

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
ЦЕНТРАЛЬНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ  
УНІВЕРСИТЕТ

Факультет  
будівництва та транспорту

Кафедра електротехнічних  
систем та енергетичного менеджменту

# **МЕТОДИ ПЛАНУВАННЯ ТА ОРГАНІЗАЦІЇ НАУКОВИХ ДОСЛІДЖЕНЬ В ЕНЕРГЕТИЦІ**

*методичні рекомендації до виконання практичних робіт*

*для здобувачів вищої освіти другого (магістерського) рівня зі спеціальності 141*

*"Електроенергетика, електротехніка та електромеханіка"*

Кропивницький  
2022

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
ЦЕНТРАЛЬНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ  
УНІВЕРСИТЕТ

Факультет  
будівництва та транспорту

Кафедра електротехнічних  
систем та енергетичного менеджменту

# **МЕТОДИ ПЛАНУВАННЯ ТА ОРГАНІЗАЦІЇ НАУКОВИХ ДОСЛІДЖЕНЬ В ЕНЕРГЕТИЦІ**

*методичні рекомендації до виконання практичних робіт*

*для здобувачів вищої освіти другого (магістерського) рівня зі спеціальності 141*

*"Електроенергетика, електротехніка та електромеханіка"*

Затверджено  
на засіданні кафедри  
електротехнічних систем та  
енергетичного менеджменту  
Протокол № 4 від 19.09.2022 р.

Кропивницький  
2022

Методи планування та організації наукових досліджень в енергетиці: методичні рекомендації до виконання практичних робіт для здобувачів вищої освіти другого (магістерського) рівня зі спеціальності 141 "Електроенергетика, електротехніка та електромеханіка" / [уклад.: К.Г. Петрова, С.В. Серебренніков], Центральноукр. нац. техн. ун-т. – Кропивницький: ЦНТУ, 2022. – 52 с.

Укладачі: К.Г. Петрова – доц., к.т.н., доц. каф. ЕТС та ЕМ;  
С.В. Серебренніков – проф., к.т.н., доц. каф. ЕТС та ЕМ.

Рецензент: О.П. Голик – доцент, к.т.н., доц. каф. АВП ЦНТУ

© Петрова К.Г.,  
Серебренніков С.В., 2022  
© Центральноукраїнський  
національний технічний університет,  
2022

## ЗМІСТ

Вступ.....	5
Вимоги до виконання та оформлення звітів.....	6
Практична робота № 1. Формулювання і перевірка статистичних гіпотез...	7
Практична робота № 2. Перевірка гіпотези про рівність двох дисперсій.....	12
Практична робота № 3. Перевірка гіпотези про рівність математичних очікувань двох випадкових величин.....	14
Практична робота № 4. Перевірка гіпотези про закон розподілу випадкової величини.....	17
Практична робота № 5. Основи кореляційного та регресійного аналізу....	23
Практична робота № 6. Однофакторний дисперсійний аналіз.....	27
Практична робота № 7. Основи планування повного факторного експерименту.....	34
Практична робота № 8. Основи планування дробового факторного експерименту.....	40
Практична робота № 9. Використання методу парних порівнянь та аналізу ієрархій.....	44
Перелік рекомендованої літератури.....	50
Додаток А. Титульний аркуш звіту з практичної роботи.....	52

## ВСТУП

Під час підготовки здобувачів другого (магістерського) рівня вищої освіти важливе місце посідають освітні компоненти, які інтегрують навчальний процес з дослідницькою діяльністю. Одним з них є «Методи планування та організації наукових досліджень в енергетиці», який входить до циклу обов'язкових навчальних дисциплін.

Згідно з робочою програмою курсу «Методи планування та організації наукових досліджень в енергетиці» для здобувачів другого (магістерського) рівня вищої освіти спеціальності 141 «Електроенергетика, електротехніка та електромеханіка» (освітні програми "Електротехнічні системи електроспоживання" та "Енергетичний менеджмент") передбачено проведення практичних робіт, основна мета яких полягає у комплексному застосуванні набутих теоретичних знань.

