МЕХАНІЧНИЙ ФАКУЛЬТЕТ

Кафедра “Експлуатація та ремонт рухомого складу”

ОЦІНКА І РОЗРАХУНОК ПОКАЗНИКІВ НАДІЙНОСТІ

МЕТОДИЧНІ ВКАЗІВКИ

до практичних занять
з дисципліни

«ОСНОВИ НАДІЙНОСТІ, ТЕХНІЧНОЇ ДІАГНОСТИКИ І МОНІТОРИНГУ ТЕХНІЧНОГО СТАНУ ТЯГОВОГО РУХОМОГО СКЛАДУ»

Харків 2009

Методичні вказівки розглянуто та рекомендовано до друку на засіданні кафедри «Експлуатація та ремонт рухомого складу» 24 вересня 2007 р., протокол № 28.

Рекомендуються для студентів спеціальності "Рухомий склад та спеціальна техніка залізничного транспорту" спеціалізації 7.100501.01 "Виробництво, експлуатація та ремонт локомотивів" денної та заочної форм навчання.

Укладачі:

професори Е. Д. Тартаковський,
А. П. Фалендиш,
доценти О. В. Устенко,
О. С. Крашенінін,
асист. С. В. Михалків

Рецензент

проф. В. Ф. Головко

ОЦІНКА І РОЗРАХУНОК ПОКАЗНИКІВ НАДІЙНОСТІ

МЕТОДИЧНІ ВКАЗІВКИ

до практичних занять
з дисципліни

«ОСНОВИ НАДІЙНОСТІ, ТЕХНІЧНОЇ ДІАГНОСТИКИ І МОНІТОРИНГУ ТЕХНІЧНОГО СТАНУ ТЯГОВОГО РУХОМОГО СКЛАДУ»

Відповідальний за випуск Михалків С.В.

Редактор Еткало О.О.

Підписано до друку 26.12.07 р.

Формат паперу 60х84 1/16 . Папір писальний.

Умовн.-друк.арк. 2,5. Обл.-вид.арк. 2,75.

Замовлення № Тираж 100. Ціна

Видавництво УкрДАЗТу, свідоцтво ДК № 2874 від. 12.06.2007 р.
Друкарня УкрДАЗТу,
61050, Харків - 50, пл. Фейєрбаха, 7

МІНІСТЕРСТВО ТРАНСПОРТУ ТА ЗВ’ЯЗКУ УКРАЇНИ

УКРАЇНСЬКА ДЕРЖАВНА АКАДЕМІЯ ЗАЛІЗНИЧНОГО ТРАНСПОРТУ

МЕХАНІЧНИЙ ФАКУЛЬТЕТ

Кафедра «Експлуатація та ремонт рухомого складу»

ОЦІНКА І РОЗРАХУНОК ПОКАЗНИКІВ НАДІЙНОСТІ

МЕТОДИЧНІ ВКАЗІВКИ

до практичних занять з дисципліни
«Основи надійності, технічної діагностики і моніторингу технічного стану тягового рухомого складу»

для студентів спеціальності
«Рухомий склад та спеціальна техніка залізничного транспорту
(Локомотиви)»
денної та заочної форм навчання

Харків 2008

Методичні вказівки розглянуті та рекомендовані до друку на засіданні кафедри «Експлуатація та ремонт рухомого складу» 2007 року, протокол № 25.

Рекомендовано для студентів денної та заочної форм навчання по спеціальності «Рухомий склад та спеціальна техніка залізничного транспорту» спеціалізації 7.100501.01 «Виробництво, експлуатація та ремонт локомотивів»

Укладачі:

доц. О. В. Устенко
доц. А. Г. Теслик
проф. А. П. Фалендиш
доц. О. С. Крашенінін
ас. С. В. Михалків

Рецензент

проф. В. Ф. Головко

ЗМІСТ

ВСТУП

У сучасних умовах особливої актуальності набувають пріоритетні дослідження щодо розроблення перспективного тягового рухомого складу, удосконалення системи обслуговування та ремонту локомотивів, впровадження засобів механізації та автоматизації, сучасних автоматизованих робочих місць (АРМ), які базуються на інформаційних технологіях. Це обумовлює особливі вимоги до інженерно-технічного персоналу, який повинен володіти знаннями у галузі розрахунку надійності і якості рухомого складу.

Дані методичні вказівки включають в себе загальні відомості з теорії ймовірності і математичної статистики, теорії надійності.

У першому розділі наведені методичні підходи щодо обробки результатів статистичних досліджень, приклади розрахунків і вихідні дані для розрахунків.

У другому розділі розглянута методика планування і реалізації багатофакторного експерименту.

Методичні вказівки призначені для студентів спеціальності «Рухомий склад і спеціальна техніка залізничного транспорту (Локомотиви)» всіх форм навчання.

1 ЗАГАЛЬНІ ВІДОМОСТІ З ТЕОРІЇ ЙМОВІРНОСТІ Й МАТЕМАТИЧНОЇ СТАТИСТИКИ

1.1 Поняття випадкової події та ймовірності

У теорії ймовірності розрізняють три види подій: достовірні, неможливі і випадкові. Достовірною називається подія, яка обов'язково відбудеться при здійсненні певних умов. Неможливою називається подія, яка напевно не відбудеться при здійсненні певних умов. Випадковою називають подію, яка може відбутися або не відбутися при здійсненні заданих умов.

Дія, у результаті якої настає або не настає випадкова подія, прийнято називати випробуванням. Якщо поява однієї випадкової події виключає появу інших у даному випробуванні, то такі події називають несумісними. Якщо жодна з подій не є більш можливою, то такі події називають рівноможливими. Єдиною можливою називають таку подію, поява якої є достовірною подією.

Випадкова подія може бути охарактеризована кількісно-імовірнісною подією. Її визначають відношенням кількості сприятливих результатів випробувань до загальної кількості випробувань [1].

Ймовірність події А визначається за формулою

$$p\{A\} = \frac{m}{N}, \tag{1.1}$$

де m – кількість сприятливих результатів;

N – загальна кількість випробувань.

З визначення ймовірності випливає, що ймовірність достовірної події дорівнює одиниці $p\{A\} = 1$, а ймовірність неможливої події дорівнює нулю $p\{A\} = 0$. Ймовірність випадкової події є позитивне число, що знаходиться між нулем і одиницею

$$0 \le p\{A\} \le 1.$$

Поняттям, досить близьким до поняття ймовірності, є відносна частота події, або частість.

Частість визначається за формулою

$$\omega(A) = \frac{\omega}{N}, \tag{1.2}$$

де ω – кількість появи події у випробуваннях.

Різниця між імовірністю та частістю полягає в тому, що для визначення частості необхідне проведення випробувань, а для визначення ймовірності вони не потрібні. Часто буває зручно користуватися статистичним визначенням імовірності. Тоді за ймовірність події приймають його частість.

1.2 Основні теореми і формули теорії ймовірностей

У теорії ймовірностей основними є теореми додавання і множення. Теорема додавання полягає в такому. Ймовірність появи одного з декількох попарно несумісних подій дорівнює сумі ймовірностей цих подій:

$$p\{A_1 + A_2 + \ldots + A_k\} = p\{A_1\} + \{A_2\} + \ldots + p\{A_k\}. \tag{1.3}$$

Протилежними називаються дві несумісні і єдино можливі події. Позначимо одну з цих подій A, а протилежну $\overline{A}$. Сума ймовірностей протилежних подій дорівнює одиниці

$$p\{A\} + p\{\overline{A}\} = 1. \tag{1.4}$$

Якщо в результаті випробувань повинна наступити одна із системи подій A_1, $A_2 \cdots$, A_k, то така система називається повною. Сума ймовірностей подій, що утворює повну систему, дорівнює одиниці.

Незалежними називаються події, у яких імовірність появи однієї з них не залежить від імовірності появи інших.

Теорема множення ймовірностей для незалежних подій визначає, що ймовірність спільної появи декількох подій, які є незалежними у сукупності, дорівнює добутку ймовірностей цих подій

$$p\{A_1 A_2 \ldots A_k\} = p\{A_1\} p\{A_2\} \ldots p\{A_k\}. \tag{1.5}$$

Позначимо ймовірність прямої події p, а протилежної – q. Можна показати, що існують події $A_1 A_2 \ldots A_k$, незалежні у сукупності, то ймовірність появи хоча б однієї з цих подій дорівнює

$$p\{A\} = 1 - q_1 q_2 \ldots q_k. \tag{1.6}$$

Якщо події $A_1 A_2 \ldots A_k$ мають однакову ймовірність, то ймовірність появи хоча б одного з цих подій дорівнює

$$p\{A\} = 1 - q^k. \quad (1.7)$$

Для аналізу залежних подій впроваджується поняття умовної ймовірності. Запис $p_A\{B\}$ означає умовну ймовірність події В у припущенні, що настала подія А.

Для декількох залежних подій імовірність їхньої спільної появи дорівнює добутку ймовірності одного з них на умовні ймовірності всіх інших; при цьому ймовірність кожної наступної події обчислюється в припущенні, що всі попередні події вже з'явилися

$$p\{A_1A_2 \dots A_k\} = p\{A_1\}p_{A_1}\{A_2\} + p_{A_1A_2}\{A_3\} + \dots + p_{A_1 \dots A_{k-1}}\{A_k\}. \quad (1.8)$$

Для двох спільних подій імовірність появи хоча б однієї з них дорівнює

$$p\{A + B\} = p\{A\} + p\{B\} - p\{AB\}. \quad (1.9)$$

Формула повної ймовірності має вигляд

$$p\{A\} = p\{B_1\}p_{B1}\{A\} + p\{B_2\}p_{B2}\{A\} + \dots + p\{B_k\}p_{Bk}\{A\}. \quad (1.10)$$

Важливе значення в теорії ймовірностей займає інтегральна теорема Лапласа: якщо ймовірність p настання події А в кожному випробуванні постійна й відмінна від нуля та одиниці, то ймовірність того, що подія А з'явиться у випробуваннях від k_1 до k_2 разів, приблизно дорівнює

$$p_n\{k_1k_2\} \approx \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{x^2}{2}} dx. \quad (1.11)$$

Значення інтеграла $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{x^2}{2}} dx$ наведені в довідкових таблицях. Функцію $\Phi(x)$ називають функцією Лапласа. За її допомогою можна визначити ймовірність того, що відхилення частості від постійної ймовірності за абсолютною величиною не перевищує заданого числа $\varepsilon > 0$. Ця ймовірність дорівнює

$$p\left\{\left|\frac{\omega}{N}-p\right|\le\varepsilon\right\}\approx 2\Phi(x), \qquad (1.12)$$

де

$$x=\varepsilon\sqrt{\frac{n}{pq}}.$$

1.3 Випадкові величини

Випадковою називається величина, яка у результаті випробувань може набувати того або іншого значення. Випадкові величини можуть бути дискретними і безперервними. Для опису випадкової величини треба знати ймовірності набування нею різних значень.

