$

$$\begin{cases} 10a_0 + 5{,}5a_1 + 3{,}85a_2 = 91{,}2 \\ 5{,}5a_0 + 3{,}85a_1 + 3{,}025a_1 = 33{,}67 \\ 3{,}85a_0 + 3{,}025a_1 + 2{,}533a_1 = 19{,}473, \end{cases}$$

$$D = \begin{vmatrix} 10 & 5{,}5 & 3{,}85 \\ 5{,}5 & 3{,}85 & 3{,}025 \\ 3{,}85 & 3{,}025 & 2{,}533 \end{vmatrix} \to 0{,}436$$

$$a_0 = 30{,}53; a_1 = -72{,}07; a_2 = 47{,}35.$$

б) визначаємо стандартну похибку:

$$S(\delta)=\sqrt{\frac{\sum\delta_i^2}{m-n-1}};m=10,n=2,$$
$$\delta_i=y_i-\hat{y}_i,$$
$$\hat{y}=30{,}53-72{,}07x+47{,}35x^2,$$
$$\delta_1=y_1-\hat{y}_1=2{,}7,$$
$$\delta_2=y_2-\hat{y}_2=-0{,}7,$$
$$\delta_3=y_3-\hat{y}_3=-3{,}37.$$

Як приклад, обчислення стандартної похибки наведені тільки для трьох точок, для інших точок розрахунки виконуються аналогічно.

$$S(\delta)=\sqrt{\frac{31{,}1}{10-2-1}}\approx 2{,}11.$$

в) оцінка середньоквадратичного відхилення коефіцієнтів a_n виражається формулою:

$$S(a_i)=S(\delta)\sqrt{\frac{D_{(i+1)(i+1)}}{D}}.$$

Оцінка середньоквадратичного відхилення коефіцієнта a_0:

$$S(a_0)=S(\delta)\sqrt{\frac{D_{11}}{D}},$$
$$D=0{,}436;D_{11}=0{,}603,$$
$$S(a_0)=2{,}11\sqrt{\frac{0{,}603}{0{,}436}}=2{,}48.$$

г) визначаємо надійний інтервал $\Delta a_0=t_{(P,k)}S(a_0)$ для коефіцієнта a_0, Коефіцієнт Стьюдента $t_{(P,k)}$ знаходиться за спеціальною таблицею, якщо $k=m\text{-}n\text{-}1=7$, $p=0{,}95$ коефіцієнт Стьюдента дорівнює 2,36.

$$\Delta a_i=t_{(P,k)}S(a_i),$$
$$\Delta a_0=2{,}36\cdot 2{,}48=5{,}85,$$
$$a_0=30{,}53\pm 5{,}85;\;\; p=0{,}95.$$

Аналогічно виконуються розрахунки коефіцієнтів a_1, a_2.

д) оцінку дисперсії функції $y=a_0+a_1x+a_2x^2$ визначаємо за формулою:

$$S^2(y)=\sum\left(\frac{\partial y}{\partial a_i}\right)^2 S^2(a_i)+2\sum\sum\frac{\partial y}{\partial a_i}\frac{\partial y}{\partial a_j}S(a_i)S(a_j)k(a_i a_j).$$

Оцінка кореляційного моменту розраховується за формулою:

$$k(a_i a_j)=S^2(\delta)\cdot\frac{D(i+1)(j+1)}{D}.$$

Визначимо оцінку кореляційного моменту між коефіцієнтами a_0 і a_1:

$$k(a_0 a_1)=S^2(\delta)\frac{-D_{12}}{D}=2{,}11^2\,\frac{-2{,}287}{0{,}436}=-23{,}35\,.$$

Аналогічно обчислюються оцінка кореляційних моментів $k(a_0, a_2)$ та $k(a_1, a_2)$.

Відповідні частинні похідні $\frac{\partial y}{\partial a_0}=1; \frac{\partial y}{\partial a_1}=x; \frac{\partial y}{\partial a_2}=x^2$, тоді:

$$\begin{aligned}S^2(y)&=S^2(a_0)+x^2S^2(a_1)+x^4S^2(a_2)+2xS(a_0)S(a_1)k(a_0a_1)+\\&+2x^2S(a_0)S(a_2)k(a_0a_2)+2x^3S(a_1)S(a_2)k(a_1a_2)=\\&=6{,}15-286{,}79x+949{,}43x^2-17587{,}11x^3+84{,}16x^4.\end{aligned}$$

е) кореляційне відношення R' визначаємо за формулою (2.16). Результати розрахунків наведені в таблиці 3.4:

Таблиця 3.4

i	x_i	y_i	$\hat{y}_i=30{,}53-72{,}07x_i+47{,}35x_i^2$
1	0,1	26,5	23,8
2	0,2	17,3	18,01
3	0,3	9,8	13,17
4	0,4	7,8	9,28
5	0,5	6,5	6,33
6	0,6	5,5	4,33
7	0,7	5,1	3,28
8	0,8	4,5	3,18
9	0,9	4,2	4,02
10	1,0	4,0	5,1
		$\bar{y}=9{,}12$	

тоді $R'=\sqrt{1-\frac{\sum(y_i-\hat{y}_i)^2}{\sum(y_i-\bar{y})^2}}=\sqrt{1-\frac{\sum(y_i-\hat{y}_i)^2}{\sum(y_i-9{,}12)^2}}=0{,}967\,.$

1.2 Визначимо всі необхідні параметри за допомогою матричного обчислювання,

а) коефіцієнти рівняння регресії визначаємо за виразом:

$$A = \left(X^T X\right)^{-1} X^T Y,$$

$$\text{де} - A = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}; \quad X = \begin{pmatrix} 1 & 0,1 & 0,01 \\ 1 & 0,2 & 0,04 \\ 1 & 0,3 & 0,09 \\ 1 & 0,4 & 0,16 \\ 1 & 0,5 & 0,25 \\ 1 & 0,6 & 0,36 \\ 1 & 0,7 & 0,49 \\ 1 & 0,8 & 0,64 \\ 1 & 0,9 & 0,81 \\ 1 & 1,0 & 1,0 \end{pmatrix}; \; Y = \begin{pmatrix} 26,5 \\ 17,3 \\ 9,8 \\ 7,8 \\ 6,5 \\ 5,5 \\ 5,1 \\ 4,5 \\ 4,2 \\ 4,0 \end{pmatrix}.$$

Виконуємо розрахунки:

$$\overline{a} = (X^T X)^{-1} X^T Y = \begin{pmatrix} 30,53 \\ -72,07 \\ 47,35 \end{pmatrix},$$

таким чином маємо: a_0 = 30,53; a_1 = -72,07; a_2 = 47,35.