У методичних рекомендаціях наведено дев'ять практичних робіт, кожна з яких має: теоретичну частину, детальне викладення типових прикладів розв'язування задач, що забезпечує теоретичну та методичну основу для подальшого самостійного вирішення поставлених завдань та безпосередньо завдання для самостійного виконання згідно варіанту.

Метою виконання практичних завдань є закріплення та поглиблення теоретичних знань, отриманих здобувачами вищої освіти на лекційних заняттях з дисципліни «Методи планування та організації наукових досліджень в енергетиці», а також набуття практичних навичок, зокрема, в таких питаннях як:

- формулювання та перевірки статистичних гіпотез;
- перевірка гіпотези про рівність двох дисперсій;
- перевірка гіпотези про рівність математичних очікувань двох випадкових величин;
- перевірка гіпотези про закон розподілу випадкової величини;
- кореляційний та регресійний аналізи;
- однофакторний дисперсійний аналіз;
- основи планування повного факторного експерименту;
- основи планування дробового факторного експерименту;
- використання методу парних порівнянь та аналізу ієрархій.

## ВИМОГИ ДО ВИКОНАННЯ ТА ОФОРМЛЕННЯ ЗВІТІВ

Зміст звітів з практичних робіт полягає в тому, що здобувачі вищої освіти на підставі наданих їм індивідуальних вихідних даних повинні виконати необхідні розрахунки практичних завдань.

Звіти з практичних робіт повинні мати:

- титульний аркуш;
- зміст;
- вступ (короткі теоретичні відомості);
- основну частину (завдання та його результати його виконання);
- висновки;
- список літератури (перелік посилань);
- додатки (за необхідності).

Приклад оформлення *титульного аркуша* наведено у додатку А.

У *вступі* доцільно навести короткі теоретичні відомості проблеми/питання, яке розглянуто у практичному завданні.

*Основна частина* містить виконання завдання згідно варіанту з необхідними розрахунками, поясненнями, графіками та обґрунтуванням прийнятих рішень.

У *висновках* повинні бути відображені підсумки виконання роботи.

*Вимоги до оформлення.* Звіт з практичної роботи друкують за допомогою ПК та принтера на одній сторінці аркуша білого паперу формату А4 (210х297мм). Обсяг роботи повинен становити до 10 сторінок.

Параметри друку звіту: міжрядковий інтервал - 1,0; шрифт – Times New Roman 13-14 пт; розмір абзацного відступу – 1,25-1,27 см; відступи: лівий – 25 мм, правий – 10 мм, верхній – 20 мм, нижній – 20 мм; кількість знаків у рядку – 70-80; рядків на сторінці – не більше 40.

## Практична робота № 1

### Тема. Формулювання і перевірка статистичних гіпотез

**Мета:** систематизація знань та навичок у галузі статистичного аналізу даних під час планування та організації експерименту, обчислення точкових та інтервальних оцінок розподілів, формулювання і перевірка статистичних гіпотез.

### Короткі теоретичні відомості

Математичну статистику можна визначити як науку про прийняття рішень в умовах невизначеності. Статистичні висновки належать до процесу отримання висновку щодо генеральної сукупності за властивостями вибірки з цієї сукупності. Зазвичай сукупність характеризується одним або кількома параметрами. Параметри генеральної сукупності будемо позначати грецькими літерами, наприклад,  $\mu$  – середнє значення сукупності,  $\sigma^2$  – дисперсія сукупності,  $\sigma$  – середнє квадратичне відхилення сукупності. Величини, обчислені за вибірковими значеннями, узятими з сукупності, називають оцінками, вибірковими статистиками або просто статистиками.

Математична статистика ґрунтується на припущенні, що вибірки, узяті з генеральної сукупності, є випадковими, тобто кожен елемент сукупності може з однаковою імовірністю виявитися уключеним у вибірку, і що видалення  $n$  елементів не впливає на модель мінливості сукупності.