Законом розподілу випадкової величини називається залежність між можливими значеннями і відповідними ймовірностями.

Для дискретної випадкової величини закон розподілу може бути заданий у вигляді таблиці 1.1, перший рядок якої містить усі можливі значення випадкової величини, а другий – відповідні ймовірності.

Таблиця 1.1

x_1	x_2	x_3	• • •	x_n
p_1	p_2	p_3	• • •	p_n

Сума ймовірностей усіх можливих подій дорівнює одиниці

$$p_1+p_2+\ldots+p_n=1 \quad . \qquad (1.13)$$

Ряд розподілу можна зобразити графічно: по осі абсцис відкладають можливі значення випадкової величини, а по осі ординат – ймовірності цих значень. З'єднавши отримані точки, отримують багатокутник розподілу (рисунок 1.1).

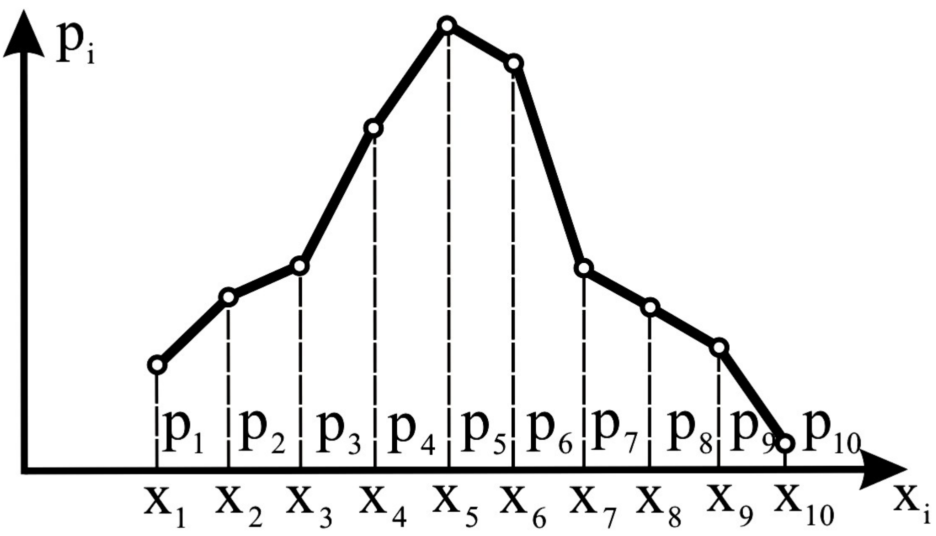

Рисунок 1.1 – Багатокутник розподілу

Для характеристики дискретної випадкової величини зручно користуватися числовими характеристиками. Однією з основних числових характеристик є математичне очікування M(x)

$$M(x) = \sum_{i=1}^{n} x_i p_i. \tag{1.14}$$

Математичне очікування приблизно дорівнює середньому арифметичному значенню випадкової величини $\overline{x}$ (тим точніше, чим більша кількість випробувань)

$$M(x) \approx \overline{x}. \tag{1.15}$$

Дисперсією називається математичне очікування квадрата відхилення випадкової величини від її математичного очікування

$$D(x) = M[x - M(x)]^2. \tag{1.16}$$

Для обчислення зручніше користуватися іншою формулою

$$D(x) = M(x^2) - [M(x)]^2. \tag{1.17}$$

Для характеристики розсіювання випадкової величини часто зручніше користуватися середнім квадратичним відхиленням, що являє собою корінь квадратний з дисперсії

$$\sigma(x) = \sqrt{D(x)}. \tag{1.18}$$

Середнє квадратичне відхилення має розмірність випадкової величини. Це не завжди зручно, наприклад при порівнянні розсіювання різних випадкових величин. У таких випадках зручніше користуватися показником розсіювання випадкової величини, вираженим у відносних величинах. Такою характеристикою є коефіцієнт варіації v

$$v = \frac{\sigma}{\bar{x}} \cdot 100\%. \tag{1.19}$$

Будь-які закони розподілу зручно характеризувати за допомогою моментів розподілу. Поняття моменту в статистиці аналогічно цьому поняттю в механіці. Початковим моментом порядку n випадкової величини називають математичне очікування величини x^n

$$\upsilon_n = M(x^n). \tag{1.20}$$

Використовуючи початкові моменти, формулу для дисперсії можна записати таким чином:

$$D(x) = \nu_2 - \nu_1^2. \tag{1.21}$$

Відхилення випадкової величини зручно характеризувати центральними моментами. Центральним моментом порядку n називається математичне очікування величини $[x - M(x)]^n$

$$\mu_n = M[(x - M(x)]^n. \tag{1.22}$$

Зокрема,

$$D(x) = \mu_2. \tag{1.23}$$

Початкові і центральні моменти пов'язані такими співвідношеннями

$$\left.\begin{aligned}\mu_2 &= \nu_2 - \nu_1^2;\\ \mu_3 &= \nu_3 - 3\nu_1\nu_2 + 2\nu_1^3;\\ \mu_4 &= \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4\end{aligned}\right\}. \tag{1.24}$$

Для кількісної оцінки безперервних і дискретних імовірних випадкових величин зручно користуватися не ймовірністю події $X = x$ (X –випадкова величина, а x – дійсне число), а ймовірністю події $X < x$. Ймовірність такої події є функція від x. Ця функція називається функцією розподілу випадкової величини X, або інтегральним законом розподілу; її позначають $F(x)$

$$F(x) = p\{X < x\}. \tag{1.25}$$

Значення інтегральної функції розташовані в межах

$$0 \le F(x) \le 1. \tag{1.26}$$

Інтегральна функція є тією, що не убуває. Ймовірність прийняття випадкової величини значення в інтервалі (a, b) дорівнює прирощенню інтегральної функції на цьому інтервалі

$$p\{a \le x < b\} = F(b) - F(a). \tag{1.27}$$

Для безперервної випадкової величини графік інтегральної функції часто має вигляд, зображений на рисунку 1.2.

Безперервну випадкову величину можна також задати, користуючись диференційною функцією розподілу ймовірностей $f(x)$, яку називають першою похідною від інтегральної функції

$$f(x) = F'(x). \tag{1.28}$$

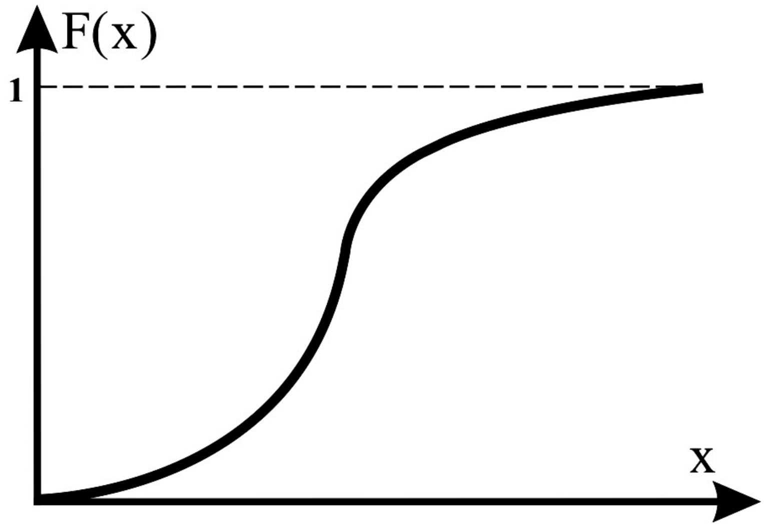

Рисунок 1.2 – Інтегральна функція безперервної випадкової величини

Знаючи диференційну функцію (щільність імовірності), можна обчислити ймовірність знаходження значення безперервної випадкової величини в заданому інтервалі від a до b

$$p\{a < x < b\} = \int_a^b f(x)dx. \tag{1.29}$$

Основні властивості диференційної функції

$$f(x) \geq 0, \tag{1.30}$$

$$\int_{-\infty}^{\infty} f(x)dx = 1. \tag{1.31}$$

На рисунку 1.3 зображена диференційна функція розподілу.

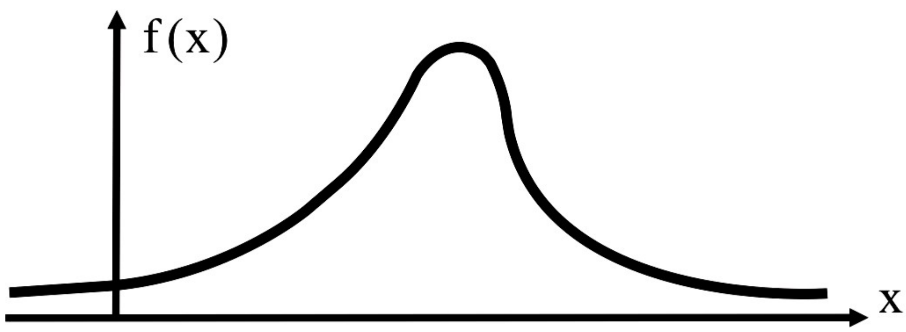

Рисунок 1.3 – Диференційна функція розподілу

1.4 Система випадкових величин

Декілька випадкових величин утворюють систему, у загальному випадку n-вимірну. Розглянемо детальніше систему із двох випадковий величин.