б) визначаємо оцінки дисперсії параметрів рівняння регресії і кореляційних моментів за допомогою кореляційної матриці:

$$K_{ij} = (X^T X)^{-1} S^2(\delta) = \begin{vmatrix} 6,15 & -23,33 & 18,52 \\ -23,33 & 107,22 & -92,57 \\ 18,52 & -92,57 & 84,16 \end{vmatrix}.$$

в) на підставі одержаних розрахунків записуємо оцінку дисперсії функції $y = a_0 + a_1 x + a_2 x^2$:

$$S^2(y) = S^2(a_0) + x^2 S^2(a_1) + x^4 S^2(a_2) + 2xS(a_0)S(a_1)k(a_0 a_1) +$$
$$+ 2x^2 S(a_0)S(a_2)k(a_0 a_2) + 2x^3 S(a_1)S(a_2)k(a_1 a_2) =$$
$$= 6,15 - 286,79x + 949,43x^2 - 17587,11x^3 + 84,16x^4.$$

г) кореляційне відношення R' визначаємо за формулою:

$$R' = \sqrt{1 - \frac{Y^T Y - A^T X^T Y}{Y^T Y - m\bar{y}^2}} = \sqrt{\frac{A^T X^T Y - m\bar{y}^2}{Y^T Y - m\bar{y}^2}} = 0{,}967 .$$

2. Визначаємо статистичні оцінки a_0^*, a_1^* параметрів a_0, a_1 рівняння нелінійної регресії $y = b_0 + {b_1}/{\sqrt{x}}$.

а) для перетворення даної функції до лінійної, вводимо заміну $x^* = {1}/{\sqrt{x}}$ і одержуємо $y = b_0 + b_1 x^*$. Результати розрахунків наведені в таблиці 3.5

Таблиця 3.5

i	*x*	*y*	*x**
1	0,1	26,5	3,16
2	0,2	17,3	2,24
3	0,3	9,8	1,83
4	0,4	7,8	1,58
5	0,5	6,5	1,41
6	0,6	5,5	1,29
7	0,7	5,1	1,19
8	0,8	4,5	1,12
9	0,9	4,2	1,05
10	1,0	4,0	1,0

б) визначаємо оцінки параметрів a_0 і a_1 лінійної залежності методом найменших квадратів. Виконуємо заміну $a_0 = b_0, a_1 = b_1$, тоді знаходимо розв'язок системи лінійних рівнянь:

$$\begin{cases} ma_0 + a_1 \sum x_i^* = \sum y_i \\ a_0 \sum x_i^* + a_1 \sum \left(x_i^*\right)^2 = \sum y_i x_i^*, \end{cases}$$

$$a_1 = \frac{\dfrac{\sum x_i^* y_i}{m} - \bar{x}^* \bar{y}}{\dfrac{\sum \left(x_i^*\right)^2}{m} - \left(\bar{x}^*\right)^2} = 10{,}73,$$

$$a_0 = \bar{y} - a_1 \bar{x}^* = -7{,}91,$$

де $\bar{x}^*$ – середнє експериментальне значення величин x_i^*.

в) визначаємо кореляційне відношення R'. Результати розрахунків наведені в таблиці 3.6:

Таблиця 3.6

i	x_i	y_i	$\hat{y}_i = -7{,}91 + 10{,}73\frac{1}{\sqrt{x_i}}$
1	0,1	26,5	26,01
2	0,2	17,3	16,07
3	0,3	9,8	11,67
4	0,4	7,8	9,05
5	0,5	6,5	7,26
6	0,6	5,5	5,94
7	0,7	5,1	4,91
8	0,8	4,5	4,08
9	0,9	4,2	3,4
10	1,0	4,0	2,82
		$\bar{y} = 9{,}12$	

тоді $R' = \sqrt{1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}} = \sqrt{1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - 9{,}12)^2}} = 0{,}9898$.

3. Кореляційне відношення R', яке визначене в прикладі 1.1, е (с. 30) становить R' = 0,967, а за результатами розрахунків прикладу 2, в кореляційне відношення – 0,9898.

Висновок: міра щільності зв'язку нелінійної регресії $y = b_0 + {b_1}/{\sqrt{x}}$ більша, ніж для поліноміальної регресії. Чим ближче значення R' до +1, тим краще вибрано функцію регресії.

4. Графіки експериментальних точок – (1), поліноміальної регресії – (2) та нелінійної регресії – (3), подані на рисунку 3.1.

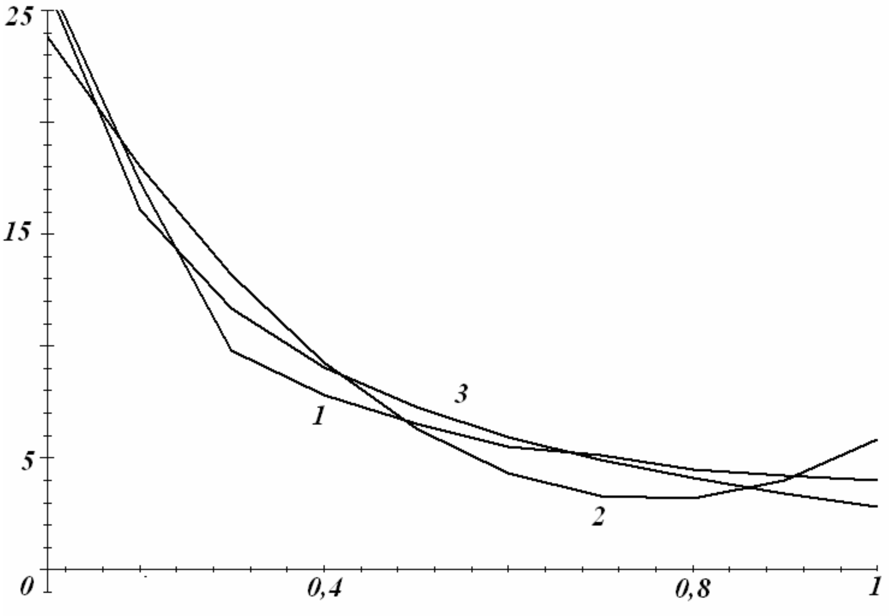

Рисунок 3.1

3.2 Елементи дисперсійного аналізу

Для рівня значущості $\alpha = 0,05$ визначити вплив факторів A і B та їх сумісний вплив на досліджувану ознаку.