Поняття *статистичного висновку* можна розділити на дві частини:

- 1) оцінювання параметрів генеральної сукупності;
- 2) перевірка статистичних гіпотез.

*Оцінювання.* Метою статистичного оцінювання є відшукування оцінки параметра генеральної сукупності на основі вибіркової статистики, узятої з цієї сукупності. Зазвичай потрібні два типи оцінок: *точкові оцінки* та *інтервальні оцінки*.

*Точкова оцінка* – це чисельне значення статистики, використовуване для оцінки параметра розподілу генеральної сукупності. Наприклад, вибіркове середнє  $\bar{x}$  – точкова оцінка математичного очікування генеральної сукупності  $\mu$ . Зазвичай передбачається, що точкові оцінки мають певні властивості. Оцінки мають відповідати таким вимогам:

- зі збільшенням обсягу вибірки оцінка має збігатися з оцінюваним параметром за ймовірністю;
- оцінка має бути ефективною, тобто за даного обсягу вибірки мати мінімальну дисперсію;
- оцінка має бути незміщеною, тобто математичне очікування оцінки повинне дорівнювати оцінюваному параметру.

Загальний метод знаходження оцінок з потрібними властивостями це метод максимуму правдоподібності. Практичне значення мають такі вибіркові характеристики, які відповідають вищепереліченим вимогам:

- 1) як оцінка математичного очікування – середнє арифметичне вибірки:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i;$$

- 2) як оцінка дисперсії випадкової величини  $x$  – значення, яке визначається формулою:

$$S_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1};$$

- 3) як оцінка дисперсії середнього арифметичного  $\bar{x}$  – значення, яке визначається формулою:

$$\bar{S}_x^2 = \frac{1}{n(n - 1)} \sum_{i=1}^n (x_i - \bar{x})^2;$$

- 4) як оцінки коефіцієнтів лінійних моделей – оцінки, отримані методом найменших квадратів;

- 5) як оцінка істинного значення величини  $x$  у серії з  $k$  вимірювань, кожне з яких виконано різними методами (приладами), що мають різні дисперсії, середнє зважене значення  $\tilde{x}$  визначається формулою:



$$\tilde{x} = \sum_{j=1}^k \frac{x_j}{S_j^2} \left( \sum_{j=1}^k \frac{1}{S_j^2} \right)^{-1},$$

де  $x_j$  – результат вимірювання  $j$ -м методом (приладом);  $S_j^2$  – оцінка дисперсії цього результату;

б) як оцінка дисперсії середньозваженого значення  $\tilde{x}$  – значення, яке визначається формулою:

$$\tilde{S}_x^2 = \frac{1}{n-1} \sum_{j=1}^k \frac{(x_j - \tilde{x})^2}{S_j^2} \left( \sum_{j=1}^k \frac{1}{S_j^2} \right)^{-1}.$$

*Інтервальна оцінка* полягає у визначенні інтервалу, який з заданою довірчою ймовірністю включає в себе оцінюваний параметр. Указати ймовірність попадання у деякий інтервал можна лише для випадкової величини з відомим законом розподілу.

*Перевірка статистичних гіпотез.* Загальний алгоритм статистичної перевірки гіпотез такий:

1. Формулюється гіпотеза  $H_0$ , яку необхідно перевірити.
2. Вибирається критерій перевірки –  $X$ . Критерій – це величина, закон розподілу якої за умови справедливості гіпотези  $H_0$  нам відомий.
3. Вибирається рівень значущості  $\alpha$  і критична область  $Q$  так, щоб умовна вірогідність попадання критерію в  $Q$  за умови справедливості гіпотези дорівнювала  $\alpha$ , тобто  $P(X \in Q | H_0) = \alpha$ .
4. Виконуємо експеримент і знаходимо експериментальне значення критерію  $X_c$ .
5. Якщо критерій не потрапляє у критичну область, гіпотеза приймається, якщо  $X \in Q$  – то відкидається.
6. Результат оформляється так: гіпотеза  $H_0$  перевірена умовою  $X$  на рівні значущості  $\alpha$  і прийнята (або відкинута).