Законом розподілу дискретної двомірної випадкової величини називають перелік можливих значень цієї величини (x_i, y_i) та її ймовірності $p\{x_i, y_i\}$. Закон розподілу зручно задати у вигляді таблиці 1.2.

Таблиця 1.2

Y \ X	x_1	x_2	....	x_i	x_n	
y_1	$p\{x_1, y_1\}$	....	....	....	$p\{x_n, y_1\}$	
....	....	....	....	....	....	....
y_j	$p\{x_1, y_j\}$	....	....	$p\{x_i, y_j\}$	$p\{x_n, y_j\}$	
....	....	....	....	....	....	
y_m	$p\{x_1, y_m\}$	....	....	$p\{x_i, y_m\}$	$p\{x_n, y_m\}$	

У першому рядку таблиці перераховані усі можливі значення X, а в першому стовпчику – значення Y. В клітинці на перетині значень x_i й y_i вказана ймовірність спільної появи цих значень двовимірної випадкової величини. Сума ймовірностей, розташованих у всіх клітинках таблиці, дорівнює одиниці.

Аналогічно одновимірній випадковій величині інтегральною функцією розподілу двовимірної випадкової величини X, Y називають функцію $F(x, y)$.

Ймовірність потрапляння випадкової точки в напівсмугу $x_1 < X < x_2$ та $Y < y$ або $X < x$ й $y_1 < Y < y_2$ дорівнює прирощенню інтегральної функції по одному з аргументів

$$p\{x_1 < X < x_2, Y < y\} = F(x_2, y) - F(x_1, y). \tag{1.32}$$

Для безперервної двовимірної випадкової величини можна користуватися диференційною функцією розподілу

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \, \partial y}. \quad (1.33)$$

Геометрично диференційна функція являє собою поверхню розподілу.

Випадкові величини можуть бути залежними й незалежними. Дві випадкові величини називаються незалежними, якщо інтегральна функція системи цих величин дорівнює добутку інтегральних функцій складових або диференційна функція системи дорівнює добутку диференційних функцій складових [2].

Для системи двох випадкових величин, окрім відомих уже характеристик – математичних очікувань і дисперсій складових, досить істотні ще дві характеристики – кореляційний момент і коефіцієнт кореляції. Кореляційним моментом називають математичне очікування відхилень випадкових величин X та Y

$$\mu_{xy} = M[(X - M(x_i))(Y - M(y_j))]. \quad (1.34)$$

Для дискретних випадкових величин

$$\mu_{xy} = \sum_{i=1}^{n}\sum_{j=1}^{m}[x_i - M(x)][y_j - M(y)]p\{x_i, y_i\}. \quad (1.35)$$

Для безперервних випадкових величин

$$\mu_{xy} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[x - M(x)][y - M(y)]\, f(x, y)dx\, dy. \quad (1.36)$$

Кореляційний момент служить для характеристики зв'язку між величинами X та Y. Якщо випадкові величини незалежні, то кореляційний момент дорівнює нулю.

Зв'язок між величинами X та Y може бути виражений відносною числовою характеристикою – коефіцієнтом кореляції

$$r_{xy} = \frac{\mu_{xy}}{\sigma_x \sigma_y}. \quad (1.37)$$

Якщо коефіцієнт кореляції дорівнює нулю, то дві випадкові величини називаються некорельованими. Коефіцієнт кореляції характеризує не будь-яку залежність, а тільки лінійну. Для нелінійної залежності застосовується інший показник статистичного зв'язку – кореляційне відношення.

1.5. Теорія вибірок

Процес дослідження є вибірковим, тобто із усієї сукупності, яку називають генеральною, проводять випадковий відбір – вибірку. Кількість об’єктів генеральної або вибіркової сукупності називають об’ємом сукупності [3].

Вибірки можуть бути повторними або безповторними. У першому випадку відібраний об’єкт повертається до генеральної сукупності, у другому – не повертається. На практиці у більшості випадків застосовується безповторна вибірка. Вибірка повинна бути репрезентативною (представницькою); для цього кожний об'єкт повинен бути обраний випадково. Результати вибірки записують у вигляді таблиці 1.3.

Таблиця 1.3

X	x_1	x_2	...	x_k
ω	ω_1	ω_2	...	ω_k

У першому рядку таблиці вказані значення досліджуваної ознаки (у порядку їхнього зростання), у другому – відносні частоти.

Сума відносних частот статистичного розподілу дорівнює одиниці.

Аналогічно інтегральній функції генеральної сукупності існує поняття емпіричної функції розподілу.

Для графічного зображення статистичного розподілу використовують полігон і гістограму.

Площа гістограми відносних частот дорівнює одиниці.

Для вибірки можуть бути розраховані параметри закону розподілу. Ці значення будуть відрізнятися від значень параметрів у генеральній сукупності, однак вони дають уявлення про величину останніх. Таке наближене значення випадкового параметра генеральної сукупності називається оцінкою параметра.

Для того, щоб статистичні оцінки давали найкраще наближення параметрів, що оцінюються, вони повинні задовольняти такі вимоги.

При збільшенні кількості дослідів оцінка повинна сходитися за ймовірністю до шуканого параметра.

Необхідно, щоб оцінка не давала систематичної помилки в бік завищення або заниження. Оцінка, що задовольняє таку умову, називається незміщеною. Бажано, щоб обрана оцінка мала у порівнянні з іншими найменшу дисперсію. Оцінка, що має таку властивість, називається ефективною.

За оцінку математичного очікування або середню арифметичної генеральної сукупності (генеральної середньої) приймається вибіркова середня.

Можна показати, що вибіркова дисперсія є зміщеною оцінкою генеральної дисперсії і тому не може бути рекомендована як оцінка. Для одержання незміщеної оцінки генеральної дисперсії слід користуватися виправленою дисперсією вибірки s^2

$$s^2 = \frac{n}{n-1} D_в = \frac{\sum_{i=1}^{n}(x_i - \overline{x}_в)^2}{n-1}, \tag{1.38}$$

де n – об'єм вибірки;

$D_в$ – дисперсія вибірки;

$\overline{x}_в$ – середнє арифметичне вибірки.

Для оцінки середнього квадратичного відхилення генеральної сукупності слід користуватися виправленим середнім квадратичним відхиленням

$$s=\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x}_в)^2}{n-1}}. \tag{1.39}$$

Для оцінки точності оцінки параметрів користуються довірчими інтервалами та довірчими ймовірностями.

Якщо параметр N оцінений вибірковим параметром $N_в$, то чим менша різниця $|N-N_в|=\varepsilon$, тим точніша оцінка. Таким чином, число ε дозволяє визначити довірчий інтервал для параметра N. Проте довірчий інтервал може бути знайдений тільки з певною ймовірністю і залежить від неї. Ця ймовірність β називається довірчою.

Тоді

$$p\{N_в-\varepsilon<N<N_в+\varepsilon\}=\beta. \tag{1.40}$$

Ця рівність означає, що з довірчою ймовірністю β значення параметра N перебуває в довірчому інтервалі $(N_в-\varepsilon,\ N_в+\varepsilon)$.

Довірчі інтервали для математичного очікування нормального розподілу $\bar{x}$ при відомому середньому квадратичному відхиленні σ з імовірністю β можуть бути визначені із співвідношення

$$p\left\{\bar{x}-t\frac{\sigma}{\sqrt{n}}<\bar{x}<\bar{x}+t\frac{\sigma}{\sqrt{n}}\right\}=2\Phi(t)=\beta., \tag{1.41}$$

де $t=\frac{\varepsilon\sqrt{n}}{\sigma}$ – визначається з рівності

$$2\Phi(t)=\beta,\ \text{або}\ \Phi(t)=\frac{\beta}{2},$$

де $\Phi(t)$ – функція Лапласа (табульована).

Довірчі інтервали для математичного очікування нормального розподілу при невідомому середньому квадратичному відхиленні з імовірністю β можуть бути визначені із співвідношення

$$p\left\{\bar{x}_в-t_\beta\frac{s}{\sqrt{n}}<\bar{x}<\bar{x}_в+t_\beta\frac{s}{\sqrt{n}}\right\}=\beta, \tag{1.42}$$

де t_β – параметр розподілу Стьюдента.

Розподіл Стьюдента залежить тільки від об'єму вибірки і не залежить від x та σ.

Існують готові таблиці, за якими для заданих t_β та n можна знайти β, а за заданими β та n можна знайти t_β.

Довірчі інтервали для середнього квадратичного відхилення σ з імовірністю β можуть бути визначені із співвідношення

$$s(1-q) < \sigma < s(1+q), \tag{1.43}$$

де s – виправлене вибіркове середнє квадратичне відхилення;
$q = \varepsilon / s$ – може бути знайдене з таблиць за заданими n та β.

1.6. Критерій розходження

У математичній статистиці та у теорії надійності часто виникає необхідність оцінити істотність розбіжності між двома розподілами або їхніми статистичними параметрами.

Виникають задачі таких трьох типів:

1 Чи є істотним розходження даного емпіричного розподілу від того або іншого теоретичного?

2 Чи є дані емпіричні сукупності вибірками з однієї й тієї ж генеральної сукупності?

3 Чи відносяться отримані значення випадкової величини до даної статистичної сукупності?

У всіх перерахованих випадках задача може бути зведена до перевірки гіпотези про відсутність реального розходження. Цю гіпотезу називають нульовою. Позначимо її H_0. Правильність нульової гіпотези перевіряється у такий спосіб. Визначається ймовірність того, що внаслідок випадковості вибірки розходження може досягти фактичної зафіксованої величини. Якщо ця ймовірність дуже мала, то нульова гіпотеза відкидається.

Значення ймовірності, починаючи з якого ймовірність можна вважати малою, називають рівнем значущості. Прийнято вважати, якщо ймовірність нульової гіпотези $\alpha \le 1\%$, то вона відкидається, якщо $\alpha \ge 5\%$, то гіпотеза приймається, при $\alpha = 1 \div 5\%$ гіпотеза сумнівна [4].