1. Досліджується вплив на зносостійкість деталей таких факторів: A – матеріал для виготовлення деталей(застосовували три види сталі) і B – технологія виготовлення деталей (дві технології). Результати експерименту наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	10; 7; 8; 6; 12; 8; 11; 10; 14; 13	8; 14; 6; 10; 16; 14; 13; 12; 11; 15	15; 12; 11; 9; 8; 13; 11; 12; 16; 14;
B_2	12; 13; 6; 9; 8; 11; 10; 10; 13; 17	11; 12; 12; 16; 13; 8; 10; 9; 8; 15	13; 12; 14; 8; 6; 8; 16; 12; 14; 16

2. Досліджується вплив на врожайність ячменю таких факторів A: посів здійснюється після чорного пару – A_1; після коренеплодів – A_2; після колосових культур – A_3; B – сортність ячменю (три сорти). Результати експерименту наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	34,2; 30,6; 36,8; 35; 32,5; 34,2; 33,4; 36	42,5; 40,4; 44,6; 46,8; 39,4; 38,6; 45,8; 49,3	44,2; 46; 45,6; 48; 49,3; 45,8; 42,3; 40,8
B_2	32,5; 30,4; 39,4; 40,3; 36,4; 38,9; 39,8; 42	30,3; 35,3; 36,8; 40,5; 28,4; 33,2; 39,1; 26,9	40,3; 45; 46,8; 30,2; 48,8; 50,2; 39; 38,5
B_3	33,3; 34,8; 39,2; 35; 32,4; 34; 39,8; 40,8	30,4; 36; 40,5; 44,4; 30,8; 42,5; 46; 33,5	32,3; 29,8; 34,3; 42; 34,8; 31,6; 40; 29,6

3. Досліджується вплив на міцність чавуну двох факторів A і B: фактор A – вміст кремнію в чавуні, а саме: A_1 – 0,24%; A_2 – 0,42%; A_3 – 0,52%; B – температурний режим плавлення (два режими). Результати експерименту наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	40,2; 40,8; 38,2; 39,6; 42,4; 44,5; 40,1; 38,8 41,5; 40,8	42,5; 43,4; 44,5; 46,4; 40,1; 36,5; 40,3; 41,8; 38; 43,5	49,2; 50,2; 48,4; 50; 52,5; 38,4; 49,8; 50,4; 51,8; 49
B_2	33,4; 36,5; 34,4; 40,2; 42; 30,2; 31,8; 35,5; 34; 41,8	31,6; 33,4; 38,4; 35; 38,9; 29,5; 28,4; 30,6; 32,9; 43	29,3; 35,6; 36; 26,8; 38; 28,5; 30,6; 40,2; 33,3; 30,5

4. Досліджується вплив на врожайність кукурудзи двох факторів: A – внесення добрив у ґрунт (три добрива); B – глибина поливу землі (три глибини поливу). Результати експерименту наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	30,2; 30,8; 31,6; 32; 32,6; 28,9; 30,5; 32,6; 33; 31,5	28,4; 29,9; 30,6; 44,3; 36,2; 42,3; 28,2; 26,5; 34,3; 26,5	40,2; 42,3; 42,7; 43,5; 44; 36,8; 38,9; 45,3; 46,2; 45,4
B_2	44,2; 42,8; 43,7; 46,5; 46,9; 40,5; 45,6; 38,4; 32,5; 44,6	42,4; 43,5; 40,6; 36,8; 40; 36,4; 38,5; 43,2; 34,6; 39,8	42,3; 43,4; 45,2; 44; 36,5; 29,8; 25,4; 43,2; 45; 46,8
B_3	40,2; 36,4; 36,9; 41,8; 40,4; 34,8; 38,5; 35; 38,6; 42,4	38,5; 33,4; 30,2; 29,4; 40,1; 26,2; 25,4; 44,1; 30,6; 34,5	43,2; 44,5; 39,5; 32,5; 45; 40,8; 36,3; 43,5; 47,8; 49

5. Досліджується вплив факторів A і B на кількість виготовлених втулок зі ста взятих, які відповідають нормам стандарту: A – використали дві технології виготовлення; B – заготовки надходили із трьох заводів. Результати експерименту наведено в таблиці

Фактор B	Фактор A	
	A_1	A_2
B_1	90; 88; 90; 96; 98; 76; 80; 95; 85; 80	100; 99; 82; 98; 95; 80; 96; 95; 99; 91
B_2	79; 88; 92; 76; 80; 83; 85; 90; 96; 75	81; 82; 100; 98; 89; 85; 96; 98; 75; 97
B_3	82; 78; 75; 79; 80; 81; 86; 89; 75; 90	80; 86; 90; 91; 78; 76; 75; 82; 73; 82

6. Досліджується вплив факторів A і B на продуктивність праці підприємства певної галузі промисловості A – фондозабезпеченість (три рівні); B – рівень оплати праці робітникам (два рівні). Результати експерименту наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	14,85; 11,94; 10,5; 12,35; 15,62; 13,2; 10,62; 12,82; 11,48; 13,5	6,42; 5,23; 4,96; 5,6; 9,82; 10,23; 12,44;16,5; 5,41; 6,32	7,82; 9,63; 12,92; 10,82; 9,36; 5,11; 13,52; 14,2; 8,96; 9,92
B_2	12,5; 13,8; 14,9; 12,6; 10,85; 11,96; 12,6; 13,42; 16;17,2	10,2; 10,85; 12,34; 11,95; 12,4; 14,92; 9,86; 9,62; 8,36; 13,62	13,62; 12,55; 14,7; 13,25; 14,66; 8,35; 10,96; 11,62; 6,12; 15,66

7. Експериментально досліджувався вплив на зносостійкість деталей факторів A і B: фактор A – тип сталі (три типи); фактор B – технологія виготовлення (дві технології). Результати експерименту наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	10; 8; 6; 9; 5; 12; 5; 8; 10; 11	8; 12; 12; 10; 11; 6; 10; 10; 9; 5	15; 14; 14; 8; 8; 13; 10; 11; 9; 6
B_2	12; 9; 9; 6; 6; 5; 10; 8; 8; 9	12; 13; 13; 14; 15; 8; 9; 10; 11; 11	13; 13; 10; 5; 5; 10; 15; 14; 14; 10

8. Експериментально досліджувався вплив факторів A і B на опріснення морської води: фактор A – тип опріснювача (три типи); фактор B – три різні лабораторії. Результати експерименту наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	3,6; 3,2; 3,4; 4,1; 3,5; 4,2; 3,5; 3,8; 4,2; 3	2,92; 2,84; 2,88; 3,2; 3,45; 3,02; 2,12; 2,26; 2,43; 3,5	2,7; 2,75; 2,97; 3,2; 4,15; 2,63; 2,49; 3,25; 3,4; 4,2
B_2	4,2; 4; 4,25; 4,35; 4,5; 3,6; 3,2; 3,2; 3,6; 3,8	3,33; 3,35; 4,2; 2,93; 2,65; 2,96; 2,25; 3,8; 3,96; 4,2	3,75; 3,87; 3,64; 2,95; 2,25; 3,85; 3,99; 4,2; 4,15; 3,14
B_3	3,46; 3,45; 4,25; 4,8; 4,1; 4,05; 3,81; 3,62; 3,4; 3,02	3,42; 3,49; 2,99; 4,1; 2,65; 3,11; 3,12; 4,41; 4,0; 3,8	3,63; 3,56; 2,99; 3,79; 4,12; 3,12; 2,05; 3,81; 3,79; 4,0