### Завдання на практичну роботу

Використовуючи гральні кубики (або монету) як генератор випадкових чисел, організувати перевірку генератора на відповідність технічному завданню (ТЗ): на будь-якому етапі гри генератор має виробляти незалежно від стану всієї системи і своїх попередніх станів або «0» (певна цифра на грані кубика або одна з граней монети) з імовірністю  $p = 1/6$  ( $1/2$  для монети), або «1» (будь-яка інша цифра для кубика або інша грань монети) з імовірністю  $q = 5/6$  ( $1/2$  для монети). Перевірку генератора організувати як статистичну перевірку гіпотези.

*Пояснення.* Гіпотеза – це завжди твердження. Треба вибрати для перевірки одне з двох тверджень: «Відповідає ТЗ» або «Не відповідає ТЗ». Не кожен гіпотезу можна перевірити: якщо генератор відповідає ТЗ, то подія «поява 0» має відому постійну ймовірність, і результат кожного випробування не залежить від попередніх результатів. У цьому разі число  $m$  нулів у серії з  $n$  випробувань підпорядковується відомому закону Бернуллі, а за великого  $n$  величина  $X$ , згідно з теоремою Муавра-Лапласа, розподілена нормально з середнім 0 та дисперсією 1, тобто  $X$  має відомий розподіл і може бути критерієм перевірки гіпотези. Якщо ж сформулювати протилежну гіпотезу, то зовсім невідомо, як та у який бік ми відхилилися від ТЗ, і критерій побудувати не можна. Вибір  $\alpha$  не визначається правилами статистики, а цілком визначається обставинами експерименту. Чим більше в нашому прикладі  $\alpha$ , тим суворіше приймання роботи, тобто тим більше вірогідність того, що доведеться переробляти цілком хороший генератор. Зате, зменшуючи  $\alpha$ , ми збільшуємо небезпеку того, що визнаємо придатним бракований виріб. Прийнято вибирати  $\alpha = 0.05$  або  $\alpha = 0.1$ , що полегшує використання багатьох таблиць. Вибір критичної області має проводитися з урахуванням змісту гіпотези. Якщо  $X$  потрапляє у критичну область, то гіпотеза буде відкинута. У прикладі очевидно, що ТЗ буде порушено і у разі, якщо нулів під час 100 випробувань буде набагато менше, ніж очікувано, і якщо набагато більше, тому критична область – двостороння.

### **Програма роботи**

1. Ознайомитися з теоретичними положеннями.
2. Сформулювати і перевірити гіпотезу про відповідність генератора технічному завданню.
3. Зробити висновки та оформити звіт про виконану практичну роботу.

### **Порядок виконання роботи**

1. Вибрати та сформулювати гіпотезу  $H_0$ , яку необхідно перевірити ( $H_0$ : генератор відповідає ТЗ).
2. За вказівкою викладача, для певного варіанта, вибрати певну цифру на грані кубика або певну грань монети як стан «0».
3. Вибрати критерій перевірки  $X$ . Наприклад:

$$X = \frac{m - np}{\sqrt{npq}},$$

де  $n$  – число кидків кубика (монети) під час випробувань;  $m$  – число згенерованих станів «0».

4. Вибрати рівень значущості  $\alpha$  (імовірність забракувати хороший генератор, наприклад,  $\alpha = 0,05$ ) і за таблицями нормального розподілу знайти критичну область  $Q$  (наприклад,  $P\{|X| > 1,96\} = 0,05$ , отже,  $Q: |X| > 1,96$ ).

5. Виконати експеримент ( $n = 100$ ) і знайти експериментальне значення критерію  $X_e$ . Якщо критерій не потрапляє у критичну область, гіпотеза приймається, інакше – відхиляється.

### **Оброблення результатів. Зміст звіту**

1. Звіт має містити: номер, тему та мету практичної роботи, порядок виконання роботи з отриманими результатами та висновками.