Існують два основних типи критеріїв розходження – параметричні й непараметричні; перші необхідні для оцінки істотності розбіжності між числовими значеннями будь-якого параметра (наприклад, статистик розподілу), а другі – для оцінки частот розподілу.

Для порівняння середніх значень двох емпіричних сукупностей може бути застосований критерій Стьюдента. Для цього визначається величина t

$$t = \frac{\overline{x} - \overline{x}_г}{s_{\overline{x}}}, \tag{1.44}$$

де $\overline{x}$ – середнє вибірки;

$\overline{x}_г$ – середнє генеральної сукупності;

$s_{\overline{x}}$ – оцінка середнього квадратичного відхилення вибіркового середнього.

Отримане значення t порівнюється з критичним t_α. Для $\alpha = 1\%$ у довідкових таблицях наведено t_{01}, а для $\alpha = 5\% - t_{05}$. При $t < t_{05}$ нульова гіпотеза не відкидається.

Для оцінки істотності розходження двох вибіркових сукупностей за середніми квадратичними відхиленнями (або дисперсіями) користуються критерієм Фішера

$$F = \frac{s_1^2}{s_2^2}, \tag{1.45}$$

де s_1 й s_2 – оцінки середніх квадратичних відхилень.

Отримане значення F порівнюється з табличними $F_\alpha (F_{01}, F_{05})$. При $F < F_{05}$ нульова гіпотеза не відкидається.

Найпоширенішим непараметричним критерієм є критерій Пірсона (або, як його часто називають, критерій хі-квадрат). Він застосовується для перевірки гіпотези про вид розподілу. Це може бути гіпотеза про ступінь близькості емпіричного розподілу до обраного для опису теоретичного або гіпотеза про близькість двох емпіричних розподілів. В основі критерію Пірсона лежить порівняння частот розподілу. Обчислюється сума квадратів відхилення між частотами порівнюваних розподілів n'_i та n''_i

$$\chi^2 = \sum_i \frac{(n'_i - n''_i)^2}{n'_i - n''_i}. \quad (1.46)$$

При порівнянні емпіричного розподілу з теоретичним n''_i є частотами теоретичного розподілу.

Розраховані значення χ^2 порівнюються з табличними χ^2_{01} та χ^2_{05}. Як і в попередніх випадках, нульова гіпотеза відхиляється при $\chi^2 > \chi^2_{01}$ і приймається, якщо $\chi^2 \leq \chi^2_{05}$.

Окрім перерахованих, найпоширеніших критеріїв відомі ще деякі. Не вдаючись у подробиці, зазначимо лише області їхнього використання.

При вирішенні питання про розходження емпіричних сукупностей за характеристиками середнього положення застосовується критерій Вілкоксона.

Критерії розходження мають різну чутливість, або, як прийнято в математичній статистиці, різну потужність. Потужність критерію Вілкоксона незначна, тому якщо його застосування не дозволяє впевнено прийняти або відхилити нульову гіпотезу, то слід застосувати потужніший критерій для порівняння середніх значень X – критерій Ван дер Вардена. Для виявлення різниці між двома сукупностями не тільки за характеристиками середнього, але й за іншими властивостями застосовується серійний критерій, а якщо потужність його недостатня (при малих вибірках), то критерій Колмогорова-Смірнова.

При порівнянні сукупності з попарно спряженими частотами застосовується критерій знаків.

2 ОБРОБКА РЕЗУЛЬТАТІВ ДОСЛІДЖЕНЬ

Мета заняття – набування навичок статистичної обробки результатів вимірювань.

Завдання заняття:

– побудова статистичного ряду й гістограми;

– перевірка гіпотези про наявність грубої помилки в одному з результатів першого масиву;

– перевірка гіпотези про приналежність двох масивів результатів вимірювань до однієї генеральної сукупності й можливості їх об'єднання для подальшої обробки;

– перевірка достатності кількості вимірювань;

– визначення довірчого інтервалу для математичного очікування;

– визначення довірчого інтервалу для дисперсії.

Робота виконується самостійно в зазначеній послідовності завдань з використанням залежностей, наведених у конспекті лекцій і в даних методичних вказівках.

За варіантом завдання студент отримує два масиви результатів вимірювань: один з них – у вигляді декількох десятків чисел n_1, інший масив заданий кількістю вимірювань n_2, величинами емпіричного стандарту S_2 і середньоарифметичною $\overline{x}_2$.

Достовірність висновків задається величиною $1-\alpha=0{,}95$. Масиви чисел за варіантами індивідуальних завдань наведені в таблиці 2.1; результати попередньої обробки (статистичний ряд) заносять до таблиці 2.2. Квантилі статистик дані в таблицях 2.3 - і 2.6.

Таблиця 2.1 - Масиви вихідних даних

Варіант 1 $n_1=23$								
0,85	0,91	0,92	0,92	0,94	0,93	0,96	0,96	$n_2=23$
0,96	0,98	0,98	1,02	1,03	1,03	1,04	1,04	$S_2=0{,}063$
1,04	1,04	1,06	1,08	1,08	1,14	1,35		$\overline{x}_2=1{,}03$
Варіант 2 $n_1=23$								
1,0	1,23	1,27	1,29	1,29	1,31	1,32		$n_2=22$

1,34	1,34	1,34	1,36	1,37	1,37	1,37		$S_2 = 0{,}045$
1,38	1,38	1,39	1,41	1,44	1,44	1,50		$\overline{x}_2 = 1{,}36$
Варіант 3 $n_1 = 23$								
0,85	1,07	1,12	1,12	1,14	1,14	1,16	1,17	$n_2 = 22$
1,19	1,19	1,19	1,19	1,21	1,22	1,22	1,22	$S_2 = 0{,}07$
1,22	1,24	1,24	1,26	1,26	1,28	1,35		$\overline{x}_2 = 1{,}21$
Варіант 4 $n_1 = 23$								
0,75	0,97	1,01	1,01	1,04	1,04	1,07	1,07	$n_2 = 22$
1,07	1,08	1,08	1,08	1,11	1,13	1,13	1,13	$S_2 = 0{,}049$
1,14	1,14	1,14	1,17	1,17	1,17	1,25		$\overline{x}_2 = 1{,}095$
Варіант 5 $n_1 = 23$								
0,90	0,97	0,97	0,97	0,99	1,02	1,02	1,02	$n_2 = 22$
1,04	1,04	1,04	1,06	1,06	1,06	1,08	1,08	$S_2 = 0{,}05$
1,08	1,09	1,12	1,13	1,13	1,18	1,40		$\overline{x}_2 = 1{,}08$
Варіант 6 $n_1 = 23$								
0,70	0,92	0,96	0,96	0,99	0,99	1,01	1,01	$n_2 = 22$
1,03	1,03	1,03	1,03	1,06	1,06	1,06	1,08	$S_2 = 0{,}065$
1,08	1,08	1,08	1,13	1,13	1,14	1,2		$\overline{x}_2 = 1{,}04$
Варіант 7 $n_1 = 23$								
0,80	1,03	1,07	1,07	1,09	1,09	1,12	1,12	$n_2 = 22$
1,13	1,12	1,14	1,14	1,14	1,17	1,17	1,17	$S_2 = 0{,}049$
1,17	1,19	1,19	1,22	1,23	1,23	1,30		$\overline{x}_2 = 1{,}16$
Варіант 8 $n_1 = 23$								
0,95	1,02	1,03	1,03	1,04	1,06	1,06	1,07	$n_2 = 22$
1,07	1,09	1,09	1,11	1,13	1,13	1,13	1,13	$S_2 = 0{,}065$
1,14	1,14	1,16	1,18	1,18	1,23	1,45		$\overline{x}_2 = 1{,}12$

Продовження таблиці 2.1

Варіант 9 $n_1 = 23$									
1,0	1,06	1,06	1,09	1,09	1,12	1,12	1,12		$n_2 = 22$
1,14	1,14	1,14	1,16	1,17	1,17	1,17	1,19		$S_2 = 0{,}064$
1,19	1,19	1,23	1,23	1,23	1,29	1,50			$\overline{x}_2 = 1{,}135$
Варіант 10 $n_1 = 23$									
0,65	0,87	0,91	0,93	0,94	0,94	0,97	0,98	0,98	$n_2 = 22$
0,99	0,99	0,99	1,01	1,01	1,02	1,02	1,02	1,04	$S_2 = 0{,}061$
1,04	1,06	1,09	1,09	1,15					$\overline{x}_2 = 1{,}02$
Варіант 11 $n_1 = 23$									
1,95	2,01	2,03	2,03	2,03	2,04	2,06	2,06		$n_2 = 22$
2,07	2,07	2,09	2,09	2,11	2,12	2,12	2,12		$S_2 = 0{,}073$
2,14	2,14	2,16	2,19	2,19	2,24	2,45			$\overline{x}_2 = 2{,}09$
Варіант 12 $n_1 = 23$									

1,10	1,16	1,17	1,17	1,19	1,21	1,22	1,22	$n_2 = 22$
1.22	1,22	1,24	1,26	1,26	1,27	1,27	1,27	$S_2 = 0{,}07$
1,27	1,28	1,31	1,34	1,34	1,37	1,60		$\overline{x}_2 = 1{,}26$
Варіант 13 $n_1 = 23$								
1,85	1,97	1,97	1,98	1,98	1,98	1,98	1,91	$n_2 = 22$
1,92	1,93	1,93	2,01	2,02	2,02	2,02	2,04	$S_2 = 0{,}071$
2,04	2,04	2,06	2,07	2,07	2,14	2,35		$\overline{x}_2 = 1{,}99$
Варіант 14 $n_1 = 23$								
1,65	1,87	1,93	1,93	1,93	1,94	1,96	1,96	$n_2 = 22$
1,98	1,98	1,98	1,99	2,01	2,01	2,02	2,02	$S_2 = 0{,}072$
2.03	2,03	2,03	2,06	2,07	2,08	2,15		$\overline{x}_2 = 2{,}02$
Варіант 15 $n_1 = 23$								
1,05	1,11	1,12	1,13	1,13	1,16	1,16	1,17	$n_2 = 22$
1,17	1,18	1,18	1,21	1,22	1,22	1,22	1,22	$S_2 = 0{,}07$
1.23	1,23	1,26	1,29	1,29	1,33	1,55		$\overline{x}_2 = 1{,}17$
Варіант 16 $n_1 = 23$								
1,20	1,26	1,28	1,28	1,28	1,32	1,32	1,32	$n_2 = 22$
1,34	1,34	1,34	1,36	1,36	1,37	1,39	1,39	$S_2 = 0{,}046$
1,39	1,39	1,42	1,42	1,44	1,46	1,70		$\overline{x}_2 = 1{,}37$