9. Експериментально досліджувався вплив факторів A і B на врожайність цукрових буряків в ц/га: фактор A – добрива (три види); фактор B – сума температур за період вегетації (три рівні). Результати досліджень наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	365; 360; 370; 385; 350; 340; 342; 340; 365; 380	403; 410; 412; 416; 345; 374; 450; 430; 402; 412	452; 440; 403; 395; 382; 444; 410; 420; 433; 390
B_2	379; 381; 390; 420; 400; 402; 380; 340; 410; 390	445; 436; 470; 412; 390; 396; 380; 445; 444; 389	433; 391; 340; 455; 460; 405; 399; 413; 449; 401
B_3	332; 450; 420; 445; 390; 420; 422; 444; 380; 395	330; 413; 425; 449; 385; 399; 440; 412; 405; 382	325; 340; 412; 402; 390; 399; 375; 399; 401; 455

10. Досліджувався вплив факторів A і B на пружність сталі в умовних одиницях: фактор A – % нікелю (три рівні); фактор B – % марганцю (два рівні). Результати досліджень наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	36,4; 38,7; 36,5; 37,5; 39; 40,2; 38,5; 42,6; 35,6; 38,4	41,2; 42; 42,8; 44,3; 45,2; 44,3; 42,4; 39,5; 38,4; 39,6	36,5; 39,8; 42,4; 45,8; 48,4; 49,5; 37,2; 38,4; 40,2; 40,5
B_2	39,2; 42,3; 44,5; 40,5; 38,1; 40,8; 45,3; 41,8; 38,7; 42	39,7; 38,4; 42,5; 44,3; 47,2; 48,4; 45,2; 46,4; 49,2; 49,8	40,8; 43,2; 41,8; 44,7;50; 42,8; 35,6; 38,9; 47,2; 48,2

11. Досліджувався вплив факторів A і B на твердість сплаву: фактор A – % нікелю (три рівні); фактор B – % хрому (три рівні). Результати досліджень наведено в таблиці, де твердість подається в умовних одиницях

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	53,2; 53,9; 54,8; 55,9; 62,2; 66,8; 70; 58,2; 54,4; 52,3	55,4; 66,7; 77,2; 53,2; 65,4; 66,2; 53,2; 65,4; 66,2; 53,2	68,3; 69,8; 74,7; 79,2; 53,4; 61,5; 58,4; 59,8; 76,2; 78,3
B_2	67,2; 66,2; 55,3; 53; 72,3; 52,4; 74,2; 52; 63,2; 53,2	77,2; 65,4; 53,9; 65,1; 63,4;61,2; 71,4; 74,2; 54,2; 53,8	77,9; 62,3; 68,9; 64,5; 73,2; 53,1; 55,2; 54,4; 76,8; 78,9
B_3	70,2; 72,1; 54,4; 53,1; 73,4; 74,8; 75,2; 53; 54,2; 67,2	69,2; 65,4; 70,4; 55,4; 62,3; 72,5; 74,4; 70,5; 53,1; 54,2	75,5; 76,4; 54,2; 56,1; 62,3; 64,8; 73,4; 75,6; 79,2; 53,5

12. Досліджувався вплив на продуктивність праці факторів A і B: фактор A – плинність кадрів, % (три рівні); фактор B – рівень оплати праці, тис. грн на рік (два рівні). Результати досліджень наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	15,62; 14,3; 14,25; 15,81; 16,35; 15,61; 14,3; 12,5; 11,2; 6,5	10,83; 10,2; 13,4; 16,25; 12,2; 5,45; 6,41; 8,93; 13,44; 15,66	12,44; 14,5; 7,6; 6,75; 8,96; 16,37; 9,82; 7,83; 10,53; 8,96
B_2	16,52; 14,21; 6,85; 8,7; 10,43; 13,5; 12,8; 11,6; 6,72; 8,9	13,24; 8,16; 9,44; 10,8; 14,56; 12,46; 11,83; 10,99; 16,42; 15,34	6,81; 5,74; 10,36; 14,57; 12,44; 13,47; 15,25; 13,4; 5,07;6,8

13. Досліджувався вплив факторів A і B на приріст урожаю картоплі з гектара: фактор A – внесення в ґрунт добрива (три рівні за кількістю у відсотках); фактор B – поливання ґрунту (два рівні проникнення вологи на глибину ґрунту). Результати досліджень наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	4,25; 4,5; 5,6; 6,8; 4,05; 4,8; 7,6; 8,2; 4,02; 6,06	6,25; 4,95; 4,26; 8,29; 8,8; 9,25; 7,44; 7,8; 4,82; 5,61	5,44; 5,23; 9,82; 8,9; 4,35; 6,81; 7,84; 6,51; 4,08; 6,52
B_2	7,45; 4,05; 8,25; 9,6; 4,06; 5,25; 6,73; 5,76; 9,21; 4,01	8,28; 6,44; 7,35; 4,9; 4,22; 7,42; 8,82; 9,5; 4,08; 5,8	6,32; 7,81; 8,92; 8,6; 4,02; 5,21; 4,21; 9,47; 9,81 10,22

14. Досліджувався вплив факторів A і B на масу корів у кг: фактор A – три види соковитих кормів; фактор B – два види грубих кормів з домішкою певного відсотка вітамінів. Результати досліджень наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	455,6; 460,2; 350,2; 500; 521,6; 534,2; 605; 340; 390; 395,5	435,6; 489,6; 572,5; 399; 480; 550,6; 580; 341,5; 382,6; 599,5	331,4; 340,5; 390,6; 405,6; 545,7; 596,2; 320,2; 305,8; 421,6; 399,5
B_2	446,2; 480,5; 620,8; 700; 721,6; 750,2; 440,2; 600; 430,8; 444,6	600; 595,6; 401,8; 321,8; 340,4; 600; 431,8; 549,6; 590; 300,6	443,8; 389,5; 541,3; 590,6; 555,4; 481,6; 405,6; 311,8; 300,6; 375,8

15. Досліджувався вплив факторів A і B на вміст протеїну в курячих яйцях: фактор A – три типи зернових; фактор B – два типи комбікормів. Результати досліджень наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	13,2; 15,6; 18,2; 13,4; 13,5; 16,4; 17,5; 14,9; 19,2; 13,1	13,9; 16,5; 14,4; 18,2; 13,1; 13,9; 19,1; 17,1; 13,2; 14,5	13,4; 18,9; 14,2; 13,5; 16,2; 13,1; 14,1; 19,1; 13,8; 14,2
B_2	14,9; 15,8; 19,2; 19,4; 18,5; 13,2; 16,4; 13,1; 17,6; 16,5	15,4; 13,2; 16,2; 13,1; 19,1; 16,2; 13,5; 14,5; 16,2; 14,1	15,4; 17,2; 18,4; 13,1; 19,3; 14,2; 15,2; 16,4; 13,1; 13,9