### **Контрольні запитання**

1. Надайте визначення понять: «статистичний висновок», «точкова оцінка».
2. Поясніть суть та надайте алгоритм статистичної перевірки гіпотез.
3. У яких випадках неможливо перевірити сформульовану гіпотезу?
4. Які критерії вибору рівня значущості  $\alpha$ ?

## Практична робота № 2

### Тема Перевірка гіпотези про рівність двох дисперсій

**Мета:** систематизація знань та навичок в області статистичного аналізу даних під час планування та організації експерименту, формулювання і перевірки гіпотез.

### Короткі теоретичні відомості

*Перевірка гіпотези про рівність двох дисперсій. F-розподіл.* Нехай є дві незалежні вибірки значень нормально розподіленої величини  $x$ :  $x_1, x_2, \dots, x_n$  – усього  $n$  елементів, і нормально розподіленої величини  $y$ :  $y_1, y_2, \dots, y_m$  –  $m$  елементів. Гіпотеза  $H_0$  полягає в тому, що дисперсії величин  $X$  і  $Y$  рівні, тобто  $H_0: D[x] = D[y] = \sigma^2$ .

Ця гіпотеза перевіряється за критерієм Фішера. Випадкова величина

$$F = \frac{V_1/k_1}{V_2/k_2},$$

де  $V_1 \sim \chi_{k_1}^2, V_2 \sim \chi_{k_2}^2$  розподілені за законом, який отримав назву «розподіл Фішера», або «F-розподіл». У цього розподілу два параметри  $k_1$  і  $k_2$  (число ступенів свободи). Величина  $F$  приймає тільки додатні значення. Також, якщо відома вірогідність  $P$  того, що  $F > Fq$  (деякого фіксованого числа), то, очевидно, з такою ж вірогідністю  $1/F < 1/Fq$ , отже, з імовірністю  $\alpha = 2P$  величина  $F$  виходить за межі інтервалу  $(1/Fq, Fq)$ . Тому таблиці F-розподілу містять тільки межі  $Fq > 1$  із заданим  $\alpha$  і з певних  $k_1, k_2$  (якщо експериментальне значення критерію  $F$  виявиться менше за 1, то його треба «перевернути» і порівняти з табличним  $Fq$  зворотну величину). Використовуючи оцінку дисперсії випадкової величини  $S$  і підставивши в  $F$  як  $V_1$  комбінацію  $(n-1) \cdot S_x^2/D[x]$ , яка розподілена за законом  $\chi_{n-1}^2$ , а як  $V_2$  – величину  $(m-1) \cdot S_y^2/D[y]$ , яка розподілена за законом  $\chi_{m-1}^2$ , отримаємо, у разі рівності дисперсій  $D[x]$  і  $D[y]$ , відношення:

$$F = S_x^2/S_y^2,$$

яке підпорядковується розподілу Фішера і, отже, може слугувати критерієм перевірки гіпотези про рівність дисперсій.

## Завдання на практичну роботу

Запропоновано нову технологію виготовлення ламп. Вважається, що вона забезпечує менший технологічний розкид світловіддачі, ніж стара. Для перевірки виготовлено  $n = 9$  ламп за старою технологією, які мають значення світловіддачі (лм/Вт):  $x$ : 62, 73, 80, 79, 63, 77, 81, 75, і  $m = 10$  ламп за новою технологією зі світловіддачею:  $y$ : 75, 78, 77, 68, 73, 79, 72, 71, 86.

### Програма роботи

1. Ознайомитися з теоретичними положеннями.
2. Сформулювати і перевірити гіпотезу про рівність двох дисперсій.
3. Зробити висновки та оформити звіт про виконану практичну роботу.