Продовження таблиці 2.1

Варіант 17 $n_1 = 23$								
1,75	1,98	2,02	2,02	2,03	2,03	2,06	2,08	$n_2 = 22$
2,08	2,08	2,09	2,09	2,11	2,11	2,11	2,13	$S_2 = 0{,}074$
2,13	2,13	2,14	2,16	2,18	2,18	2,25		$\overline{x}_2 = 2{,}12$
Варіант 18 $n_1 = 23$								
0,90	1,13	1,17	1,17	1,17	1,18	1,22	1,22	$n_2 = 22$
1,22	1,24	1,24	1,24	1,26	1,26	1,26	1,27	$S_2 = 0{,}05$
1,27	1,29	1,29	1,32	1,33	1,34	1,40		$\overline{x}_2 = 1{,}27$
Варіант 19 $n_1 = 23$								
0,85	0,9	0,91	0,92	0,94	0,93	0,96	0,97	$n_2 = 22$
0,96	0,98	0,99	1,01	1,02	1,03	1,04	1,04	$S_2 = 0{,}065$
1,04	1,05	1,06	1,08	1,09	1,13	1,35		$\overline{x}_2 = 1{,}04$
Варіант 20 $n_1 = 23$								
1,0	1,22	1,22	1,27	1,29	1,29	1,32	1,32	$n_2 = 22$
1,33	1,34	1,34	1,35	1,36	1,36	1,37	1,37	$S_2 = 0{,}047$
1,38	1,38	1,38	1,41	1,42	1,44	1,5		$\overline{x}_2 = 1{,}34$
Варіант 21 $n_1 = 23$								
0,85	1,06	1,12	1,12	1,13	1,13	1,16	1,16	$n_2 = 22$
1,19	1,19	1,18	1,18	1,21	1,21	1,22	1,22	$S_2 = 0{,}071$
1,22	1,23	1,24	1,26	1,26	1,28	1,35		$\overline{x}_2 = 1{,}24$

Варіант 22 $n_1 = 23$								
0,75	0,97	1,01	1,02	1,04	1,04	1,06	1,06	$n_2 = 22$
1,07	1,07	1,07	1,08	1,11	1,13	1,13	1,13	$S_2 = 0{,}051$
1,14	1,14	1,14	1,16	1,16	1,17	1,25		$\overline{x}_2 = 1{,}093$
Варіант 23 $n_1 = 23$								
0,90	0,96	0,97	0,97	0,99	1,01	1,01	1,02	$n_2 = 22$
1,03	1,04	1,04	1,04	1,06	1,07	1,07	1,08	$S_2 = 0{,}052$
1,08	1,09	1,09	1,11	1,13	1,18	1,40		$\overline{x}_2 = 1{,}08$
Варіант 24 $n_1 = 23$								
0,90	1,10	1,16	1,17	1,17	1,18	1,22	1,22	$n_2 = 22$
1,22	1,24	1,24	1,24	1,25	1,26	1,26	1,27	$S_2 = 0{,}051$
1,29	1,29	1,29	1,32	1,32	1,34	1,40		$\overline{x}_2 = 1{,}28$

Продовження таблиці 2.1

Варіант 25 $n_1 = 23$								
1,75	1,99	2,02	2,02	2,03	2,04	2,06	2,06	$n_2 = 22$
2,08	2,08	2,09	2,09	2,11	2,12	2,12	2,13	$S_2 = 0{,}075$
2,13	2,14	2,14	2,16	2,18	2,18	2,25		$\overline{x}_2 = 2{,}13$

Приймаємо такі умовні позначення:

α – рівень значущості;
D_x – дисперсія випадкової величини X;
Δ – довжина інтервалу;
ε – точність оцінки m_x за $\overline{x}(\varepsilon = |\overline{x} - m_x|)$;
F – статистика $(F = S_б^2 / S_м^2)$;
F_α – квантиль F-статистики порядку α;
f_i – щільність частоти в і-му інтервалі;
G – статистика $(G = S_{max}^2 / \sum_1^N S_i^2)$;
G_α – квантиль статистики G порядку α;
k – кількість інтервалів, на яку ділиться масив при обробці;
μ_f – масштаб щільності частоти;
μ_x – масштаб X;
m_i – кількість значень в і-му інтервалі;
m_x – математичне очікування випадкової величини X;

n – кількість вимірювань;

υ – кількість ступенів свободи розподілу;

P – ймовірність;

p_i^* – частота потрапляння в i-й інтервал;

σ_x – середньоквадратичне відхилення випадкової величини X;

S – емпіричний стандарт;

S^2 – емпірична дисперсія;

$S_б^2$; $S_м^2$ – більше й менше значення двох порівнюваних емпіричних дисперсій;

t – статистика $(t = \frac{x_i - \overline{x}}{S}; \quad t = \frac{\overline{x} - m_x}{S/\sqrt{n}})$;

t_α – квантиль t-статистики порядку α;

V – статистика $(V = \frac{(n-1)S^2}{\sigma^2})$;

V_α – квантиль V-статистики порядку α;

$x_1, x_2, \ldots, x_n$ – можливі значення випадкової величини X;

$x^п$ – результат вимірювань, підозрюваний як груба помилка.

2.1 Методика обробки результатів досліджень

Будуємо статистичний ряд і гістограми інтервалів значень масиву від x_{min} до x_{max} із подальшим поділом на 10 рівних по довжині інтервалів (k=10). Довжина кожного інтервалу: $\Delta = (x_{max} - x_{min})/k$. Будується й заповнюється таблиця або статистичний ряд за формою, показаною в таблиці 2.2. Протягом підрахунку значень в кожному з інтервалів число, що лежить на правій границі, умовно відносять до попереднього інтервалу. Лише в перший інтервал включають числа як з правої, так і з лівої границі. Після заповнення таблиці виконуються такі перевірки:

$$\sum_1^k m_i = n_1; \quad \sum_1^k p_i^* = 1.$$

Таблиця 2.2 - Статистичний ряд, отриманий попередньою обробкою результатів вимірювань

Інтервали	x_{min} x_1	x_1 x_2	x_2 x_3	x_3 x_4	x_4 x_5	x_5 x_6	x_6 x_7	x_7 x_8	x_8 x_9	x_9 x_{max}
m_i – кількість значень в і-му інтервалі										
$p_i^* = m_i / n$ – частота потрапляння в і-й інтервал										
$f_i = kp_i^* /(x_{max} - x_{min})$ – щільність частоти в і-му інтервалі										

Обчислення величин p_i^* і f_i проводиться з точністю до 0,001. Для побудови гістограми (залежності f_i від x) обираються масштаби, наприклад, $\mu_f = 0{,}1$ од./мм; $\mu_x = 0{,}005$ од./мм. У цих масштабах по осі абсцис зображують відрізок ($x_{min} \dots x_{max}$) з поділом на інтервали, а по осі ординат –значення f_i. Інтервал $\Delta = 0{,}05$ од. x зобразиться в обраному масштабі відрізком довжиною $\Delta / \mu_x = 10$ мм. Значення щільності частоти $f_i = 0{,}1$ зобразиться відрізком $f_i / \mu_f = 1$ мм. На рисунку 2.1 зображена фігура, що складається з окремих прямокутників. Значення по осі абсцис є однаковими і з урахуванням масштабу μ_x дорівнює довжині інтервалу Δ, а висоти мають різне значення і дорівнюють з урахуванням масштабу μ_f щільності частоти для відповідних інтервалів значень x.

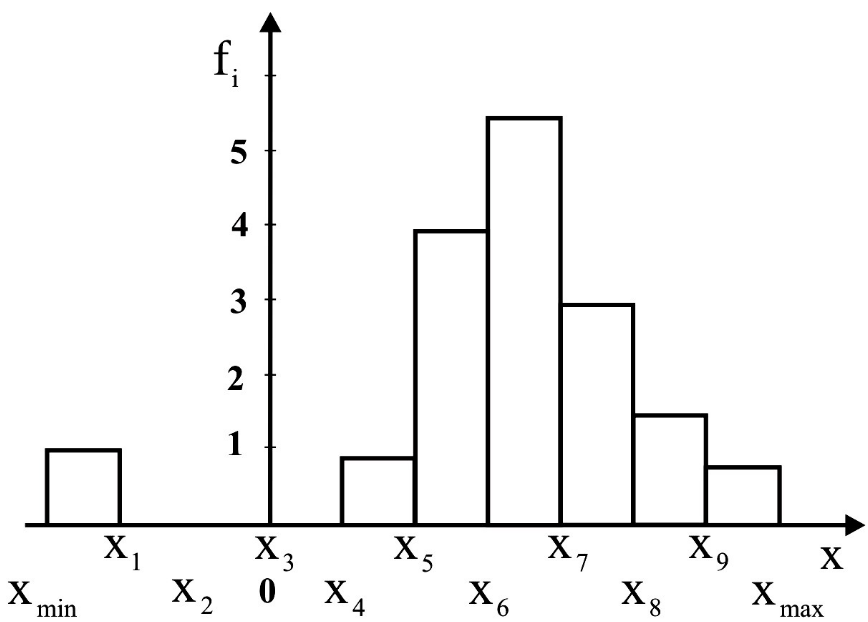

Рисунок 2.1 – Гістограма щільності розподілу

Для перевірки правильності побудови цієї фігури (гістограми) можна використовувати рівність

$$A\mu_x\mu_f = 1{,}0,$$

де А – площа фігури, обмежена залежністю $f_i = (x)$ та віссю х, мм^2.