16. Досліджувався вплив факторів A і B на приріст урожаю соняшнику, ц/га: фактор A – три види добрив; фактор B – глибина зрошення (дві глибини). Результати досліджень наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	6,2; 6,4; 6,3; 7,2; 8; 6,1; 7,2; 7,4; 8,2; 6,3	6,8; 7,2; 8,3; 9,2; 6,2; 7,1; 7,5; 6,2; 6,8; 9,4	9,4; 6,2; 6,8; 8,8; 8,5; 6,1; 9,2; 8,2; 7,6; 8,1
B_2	8,3; 9,1; 6,2; 6,8; 7,4; 8,2; 6,5; 8,3; 9,2; 6,5	7,4; 9,4; 6,5; 6,1; 7,2; 8,3; 9,2; 6,4; 9,4; 8,1	9,2; 7,5; 6,3; 8,9; 7,9; 7,2; 6,3; 9,3; 9,4; 8,2

17. Досліджувався вплив факторів A і B на жирність молока, %: фактор A – три види соковитих кормів; фактор B – два види біодомішок. Результати досліджень наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	3,25; 3,45; 3,55; 4,04; 4,08; 4,2; 3,3; 3,8; 3,45; 3,25	4,2; 3,95; 3,33; 4,1; 3,5; 3,42; 3,49; 3,59; 3,68; 3,79	3,99; 3,89; 4,32; 4,23; 4,4; 3,29; 3,25; 3,11; 4,45; 4,05
B_2	3,41; 3,45; 3,5; 4,45; 4,25; 4,33; 4,5; 4,29; 3,42; 3,41	3,3; 3; 3,42; 4,2; 4,29; 4,39; 3,8; 3,92; 3,99; 4,05	4,2; 3,21; 3,2; 3,11; 4,29; 4,41; 4,5; 4,48; 3,81; 4,29

18. Досліджувався вплив факторів A і B на річні надої молока в тоннах: фактор A – три види соковитих кормів; фактор B – два види грубих кормів із різним рівнем частки, %, спеціальних вітамінів. Результати досліджень наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	5,5; 4,9; 6; 6,5; 8,9; 2,9; 2,5; 6; 10; 9,5	9,5; 12; 6,3; 4,5; 3,9; 8,5; 8,6; 5,9; 12,4; 6,9	12,5; 8,9; 7,9; 8,7; 9,2; 10,5; 14; 7,2; 5,3; 9
B_2	3,8; 1,5; 12,9; 6,9; 4; 8; 11,2; 12,4; 4,9; 8,9	12; 11,5; 8,9; 4,4; 9,8; 6,9; 7,2; 6,2; 10,5; 9,2	14,5; 4,3; 6,7; 12,4; 13,2; 8,4; 7,9; 15,2; 3,2; 5,5

19. Досліджувався вплив факторів A і B на приріст урожаю з гектара кукурудзи, ц/га: фактор A – три види внесення азотних добрив (частка), %;

фактор B – два види глибини зрошення шляхом штучного поливу. Результати досліджень наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	6,2; 6,53; 6,82; 7,42; 6,55; 8,56; 9,49; 10,25; 9,64; 6,89	7,63; 8,53; 6,92; 9,73; 11,25; 7,33; 6,25; 10,11; 12,55; 8,93	8,35; 9,44; 8,44; 9,89; 10,99; 11,35; 15,21; 14,25; 11,6; 6,2
B_2	6,99; 8,49; 2,45; 13,4; 12,45; 6,85; 10,23; 9,51; 7,21; 11,92	9,47; 6,83; 6,74; 13,53; 15,41; 12,36; 8,79; 6,44; 12,35; 10,42	12,53; 14,5; 10,26; 8,96; 7,44; 6,72; 6,34; 14,39; 13,29; 11,95

20. Досліджувався вплив факторів A і B на норми витрат пального трактором на 100 га в умовних одиницях: фактор A – три види обробітку землі (A_1 – під злакові культури; A_2 – картопляні культури; A_3 – кормові трави); фактор B – пори року (B_1 – весна; B_2 – осінь). Результати досліджень наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	9,72; 10,55; 9,85; 12,44; 10,24; 11,24; 9,99; 11,95; 9,85; 10,95	9,92; 9,71; 9,99; 16,53; 14,91; 13,86; 12,44; 11,66; 10,31; 11,63	9,81; 10,64; 11,85; 12,44; 10,95; 13,44; 14,53; 9,84; 10,25; 10,96
B_2	11,44; 12,53; 15,64; 9,51; 9,73; 9,56; 10,41; 12,22; 9,34; 9,2	12,44; 10,64; 11,53; 9,83; 11,62; 9,71; 13,44; 16,51; 15,32; 14,95	9,96; 10,24; 16,44; 13,53; 14,83; 9,71; 9,51; 10,43; 11,22; 9,7

21. Досліджувався вплив факторів A і B на яйценосність курей у штуках за 5 місяців: фактор A – домішка до корму з моркви A_1; силосу A_2; біокорму A_3; фактор B – два режими температури в птахофермі. Результати досліджень наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	180; 195; 205; 184; 196; 220; 245; 192; 210; 230	184; 180; 256; 196; 260; 195; 210; 216; 233; 205	196; 210; 236; 186; 192; 210; 181; 244; 212; 212
B_2	189; 210; 219; 250; 180; 196; 244; 249; 260; 202	196; 199; 188; 199; 244; 236; 259; 212; 189; 204	213; 243; 189; 199; 244; 249; 254; 260; 195; 182

22. Досліджувався вплив факторів A і B на час, який витрачається на виготовлення однієї деталі в хвилинах: фактор A – розряд робітника (A_1 – четвертий; A_2 – п'ятий; A_3 – шостий); фактор B – два сорти (B_1 і B_2). Результати досліджень наведено в таблиці

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	2,2; 3,4; 2,9; 3,5; 4,2; 2,8; 3,5; 3,9; 3,6; 4,6	2,9; 2,5; 4,4; 4,2; 3,8; 3,9; 5,4; 6,4; 5,2; 6,8	2,5; 6,8; 4,6; 5,9; 6,2; 4,2; 2,4; 2,6; 3,4; 3,8
B_2	5,4; 4,4; 3,8; 4,6; 2,4; 3,5; 6,8; 4,2; 3,6; 5,8	2,9; 5,4; 6,3; 5,4; 3,8; 3,9; 2,5; 4,6; 3,7; 4,8	3,6; 2,2; 2,5; 2,9; 6,2; 4,7; 5,2; 3,2; 3,9; 4,2

Приклад. При рівні значущості $\alpha = 0,05$ визначити вплив факторів A і B та їх сумісний вплив на досліджувану ознаку. Результати експерименту наведено в таблиці 3.6:

Таблиця 3.6

Фактор B	Фактор A		
	A_1	A_2	A_3
B_1	3,6;3,8;4,1	2,9;3,1;3,0	2,7;2,5;2,9
B_2	4,2;4,0;4,1	3,3;2,9;3,2	3,7;3,5;3,6
B_3	3,8;3,5;3,6	3,6;3,7;3,5	3,2;3,0;3,4
B_4	3,4;3,2;3,2	3,4;3,6;3,5	3,6;3,8;3,7

Розв'язування

Розрахунки подамо в таблиці 3.7.