2.2 Перевірка гіпотези про наявність грубої помилки в одному з результатів вимірювань

У заданому масиві одне із чисел випадає із загального ряду (це добре видно на гістограмі). Необхідно перевірити, чи є дане число результатом грубої помилки експерименту, чи воно є наслідком випадкової природи самого досліджуваного процесу. Під час розв’язання цієї й наступної задач передбачається, що досліджувана випадкова величина X узгоджується з нормальним законом розподілу.

Гіпотеза про наявність грубої помилки перевіряється з використанням розподілу t-статистики, що добре узгоджується з розподілом Стьюдента $f(t, v)$ при кількості ступенів свободи $v = n - 2$, де n – кількість вимірювань, не враховуючи того вимірювання, у якому отриманий підозрюваний на помилку результат.

За розподілом $f(t, v)$, що є симетричним щодо осі t, визначаються квантилі $t_{\alpha/2}$ та $t_{1-\alpha/2}$, що знаходяться на границі області правдоподібних значень при заданій достовірності висновку $1 - \alpha = 0{,}95.$ Якщо величина t знаходиться лівіше границі правдоподібних значень $(t < t_{\alpha/2})$ або правіше правої $(t > t_{1-\alpha/2})$, то таке значення t з достовірністю 0,95 не вважається приналежним до даної генеральної сукупності. Враховуючи симетричність розподілу $f(t, v)$, можна обмежитися порівнянням $|t|$ із квантилем $t_{1-\alpha/2}$.

Для практичного розв’язання задачі обчислюють значення

$$t^n = \frac{\left| x^n - \overline{x}_1 \right|}{S_1}, \qquad (2.1)$$

де $x^п$ – результат вимірювань, що знаходиться під підозрою грубої помилки;

$\overline{x}_1$, S_1 – середньоарифметичне значення і середнє квадратичне відхилення, обчислені без урахування підозрюваного результату.

Для обчислення $\overline{x}_1$ та S_1 використовують залежності

$$\overline{x}_1 = \frac{1}{n}\sum_1^n x_i, \tag{2.2}$$

$$S_1 = \sqrt{\frac{\sum_1^{n_1}(x_i - \overline{x}_1)^2}{n-1}}. \tag{2.3}$$

Тут n дорівнює числу дослідів мінус одиниця (дослід, що перебуває під підозрою, не враховується). За умови малої кількості чисел $(x_i - \overline{x})$ рекомендується для зручності обчислень за формулою (2.3) величину $\overline{x}_1$, а також всі числа x_i збільшити в 100 разів, а отримане значення S_1 поділити на 100.

Якщо $t^п > t_{1-\alpha/2}$ то результат $x^п$ з імовірністю $1-\alpha$ є наслідком грубої помилки експерименту і його варто відкинути. Якщо $t^п < t_{1-\alpha/2}$, то результат $x^п$ не слід відкидати.

Наприклад, отримано $\overline{x}_1 = 0{,}9991$ і $S_1 = 0{,}0546$, а за таблицями квантилів розподілу Стьюдента при $v = 22 - 2 = 20$ обчислене значення квантиля $t_{1-\alpha/2} = t_{0,975} = 2{,}09$. Перевірити, чи є значення $x^п = 0{,}65$ наслідком грубої помилки експерименту

$$t^п = \frac{|0{,}65 - 0{,}9991|}{0{,}0546} = 6{,}4.$$

За умови $t^п = 6{,}4 > t_{1-\alpha/2} = 2{,}09$, то результат вимірювання $x^п = 0{,}65$ з достовірністю висновку 0,95 слід вважати грубою помилкою.

2.3 Перевірка гіпотези про приналежність двох масивів до однієї генеральної сукупності

Це можна здійснити за допомогою використання розподілів Фішера або Кокрена.

Статистика $F = S_б^2 / S_м^2$ для однорідних вибірок добре узгоджується із розподілом Фішера із кількістю ступенів свободи: для чисельника $v_б = n_б - 1$, для знаменника $v_м = n_м - 1$. Тут $S_б^2$ та $S_м^2$ – відповідно більше й менше значення емпіричних дисперсій вибірок, що порівнюються; $n_б$ та $n_м$ –відповідні їм числа вимірювань. За таблицею квантилів розподілу Фішера визначаємо квантиль $F_{0,95}$. Якщо $S_б^2 / S_м^2 > F_{0,95}$, то з достовірністю 0,95 доцільно визнати вибірки неоднорідними. Якщо $S_б^2 / S_м^2 < F_{0,95}$, то вибірки визнаються однорідними, їх можна об'єднати в одну й подальшу обробку проводити для об'єднаної вибірки.

Наприклад, при $n_б = n_м = 22$ квантиль $F_{0,95} = 2{,}07$, а відношення $S_б^2 / S_м^2 = 0{,}061^2 / 0{,}0546^2 = 1{,}249$. Так як $S_б^2 / S_м^2 < F_{0,95}$, то з достовірністю 0,95 можна зробити висновок, що вибірки є однорідними і їх можна об'єднати в одну.

Статистика $G = S_{max}^2 / \sum_{1}^{N} S_i^2$ при однорідності вибірок узгоджується з розподілом Кокрена із числом ступенів свободи $v_1 = n - 1$ для чисельника та $v_2 = N$ для знаменника. Тут n – кількість результатів у масиві, що має найбільшу емпіричну дисперсію; N – кількість порівнюваних дисперсій (у нашому випадку $N = 2$).

За таблицями квантилів розподілу Кокрена при $v_1 = n - 1$ та $v_2 = 2$ знаходять квантиль $G_{0,95}$. Якщо $G > G_{0,95}$, то з достовірністю 0,95 вибірки визнають неоднорідними. Якщо $G < G_{0,95}$, то вибірки об'єднують і подальшу обробку ведуть для об'єднаного масиву.

У даній роботі при однорідності масивів хоча б за одним з критеріїв доцільно об'єднати вибірки й обчислити величини $\overline{x}$ та S^2 для об'єднаної вибірки

$$\overline{x} = \frac{n_1 \overline{x}_1 + n_2 \overline{x}_2}{n_1 + n_2}, \tag{2.4}$$

$$S^2 = \frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2}. \tag{2.5}$$

2.4 Перевірка достатності кількості вимірювань

Припустимо, що масиви визнані однорідними і для об'єднаної вибірки отримано

$$n = n_1 + n_2 = 44, \ \bar{x} = 1{,}009, \ S = 0{,}0575.$$

Розв’язується задача про достатність кількості вимірювань $n = 44$ для того, щоб з достовірністю 0,95 можна було сказати, що математичне очікування m (дійсне середнє) випадкової величини відхиляється від знайденого середньоарифметичного $\bar{x}$ не більше, ніж на задану малу величину $\pm\varepsilon$.

Ця задача розв’язується з використанням статистики $t = (\bar{x} - m)\sqrt{n}/S$, що добре узгоджується з розподілом Стьюдента із кількістю ступенів свободи $v = n - 2$, якщо самі результати вимірювань узгоджуються з нормальним законом розподілу.

Якщо за таблицями квантилів розподілу Стьюдента визначити квантиль $t_{1-\alpha/2}$, то необхіднц кількість вимірювань можна визначити за залежністю

$$n \ge \left(\frac{St_{1-\alpha/2}}{|\varepsilon|}\right)^2, \tag{2.6}$$

отриманою з виразу для визначення t-статистики.

Наприклад, $t_{0,975} = 2{,}02$ при $v = 42$. Тоді при $S = 0{,}0575$ та $|\varepsilon| = 0{,}01$ маємо $n \ge \left(\frac{0{,}0575 \cdot 2{,}02}{0{,}01}\right)^2 = 136$.

Звідси можна зробити висновок, що проведена кількість вимірювань $n = 44$ не забезпечує з достовірністю 0,95 бажаної точності оцінки m по $\bar{x}$, рівної 0,01. Якщо нам потрібна точність оцінки не більше, ніж 0,01, і надійність висновку не менше, ніж 0,95, то необхідно продовжити експеримент і провести ще $136 - 44 = 92$ досліди. Якщо нас влаштовує точність оцінки $\varepsilon = |\bar{x} - m_x| = 0{,}02,$ то необхідна кількість вимірювань становить

$$n \ge \left(\frac{0{,}0575 \cdot 2{,}02}{0{,}02}\right)^2 = 34$$

і експеримент можна вважати закінченим.

2.5 Визначення довірчого інтервалу для математичного очікування

З виразу (2.6) можна визначити точність оцінки m_x за $\overline{x}$ при виконаній кількості вимірювань n і заданої достовірності висновку

$$\varepsilon = \frac{t_{1-\alpha/2}S}{\sqrt{n}}. \tag{2.7}$$

Тоді інтервал значень, у який з достовірністю $1-\alpha$ потрапляє математичне очікування, складе

$$\left\{\overline{x} - St_{1-\alpha/2}/\sqrt{n};\quad \overline{x} + St_{1-\alpha/2}/\sqrt{n}\right\}. \tag{2.8}$$

Наприклад, при $n = 44$ отримуємо

$$\varepsilon = \frac{2{,}02 \cdot 0{,}0575}{\sqrt{44}} = 0{,}0175.$$

Довірчий інтервал для m_x при $\overline{x} = 0{,}9991$

$$\{0{,}9816;\quad 1{,}0166\}.$$

Імовірність того, що дійсне середнє досліджуваної випадкової величини не вийде за межі даного інтервалу, дорівнює 0,95 або $P(0{,}9816 < m_x < 1{,}0166) = 0{,}95.$

2.6 Визначення довірчого інтервалу для дисперсії

Необхідно знайти такий інтервал значень, всередині якого з ймовірністю 0,95 перебуває дисперсія досліджуваної випадкової величини.

Задача розв'язується з використанням статистики $V = (n-1)S^2/\sigma^2$, що узгоджується з χ^2-розподілом і кількістю ступенів свободи $v = n-1$.

За таблицями квантилів χ^2-розподілу знаходять значення квантилів $V_{1-\alpha/2}$ та $V_{\alpha/2}$. Потім з виразу статистики V визначають ліву й праву границі довірчого інтервалу для дисперсії σ^2

$$\left\{\frac{(n-1)S^2}{V_{1-\alpha/2}} < \sigma^2 < \frac{(n-1)S^2}{V_{\alpha/2}}\right\}. \qquad (2.9)$$

Наприклад, при $v = 44 - 1 = 43$ маємо

$$V_{1-\alpha/2} = V_{0,975} = 59{,}34,$$
$$V_{\alpha/2} = V_{0,025} = 24{,}43.$$

Тоді, використовуючи загальноприйняте позначення $\sigma^2 = D$, запишемо

$$P\left(\frac{(44-1)0{,}0575^2}{59{,}34} < D < \frac{(44-1)0{,}0575^2}{24{,}43}\right) = 0{,}95$$

або

$$P(0{,}0024 < D_x < 0{,}0058) = 0{,}95.$$

Інтервал $(0{,}0024;\ 0{,}0058)$ є в розглянутому випадку довірчим інтервалом для дисперсії.

Таблиця 2.3 – Квантилі $t_{v,\beta}$ розподілу Стьюдента (t-розподіл)

Кількість ступенів свободи	Ймовірність β, $\beta = 1 - \alpha$						
	0,70	0,80	0,90	0,95	0,975	0,99	0,995
5	0,56	0,92	1,48	2,01	2.57	3,36	4,03
10	0,54	0,88	1,37	1,81	2,23	2,76	3,17
15	0,54	0,87	1,34	1,75	2,13	2,60	2,95
19	0,53	0,86	1,33	1,73	2,09	2,53	2,84
20	0,53	0,86	1,32	1,72	2,09	2,53	2,84
21	0,53	0,86	1,32	1,72	2,08	2,52	2,83
22	0,53	0,86	1,32	1,72	2,07	2,51	2,82
23	0,53	0,86	1,32	1,71	2,07	2,50	2,81
40	0,53	0,85	1,30	1,68	2,02	2,42	2,70
100	0,53	0,84	1,29	1,66	1,98	2,36	2,63
500	0,52	0,84	1,28	1,65.1	1,96	2,33	2,59
∞	0,52	0,84	1,28	1,64	1,96	2,33	2,58

Таблиця 2.4 – Квантилі $F_{v_1,v_2,\beta}$ розподілу Фішера (F-розподіл), $F_{0,95}$, $P(F < F_{0,95}) = 0,95$

V_2	V_1, β=0.95, β=1-α						
	5	10	15	20	24	40	200
5	5,05	4,74	4,62	4,56	4,53	4,46	4,39
10	3,33	2,98	2,85	2,77	2,74	2,66	2,54
15	2,90	2,54	2,40	2,33	2,29	2,20	2,10
20	2,71	2,35	2,20	2,12	2,08	1,99	1,88
22	2,66	2,30	2,15	2,07	2,03	1,94	1,82
40	2,45	2,08	1,92	1,84	1,79	1,69	1,55
200	2,26	1,88	1,72	1,62	1,57	1,46	1,26

Таблиця 2.5 – Квантилі $G_{v_1,v_2,\beta}$ розподілу Кокрена,
$P(G < G_{0,95}) = 0,95, \quad v_1 = n - 1$

V_2	V_1, β=0.95, β=1-α				
	5	10	16	36	144
2	0,8772	0,7880	0,7341	0,6602	0,5813 0,5000
4	0,5895	0,4884	0,4366	0,3720	0,3093 0,2500
10	0,3029	0,2353	0,2032	0,1655	0,1308 0,1000

Таблиця 2.6 – Квантилі $\chi^2_{v,\beta}$ розподілу χ^2

Кількість ступенів свободи	Ймовірність β, $\beta = 1 - \alpha$				
	0,90	0,95	0,975	0,99	0,995
5	9,24	11,07	12,83	15,09	16,75
10	15,99	18,31	20,48	23,21	25,19
15	22,31	25,00	27,49	30,58	32,80
20	28,41	31,41	34,17	37,57	40,00
24	33,20	36,42	39,36	42,98	45,56
30	40,26	43,77	46,98	50,89	53,67
40	51,81	55,76	59,34	63,69	66,77
60	74,40	79,08	83,30	88,38	91,95
120	140,23	146,57	152,21	158,95	163,64

3 ПЛАНУВАННЯ Й РЕАЛІЗАЦІЯ БАГАТОФАКТОРНОГО ЕКСПЕРИМЕНТУ

Мета роботи – оволодіння алгоритмом планування багатофакторного експерименту за методом колового сходження (метод Бокса-Уілсона).

Завдання – визначити поєднання значень факторів, при яких параметр оптимізації досягає мінімуму [5].

Студенту задаються початкові рівні (таблиця 3.1) і пропонується знайти таке поєднання значень факторів, при якому значення функції досягає мінімуму.

Таблиця 3.1 – Натуральні значення початкових рівнів факторів

Варіант	$\tilde{x}_1^0$	$\tilde{x}_2^0$	$\tilde{x}_3^0$
1 – 10	15	15	35
11 – 20	20	10	10
21 – 30	10	25	10
Область існування факторів 0...40			

Об'єкт, що досліджується, може бути заданий моделлю виду

$$y = 3(x_1 - C_1)^2 + 1{,}5(x_2 - C_2)^2 + (x_3 - C_3)^2 + C_4,$$

де $C_1 = 0{,}7B,\ C_2 = 30 - 0{,}7B,\ C_3 = 30/B,\ C_4 = 0{,}5B + 5,\ B$ – номер варіанта.

Алгоритм планування багатофакторного експерименту дає переваги в трудомісткості при кількості факторів більше трьох. Застосування алгоритму в трифакторному експерименті здійснюється з метою спрощення.

Прийняті такі умовні позначення

b_j – коефіцієнт рівняння регресії при j-му факторі;
b_0 – вільний член лінійного рівняння регресії;
I_j – інтервал варіювання j-го фактора;
k – кількість факторів;

N – кількість дослідів для визначення коефіцієнтів рівняння регресії (кількість пробних дослідів);

x_j – кодоване значення j-го фактора;

$\tilde{x}_j^o$ – натуральне значення початкового рівня j-го фактора;

$\tilde{x}_j^в$, $\tilde{x}_j^н$ – натуральне значення верхнього й нижнього рівнів j-го фактора;

$\Delta\tilde{x}_j$ – натуральне значення кроку j-го фактора під час пошукового руху оптимуму;

+1, –1 – кодування значення верхніх і нижніх рівнів факторів;

y_i – значення параметра оптимізації в і-му досліді.

Метою експериментального дослідження технічного об'єкта (машини, механізму, апарата) є пошук поєднання значень факторів (розміри окремих елементів, швидкості, сили, тиску та ін.), при якому об'єкт функціонує якнайкраще. Показник ефективності функціонування об'єкта називається параметром оптимізації, і ним може бути собівартість одиниці продукції, продуктивність, енергоємність та ін.

Оскільки математична модель об'єкта дослідження невідома, знайти таке поєднання факторів можна тільки експериментально. При цьому, якщо кількість факторів є значною, то вирішити завдання методом послідовного перебору поєднань їхніх значень практично не можливо внаслідок великих витрат засобів і часу. Експеримент вдається провести з меншими витратами, якщо пошук вести за спеціальним алгоритмом, розробленим у теорії планування екстремального багатофакторного експерименту.

Спочатку проводять малу серію дослідів за матрицею дробового факторного експерименту (так звані пробні досліди) з варіюванням кожного фактора на двох рівнях і знаходять спрощену лінійну математичну модель об'єкта. Знаки й величини коефіцієнтів при факторах у цій моделі дозволяють визначити кроки й напрямки зміни кожного фактора під час руху до оптимуму. Змінюючи одночасно всі фактори на відповідні кроки, проводять досліди доти, доки поліпшиться (у нашому випадку зменшиться) значення параметра оптимізації. Досліди припиняють, як тільки параметр оптимізації почне погіршуватися. Значення факторів, при яких отримане найкраще значення, приймають за початкові рівні нової серії пробних дослідів. Ці досліди дозволять обчислити нові значення коефіцієнтів лінійної моделі та кроки подальшого руху для кожного з факторів.

Під час проведення трифакторного обчислювального експерименту для обчислення чотирьох коефіцієнтів моделі виду

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 \tag{3.1}$$

достатньо чотирьох пробних досліди за матрицею планування 2^{3-1}

x_1	x_2	x_3
+1	+1	+1
–1	+1	–1
+1	–1	–1
–1	–1	+1

Відповідно до цієї матриці перший дослід проводять, приймаючи натуральні значення усіх факторів на верхніх рівнях, другий дослід –приймаючи перший фактор на нижньому рівні, другий – на верхньому, третій – на нижньому і т.д.

Верхній рівень фактора

$$\tilde{x}_j^{в} = \tilde{x}_j^{0} + I_j. \tag{3.2}$$

Нижній рівень фактора

$$\tilde{x}_j^{н} = \tilde{x}_j^{0} - I_j. \tag{3.3}$$

Інтервали варіювання I_j для всіх факторів у даній роботі можна приймати однаковими й рівними, наприклад, у першій серії дослідів 1,0, у другій 0,5, у третій 0,25, у четвертій і наступних 0,1.

Провівши зазначені чотири досліди й отримавши чотири значення параметра оптимізації, обчислюють коефіцієнти лінійної моделі (рівняння регресії)

$$b_0 = \frac{1}{4}\sum_1^4 y_i; \quad b_1 = \frac{1}{4}\sum_1^4 y_i x_{1i}; \quad b_2 = \frac{1}{4}\sum_1^4 y_i x_{2i}; \quad b_3 = \frac{1}{4}\sum_1^4 y_i x_{3i}, \qquad (3.4)$$

де y_i – значення параметра оптимізації, отримане в і-му досліді;

x_{1i}, x_{2i}, x_{3i} – кодовані значення факторів (±1) в і-му досліді, і=1,2,3,4.