Обчислюється розрахункове значення критерію Фішера:

$$F_A^* = \frac{S_A^2}{S_\varepsilon^2} = 27,699; F_B^* = \frac{S_B^2}{S_\varepsilon^2} = 12,942; F_{AB}^* = \frac{S_{AB}^2}{S_\varepsilon^2} = 19,728.$$

При рівні значущості α = 0,05 та при відповідній кількість ступенів вільності k_1 = 2, k_2 = 3, k_3 = 6, k_4 = 24 за таблицею критерію Фішера визначаються критичні точки: $F_{kpA}(\alpha, k_1, k_4) = 3,4$, $F_{kpB}(\alpha, k_2, k_4) = 3,0$, $F_{kpAB}(\alpha, k_3, k_4) = 2,5$.

Порівнюючи розрахункові значення критерію Фішера та значення критичних точок, робимо висновки про вплив факторів A і B та їх сумісний вплив на досліджувану ознаку.

Таблиця 3.7

Рівень фактора B q = 4	Рівень фактора A p = 3						Середня величина за рядками	Загальна середня величина
	A_1		A_2		A_3			
		середня групова		середня групова		середня групова		
B_1	3,6; 3,9; 4,1	$\bar{x}_{11}$ = 3,83	2,9, 3,1; 3,0	$\bar{x}_{12}$ = 3	2,7, 2,5; 2,9	$\bar{x}_{13}$ = 2,7	$\bar{y}_1$ = 3,18	$\bar{x} = \dfrac{\sum_{i=1}^{n}\sum_{j=1}^{p}\sum_{k=1}^{q} x_{ijk}}{N} = \dfrac{\sum_{i=1}^{n}\sum_{j=1}^{p}\sum_{k=1}^{q} x_{ijk}}{npq} = 3{,}44$
B_2	4,2, 4,0; 4,1	… = 4,1	3,3, 2,9; 3,2	… = 3,13	3,7, 3,5; 3,6	… = 3,6	$\bar{y}_2$ = 3,61	
B_3	3,8, 3,5; 3,6	… = 3,63	3,6, 3,7; 3,5	… = 3,6	3,2, 3,0; 3,4	… = 3,2	$\bar{y}_3$ = 3,48	
B_4	3,4, 3,2; 3,2	$\bar{x}_4$ = 3,27	3,4, 3,6; 3,5	… = 3,5	3,6, 3,8; 3,7	$\bar{x}_{43}$ = 3,7	$\bar{y}_4$ = 3,49	
Середня величина за стовпцями	$\bar{z}$ = 3,71		$\bar{z}_2$ = 3,31		$\bar{z}_3$ = 3,3			

Джерело, що спонукає до розсіювання	Сума квадратів відхилень	Кількість ступенів вільності	Статистичні оцінки дисперсії (виправлені дисперсії)
Фактор A	$Q_A = nq\sum_{j=1}^{p}(\bar{z}_j - \bar{x})^2 = 1{,}307$	$k_1 = p$-1 = 2	$S_A^2 = \dfrac{Q_A}{p-1} = 0{,}65405$
Фактор B	$Q_B = np\sum_{k=1}^{q}(\bar{y}_k - \bar{x})^2 = 0{,}917$	$k_2 = q$-1 = 3	$S_B^2 = \dfrac{Q_B}{q-1} = 0{,}30561$
Одночасна дія факторів A і B	$Q_B = n\sum_{=}^{q}\sum_{k=}\left(\bar{x}_k - \bar{z} - \bar{y}_k + \bar{x}\right)^2 = 2{,}795$	k_3 = (p-1)(q-1) = 6	$S_B^2 = \dfrac{Q_B}{(p-1)(q-1)} = 0{,}4658$
Дія випадкової компоненти ε_{ijk}	$Q_\varepsilon = \sum_{i=1}^{q}\sum_{j=1}^{p}\sum_{k=1}^{n}(x_{ijk} - \bar{x}_{jk})^2 = 0{,}567$	$k_4 = N$-pq = 24	$S_\varepsilon^2 = \dfrac{Q_\varepsilon}{N - pq} = 0{,}0236$
Загальне відхилення	$Q_0 = \sum_{i=1}^{q}\sum_{j=1}^{p}\sum_{k=1}^{n}(x_{ijk} - \bar{x})^2 = 5{,}586$	$k = N$-1 = 35	$S_0^2 = \dfrac{Q_0}{N-1} = 0{,}1596$

3.3 Перевірка статистичної гіпотези

Визначити можливий закон розподілу випадкової величини *X*, користуючись критеріями Пірсона та Колмогорова, для рівня значущості 0,05 перевірити правильність висунутої нульової гіпотези.

1.

границі інтервалу		емпірична частота
x_i	x_{i+1}	n_i
0	0,48	47
0,48	0,96	28
0,96	1,44	9
1,44	1,92	4
1,92	2,40	5
2,40	2,88	3
2,88	3,36	4

2.

границі інтервалу		емпірична частота
x_i	x_{i+1}	n_i
0	0,54	45
0,54	1,08	21
1,08	1,62	18
1,62	2,16	3
2,16	2,70	4
2,70	3,24	4
3,24	3,78	5

3.

x_i	x_{i+1}	n_i
0	0,5	51
0,5	1,0	20
1,0	1,5	10
1,5	2,0	7
2,0	2,5	4
2,5	3,0	3
3,0	3,5	5

4.

x_i	x_{i+1}	n_i
0	0,42	48
0,42	0,84	29
0,84	1,26	9
1,26	1,68	4
1,68	2,10	4
2,10	2,52	3
2,52	2,94	3

5.

x_i	x_{i+1}	n_i
0	0,47	48
0,47	0,94	20
0,94	1,41	14
1,41	1,88	6
1,88	2,35	5
2,35	2,82	4
2,82	3,29	3

6.

x_i	x_{i+1}	n_i
0	3,58	52
3,58	7,16	24
7,16	10,74	10
10,74	14,32	4
14,32	17,9	3
17,9	21,48	3
21,48	25,06	4

7.

x_i	x_{i+1}	n_i
0	3,38	48
3,38	6,76	29
6,76	10,14	8
10,14	13,52	5
13,52	16,9	3
16,9	20,28	4
20,28	23,66	3