Коефіцієнти обчислюють із точністю трьох значущих цифр. Величини кроків зміни кожного з факторів для руху до оптимуму

$$\Delta\tilde{x}_1 = b_1 I_1; \quad \Delta\tilde{x}_2 = b_2 I_2; \quad \Delta\tilde{x}_3 = b_3 I_3. \qquad (3.5)$$

Якщо крок для одного з факторів є занадто значним (наприклад, більше 1...2), то кроки всіх факторів слід зменшити в однакову кількість разів (в 2,5,10…). Якщо кроки всіх факторів виявилися занадто малими (наприклад, менше 0,05), то рекомендується всі кроки збільшити в однакову кількість разів (в 2,5,10...). При виконанні даної роботи рекомендується кроки приймати в межах від 0,1 до 2,0, причому в кожному наступному циклі руху до оптимуму кроки, як правило, приймають меншими за величиною, ніж у попередньому циклі.

Якщо кроки факторів виявилися в рекомендованому діапазоні, а крок третього фактора є меншим 0,01, то значення цього фактора при русі у бік оптимуму можна залишати незмінним, а змінювати значення тільки перших двох факторів.

Послідовні значення факторів для руху до оптимуму обчислюють шляхом додавання або віднімання відповідного кроку до значення фактора, що прийнятий як початковий в даному циклі. Дія додавання або віднімання залежить від знака коефіцієнта при факторі та від того, до максимуму чи мінімуму параметра оптимізації ми прагнемо.

Необхідно обрати таке поєднання значень факторів, що мінімізує значення параметра оптимізації. Отже, якщо коефіцієнт при факторі позитивний, то цей фактор треба зменшувати (крок з початкового рівня послідовно віднімається), якщо від'ємний – збільшувати (крок до початкового рівня додається). Наприклад, при $b_1 < 0,\ b_2 < 0,\ b_3 > 0$ розрахунок значень факторів для руху до мінімуму заносять в таблицю 3.2.

Таблиця 3.2

Досліди руху до оптимуму	$\tilde{x}_1$	$\tilde{x}_2$	$\tilde{x}_3$
1	$\tilde{x}_1^0 + \Delta\tilde{x}_1$	$\tilde{x}_2^0 + \Delta\tilde{x}_2$	$\tilde{x}_3^0 - \Delta\tilde{x}_3$
2	$\tilde{x}_1^0 + 2\Delta\tilde{x}_1$	$\tilde{x}_2^0 + 2\Delta\tilde{x}_2$	$\tilde{x}_3^0 - 2\Delta\tilde{x}_3$
І т.д.			

Якщо під час руху до оптимуму один з факторів досяг границі області існування, то його величину слід зафіксувати на цьому значенні і далі змінювати тільки два інших фактори.

Якщо студент не припускається помилок у процедурі або обчисленнях, є достатнім провести 3 – 4 цикли, щоб з достатньою точністю (менше, дорівнює 0,5) визначити мінімальне значення параметра оптимізації й відповідне йому поєднання значень факторів. На необхідність припинення дослідів указує, наприклад, одержання в рівнянні регресії значень коефіцієнтів, близьких до нуля (наприклад, менше 0,01), або погіршення значення параметра оптимізації при будь-яких спробах зміни значень факторів [6].

По ходу виконання роботи кожним студентом послідовно заповнюється план експерименту, зразок якого наведений у таблиці 3.3.

Таблиця 3.2 – План проведення трифакторного екстремального експерименту

Перша серія дослідів				
Рівні	Натуральні значення			Пояснення
	$\tilde{x}_1$	$\tilde{x}_2$	$\tilde{x}_3$	
Основний $\tilde{x}_j^0$	5	10	20	
Інтервал варіювання I_j	1	0,5	1	
Верхній рівень $\tilde{x}_j^в$	6	11	21	
Нижній рівень $\tilde{x}_j^н$	4	9	19	

Досліди для пошуку коефіцієнтів лінійної моделі (пробні досліди)		Кодовані значення			Параметр оптимізації	Примітка
		x_1	x_2	x_3	y_i	
$N = 2^{3-1} = 4$ (напіврепліка плану 2^3)	1	+1	+1	+1		Для обчислення вводяться натуральні значення факторів
	2	–1	+1	–1		
	3	+1	–1	–1		
	4	–1	–1	+1		

$$b_j = \frac{1}{4}\sum_1^4 y_i x_{ji}; \quad b_1 = 1{,}02; \quad b_2 = -8{,}1; \quad b_3 = 1{,}63.$$

Крок $\Delta\tilde{x}_j = |b_j I|$; $\Delta\tilde{x}_1 = 1{,}02$; $\Delta\tilde{x}_2 = 4{,}05$; $\Delta\tilde{x}_3 = 1{,}63$

Кроки руху (після корегування й округлення) $\Delta\tilde{x}_1 = 0{,}5$; $\Delta\tilde{x}_2 = 2$; $\Delta\tilde{x}_3 = 0{,}8$

Досліди руху до оптимуму	$\tilde{x}_1$	$\tilde{x}_2$	$\tilde{x}_3$	y_i
5	$\tilde{x}_1^0 - \Delta\tilde{x}_1$	$\tilde{x}_2^0 + \Delta\tilde{x}_2$	$\tilde{x}_3^0 - \Delta\tilde{x}_3$	$y_5 = 12$
6	$\tilde{x}_1^0 - 2\Delta\tilde{x}_1$	$\tilde{x}_2^0 + 2\Delta\tilde{x}_2$	$\tilde{x}_3^0 - 2\Delta\tilde{x}_3$	$y_5 = 9$
.	. . . .	. . . .	. . . .	. . . .
10	2,0	22,0	15,2	$y_{10} = 5$ – кращий дослід серії
11	. . . .	. . . .	. . . .	$y_{11} = 5{,}5$ – припинення цієї серії дослідів

Закінчення таблиці 3.3

<table>
<tr><td colspan="6">Друга серія дослідів</td></tr>
<tr><td colspan="2"></td><td>$\tilde{x}_1$</td><td>$\tilde{x}_2$</td><td colspan="2">$\tilde{x}_3$</td></tr>
<tr><td colspan="2">$\tilde{x}_j^{0}$</td><td>2</td><td>22</td><td colspan="2">15,2</td></tr>
<tr><td colspan="2">I_j</td><td>0,5</td><td>0,4</td><td colspan="2">0,6</td></tr>
<tr><td colspan="2">$\tilde{x}_j^{в}$</td><td>2,5</td><td>22,4</td><td colspan="2">15,8</td></tr>
<tr><td colspan="2">$\tilde{x}_j^{н}$</td><td>1,5</td><td>21,6</td><td colspan="2">14,6</td></tr>
<tr><td rowspan="5">Пробні досліди</td><td></td><td>x_1</td><td>x_2</td><td>x_3</td><td>y_i</td></tr>
<tr><td>12</td><td>+1</td><td>+1</td><td>+1</td><td></td></tr>
<tr><td>13</td><td>–1</td><td>+1</td><td>–1</td><td></td></tr>
<tr><td>14</td><td>+1</td><td>–1</td><td>–1</td><td></td></tr>
<tr><td>15</td><td>–1</td><td>–1</td><td>+1</td><td></td></tr>
<tr><td colspan="6">b_j</td></tr>
<tr><td colspan="6">$\Delta\tilde{x}_j$</td></tr>
<tr><td colspan="2">Досліди руху до оптимуму</td><td>$\tilde{x}_1$</td><td>$\tilde{x}_2$</td><td>$\tilde{x}_3$</td><td>y_i</td></tr>
<tr><td colspan="2">16</td><td></td><td></td><td></td><td>4,90</td></tr>
<tr><td colspan="2">17</td><td></td><td></td><td></td><td>4,70</td></tr>
<tr><td colspan="2">. . . . .</td><td>. . . . .</td><td>. . . . .</td><td>. . . . .</td><td>. . . . .</td></tr>
<tr><td colspan="2">20</td><td>3,5</td><td>20</td><td>14,7</td><td>4,30 – кращий дослід серії</td></tr>
<tr><td colspan="2">21</td><td></td><td></td><td></td><td>4,33 – припинення цієї серії дослідів</td></tr>
</table>

Як звіт про виконану роботу студент наводить таблицю 3.3, заповнену в результаті планування й реалізації експерименту. Наприкінці таблиці формулюється висновок: у результаті проведених експериментів можна стверджувати, що параметр оптимізації досягає мінімуму, що дорівнює приблизно … при значеннях факторів $\tilde{x}_1 = ...,\ \tilde{x}_2 = ...,\ \tilde{x}_3 = ...$. Після цього викладач звіряє результати з дійсними значеннями, які відповідають досліджуваній функції.

СПИСОК ЛІТЕРАТУРИ

1 Биргер И.А. Прочность. Устойчивость. Колебания. – М.: Машиностроение, 1968. – Т. 1.
2 Калабро С.Р. Принципы и практические вопросы надежности. – М.: Машиностроение, 1966.
3 Румшиский Л.З. Математическая обработка результатов эксперимента. – М.: Наука, 1971.
4 Биргер И.А. Расчет на прочность деталей машин. – М.: Машиностроение, 1966.
5 Пустыльник Е.И. Статистические методы анализа и обработки наблюдений. – М.: Наука, 1968.
6 Свешников А.А. Прикладные методы теории случайных функций. – М.: Наука, 1968.
7 Решетов Д.Н. Работоспособность и надежность деталей машин. – М.: Высш. шк., 1974.
8 Смирнов Н.В., Дунин-Барковский И.В. Курс теории вероятностей и математической статистики для технических приложений. – М.: Наука, 1969. – 511 с.
9 Вентцель Е.С. Теория вероятностей. – М.: Физматгиз, 1962. – 654 с.
10 Шор Я.Б. Статистические методы анализа и контроля качества и надежности. – М.: Сов. радио, 1962. – 304 с.
11 Герцбах И.Б., Кордонский Х.Б. Модели отказов. – М.: Сов. радио, 1966. – 168 с.
12 Розенберг В.Я., Прохоров А.И. Что такое теория массового обслуживания. – М.: Сов. радио, 1965. – 256 с.