8.

x_i	x_{i+1}	n_i
0	0,31	47
0,31	0,62	27
0,62	0,93	9
0,93	1,24	6
1,24	1,55	4
1,55	1,86	4
1,86	2,17	3

9.

x_i	x_{i+1}	n_i
0	0,29	41
0,29	0,58	20
0,58	0,87	17
0,87	1,16	11
1,16	1,45	5
1,45	1,74	3
1,74	2,03	3

10.

x_i	x_{i+1}	n_i
0	0,53	43
0,53	1,06	21
1,06	1,59	14
1,59	2,12	10
2,12	2,65	5
2,65	3,18	3
3,18	3,71	4

11.

x_i	x_{i+1}	n_i
1,5	1,73	14
1,73	1,96	12
1,96	2,19	13
2,19	2,42	15
2,42	2,65	15
2,65	2,88	15
2,88	3,11	16

12.

x_i	x_{i+1}	n_i
2,5	2,82	13
2,82	3,14	15
3,14	3,46	15
3,46	3,78	17
3,78	4,10	18
4,10	4,42	12
4,42	4,74	10

13.

x_i	x_{i+1}	n_i
2,7	3,04	11
3,04	3,38	14
3,38	3,72	14
3,72	4,06	16
4,06	4,40	18
4,40	4,74	15
4,74	5,08	12

14.

x_i	x_{i+1}	n_i
2,9	3,26	12
3,26	3,62	15
3,62	3,98	15
3,98	4,34	17
4,34	4,70	17
4,70	5,06	14
5,06	5,42	10

15.

x_i	x_{i+1}	n_i
1,2	1,68	21
1,68	2,16	21
2,16	2,64	19
2,64	3,12	24
3,12	3,60	25
3,60	4,08	20
4,08	4,56	20

16.

x_i	x_{i+1}	n_i
1,4	1,86	20
1,86	2,32	20
2,32	2,78	25
2,78	3,24	21
3,24	3,70	21
3,70	4,16	18
4,16	4,62	25

17.

x_i	x_{i+1}	n_i
1,2	1,64	19
1,64	2,08	20
2,08	2,52	23
2,52	2,96	23
2,96	3,40	23
3,40	3,84	19
3,84	4,28	23

18.

x_i	x_{i+1}	n_i
1,0	1,42	18
1,42	1,84	18
1,84	2,26	21
2,26	2,68	25
2,68	3,10	25
3,10	3,52	23
3,52	3,94	20

19.

x_i	x_{i+1}	n_i
1,3	1,79	19
1,79	2,28	21
2,28	2,77	22
2,77	3,26	24
3,26	3,75	20
3,75	4,24	22
4,24	4,73	22

20.

x_i	x_{i+1}	n_i
1,7	2,13	20
2,13	2,56	20
2,56	2,99	26
2,99	3,42	23
3,42	3,85	22
3,85	4,28	20
4,28	4,71	19

21.

x_i	x_{i+1}	n_i
10,0	12,5	5
12,5	15,0	10
15,0	17,5	20
17,5	20,0	30
20,0	22,5	20
22,5	25,0	10
25,0	27,5	5

22.

x_i	x_{i+1}	n_i
9,0	11,3	7
11,3	13,6	12
13,6	15,9	22
15,9	18,2	24
18,2	20,5	20
20,5	22,8	10
22,8	25,1	5

23.

x_i	x_{i+1}	n_i
5,1	7,0	6
7,0	8,9	11
8,9	10,8	21
10,8	12,7	25
12,7	14,6	20
14,6	16,5	12
16,5	18,4	5

24.

x_i	x_{i+1}	n_i
7,0	9,1	7
9,1	11,2	10
11,2	13,3	20
13,3	15,4	24
15,4	17,5	22
17,5	19,6	12
19,6	21,7	5

Довідкова інформація

При перевірці закону розподілу випадкової величини X критеріями Пірсона або Колмогорова, зручно користуватися відповідними таблицями, в які заносяться всі проміжні результати.

Перевірка статистичної гіпотези критерієм Колмогорова

i	x_i	x_{i+1}	n_i	x_i^*	p_i	$n_i' = p_i N$	$F(x_i^*)$	$\widetilde{F}(x_i^*)$	$\left\|\widetilde{F}(x_i^*) - F(x_i^*)\right\|$
1	x_1	x_2	n_1	x_1^*	p_1	n_1'	$F(x_1^*)$	$\widetilde{F}(x_1^*)$	
2	x_2	x_3	n_2	x_2^*	p_2	n_2'	$F(x_2^*)$	$\widetilde{F}(x_2^*)$	
…	…	…	…	…	…	…	…	…	
k	x_k	x_{k+1}	n_k	x_k^*	p_k	n_k'	$F(x_k^*)$	$\widetilde{F}(x_k^*)$	
			$\sum_{i=1}^{k} n_i = N$						

Перевірка статистичної гіпотези критерієм Пірсона

Початкові дані				Результати розрахунків					
i	x_i	x_{i+1}	n_i	x^*	$\Phi(z_i)$	$\Phi(z_{i+1})$	p_i	$n_i' = p_i N$	$\chi_i^2 = \frac{(n_i - n_i')^2}{n_i'}$
1	x_1	x_2	n_1	x^*	$\Phi(-\infty)$	$\Phi(z_2)$	p_1	n_1'	χ_1^2
2	x_2	x_3	n_2	x^*	$\Phi(z_2)$	$\Phi(z_3)$	p_2	n_2'	χ_2^2
…	…	…	…	…	…	…	…	…	…
k	x_k	x_{k+1}	n_k	x_k^*	$\Phi(z_k)$	$\Phi(\infty)$	p_k	n_k'	χ_k^2
			$\sum_{i=1}^{k} n_i = N$						$\chi^2 = \sum_{i=1}^{k} \chi_i^2$

де k – кількість інтервалів;

x_k – випадкова величина;

x_i^* – середина інтервалу;

$z_i = \dfrac{x_i - \bar{x}}{S_x}$ – приведене значення випадкової величини X;

$\Phi(z_i)$ – нормована функція Лапласа;

$F(x_i^*)$ – теоретична функція розподілу;

$\widetilde{F}\left(x_m^*\right) = \dfrac{1}{N}\left[\sum_{i=1}^{m-1} n_i + 0{,}5 n_m\right]$ – статистична функція розподілу.

Послідовність перевірки статистичної гіпотези:

1. Будується гістограма. Після побудови гістограми висувається гіпотеза про ймовірний закон розподілу випадкової величини;

2. За емпіричними даними визначаються середньоарифметичне $\bar{x} = \dfrac{\sum_{i=1}^{k} n_i x_i^*}{N}$, оцінка дисперсії $S_X^2 = \dfrac{\sum_{i=1}^{k} n_i \left(x_i^* - \bar{x}\right)^2}{N-1} = \dfrac{\sum_{i=1}^{k} n_i \left(x_i^*\right)^2}{N-1} - \dfrac{\bar{x}^2 N}{N-1}$ та середньоквадратичне відхилення S_X;

3. За прогнозованим законом розподілу визначається теоретичне значення ймовірності знаходження випадкового числа в інтервалі за формулою $P(\alpha < x < \beta) = \int_{\alpha}^{\beta} f(x) dx = F(\beta) - F(\alpha)$.

Наприклад:

а) нормальний закон – $P(z_i < z < z_{i+1}) = \int_{z_i}^{z_{i+1}} f(z) dz = \Phi(z_{i+1}) - \Phi(z_i)$;

б) експоненціальний закон – $P(x_i < x < x_{i+1}) = \int_{x_i}^{x_{i+1}} f(x) dx = F(x_{i+1}) - F(x_i)$, параметр λ, визначається на підставі значення середньоарифметичного $\bar{x} = \dfrac{1}{\lambda}$;

в) рівномірний закон – $P(x_i < x < x_{i+1}) = \int\limits_{x_i}^{x_{i+1}} f(x)dx = F(x_{i+1}) - F(x_i)$, параметри a і b визначаються відповідно як $a = \overline{x} - \sqrt{3}S_x$, $b = \overline{x} + \sqrt{3}S_x$.

Примітка. Особливість визначення ймовірності знаходження випадкового числа для першого і останнього інтервалів:

а) нормальний закон:

для першого інтервалу $P(-\infty < z < z_2) = \Phi(z_2) - \Phi(-\infty) = \Phi(z_2) + 0{,}5$,

для останнього інтервалу $P(z_k < z < \infty) = \Phi(\infty) - \Phi(z_k) = 0{,}5 - \Phi(z_k)$;

б) експоненціальний закон:

для першого інтервалу $P(0 < x < x_2) = F(x_2) - F(0)$,

для останнього інтервалу $P(x_k < x < \infty) = F(\infty) - F(x_k)$;

в) рівномірний закон:

для першого інтервалу $P(a < x < x_2) = \dfrac{1}{b-a}(x_2 - a)$,

для останнього інтервалу $P(x_k < x < b) = \dfrac{1}{b-a}(b - x_k)$;

4. Теоретичне значення частоти попадання випадкового числа в інтервал, визначається за формулою $n'_i = p_i \sum\limits_{i=1}^{k} n_i = p_i N$, де N – загальна кількість випадкових величин (експериментальних точок);

5. Відхилення χ_i^2 визначається $\chi_i^2 = \dfrac{(n_i - n'_i)^2}{n'_i}$;

6. Критерій $\chi^2 = \sum\limits_{i=1}^{k} \chi_i^2$, порівнюється з теоретичним значенням $\chi^2_{кр}$ (критичним), яке визначається з відповідної таблиці критерію Пірсона при відповідній кількості ступенів вільності і заданій надійній ймовірності, або при заданому рівні значущості;

7. Визначається найбільше відхилення $D_q = \max\left| \widetilde{F}(x_i^*) - F(x_i^*) \right|$, і розраховується критерій $\lambda_q = D_q \sqrt{N}$, який порівнюється з теоретичним значенням $\chi^2_{кр}$ (критичним), яке визначається з відповідної таблиці критерію Колмогорова при заданому рівні значущості.

ЛІТЕРАТУРА

1. Бабак В. П. Основи теорії ймовірностей та математичної статистики : навч. посібник / Бабак В. П, Біленький А. Я., Приставка О. П. – К. : КВЩ, 2003. – 432с.

2. Бабак В. П. Теорія ймовірностей, випадкові процеси та математична статистика : підручник для студентів ВНЗ / Бабак В. П., Марченко Б. Г., Фриз М. Є.– К. : Техніка, 2004. – 288с.

3. Жлуктечко В. І. Теорія ймовірностей і математична статистика : навч.-метод. посібник. У 2-х ч. – Ч.І. Теорія ймовірностей / В. І. Жлуктечко, С. І. Наконечний – К. : КНЕУ, 2000. – 303с.

4. Жлуктечко В. І. Теорія ймовірностей і математична статистика : навч.-метод. посібник. У 2-х ч. – Ч.ІІ. Математична статистика / Жлуктечко В. І. , Наконечний С. І., Савіна С. С. – К. : КНЕУ, 2001. – 336с.

5. Захаров И. П. Теория неопределенности в измерениях: учебное пособие / И. П. Захаров, В. Д. Кукуш – Харьков: «Консум», 2002. – 256с.

6. Володарський Є. Т. Статистична обробка даних : навч. посібник / Є. Т. Володарський, Л. О. Кошева – К. : НАУ. 2008. – 308 с.

7. Вентцель Е. С. Теория вероятностей и ее инженерные приложения / Е. С. Вентцель, Л. А. Овчаров – М. : «Наука», 1988. – 480с.

8. Вентцель Е. С.Теория случайных процессов и ее инженерные приложения / Е. С. Вентцель, Л. А. Овчаров – М. : «Наука», 1991. – 384 с.

9. Сборник задач по теории вероятностей, математической статистике и теории случайных функций, под. ред. А. А. Свешникова – М. : «Наука», 1970. – 656 с.

10. Гмурман. В. Е. Руководство к решению задач по теории вероятностей и математической статистике : учебное пособие / Гмурман В. Е. – М. : Высшая школа, 1970. – 212 с.

Навчальне видання

Методичні вказівки
до проведення практичних занять
та виконання контрольних робіт
з дисципліни «Спецглави математики». Ч. 2
для студентів напряму підготовки
«Метрологія та інформаційно-вимірювальні технології» всіх форм навчання

Редактор В. Дружиніна
Коректор З. Поліщук

Укладачі: П. Кулаков
В. Присяжнюк

Оригінал-макет підготовлено В. Присяжнюком

Підписано до друку
Формат 29,7 × 42¼ . Папір офсетний.
Гарнітура Times New Roman
Друк різографічний. Ум. друк. арк.
Наклад пр. Зам. №

Вінницький національний технічний університет.
навчально-методичний відділ ВНТУ.
21021. м. Вінниця. Хмельницьке шосе. 95,
ВНТУ к. 2201.
Тел. (0432) 59-87-36.
Свідоцтво суб'єкта видавничої справи
серія ДК № 3516 від 01.07.2009 р.

Віддруковано у Вінницькому національному технічному університеті
в комп'ютерному інформаційно-видавничому центрі.
21021. м. Вінниця. Хмельницьке шосе. 95,
ВНТУ. ГНК. к. 114.
Тел.(0432) 59-85-32.
publish.vntu.edu.ua; email: kive.vntu@gmail.com.
Свідоцтво суб'єкта видавничої справи
серія ДК № 3516 від 01.07.2009 р.