### Порядок виконання роботи

1. Вибрати та сформулювати гіпотезу  $H_0$ , яку необхідно перевірити ( $H_0$ : технологічний розкид однаковий, тобто  $D[x] = D[y]$ ).
2. Вибрати критерій перевірки  $F$ .
3. Вибрати рівень значущості ( $\alpha = 0,05$ ).
4. За таблицею знайти  $F_q$ :
  - якщо виявиться, що  $S_x^2 > S_y^2, k_1 = 8, k_2 = 9$ , то  $F_q = 3,23$ ;
  - якщо  $S_x^2 < S_y^2, k_1 = 9, k_2 = 8$ , то  $F_q = 3,39$ .
5. Розрахувати:

$$S_x^2 = \frac{\sum_1^{k_1} (x_i - \bar{x})^2}{k_1 - 1},$$

$$S_y^2 = \frac{\sum_1^{k_2} (y_i - \bar{y})^2}{k_2 - 1},$$

$$F_e = \frac{S_x^2}{S_y^2}.$$

Зробити висновок. Якщо критерій не потрапляє у критичну область ( $F_e < F_q$ ), гіпотеза приймається, інакше – відхиляється. Якщо розробники все ж наполягають на перевагах нової технології, слід провести новий експеримент з вибірками великих обсягів.

### Оброблення результатів. Зміст звіту

1. Звіт має містити: номер, тему і мету лабораторної роботи, порядок виконання роботи з отриманими результатами та висновками.

### Контрольні запитання

1. Надайте визначення поняття «розподіл Фішера».
2. Поясніть суть та надайте алгоритм перевірки гіпотези про рівність двох дисперсій.
3. У яких випадках неможливо перевірити сформульовану гіпотезу?
4. Які критерії вибору рівня значущості  $\alpha$ ?

### Практична робота № 3

**Тема. Перевірка гіпотези про рівність математичних очікувань двох випадкових величин**

**Мета:** систематизація знань та навичок у галузі статистичного аналізу даних під час планування та організації експерименту, формулювання і перевірки гіпотез.

### Короткі теоретичні відомості

Нехай є дві незалежні вибірки значень нормально розподіленої величини  $x: x_1, x_2, \dots, x_n$  – всього  $n$  елементів, і нормально розподіленої величини  $y: y_1, y_2, \dots, y_m$  –  $m$  елементів.

Передбачається, що  $D[x] = D[y]$ . Гіпотеза  $H_0$  полягає в тому, що  $Mx = My$ . Це одна з найбільш поширених типів гіпотез у технологічних, біологічних і педагогічних експериментах. У обох вибірках існує однаковий розкид, але важливо визначити, чи значно відрізняються середні значення вибірок. Гіпотеза полягає в тому, що математичні очікування не відрізняються. Критерієм перевірки слугує випадкова величина  $t$ , розподілена за законом Стюдента:

$$t = \frac{z\sqrt{v}}{\sqrt{V}},$$

де  $z \sim N(0,1)$ ,  $V \sim \chi^2 v$ .

Візьмемо як  $z$  комбінацію:

$$z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_x^2}{n} - \frac{\sigma_y^2}{m}}}$$

Враховуючи, що  $x \sim N(M[x], D[x])$ ,  $y \sim N(M[y], D[y])$ , відповідно  $\bar{x} \sim N(M[x], D[x]/n)$ ,  $\bar{y} \sim N(M[y], D[y]/m)$ ,  $x$  та  $y$  незалежні, і тому дисперсія різниці їх середньоарифметичних дорівнює сумі дисперсій, а маточікування – різниці маточікувань. Ураховуючи рівність  $D[x]$  і  $D[y]$ , отримаємо, що  $z$  розподілено нормально з параметрами  $M[z] = 0$ ,  $D[z] = 1$ .

Як  $V$  візьмемо

$$V = \frac{(n-1)S_x^2}{\sigma^2} + \frac{(m-1)S_y^2}{\sigma^2} \sim \chi_{n+m-2}^2,$$

що є наслідком визначення  $\chi^2$ .

Як результат:

$$t = \frac{(\bar{x} - \bar{y})\sqrt{mn}\sqrt{m+n-2}}{\sqrt{m+n} \cdot \sqrt{(n-1)S_x^2 + (m-1)S_y^2}}.$$

Критична область – двостороння, тобто гіпотеза відхиляється, якщо  $|t| > t_q$ .

### Завдання на практичну роботу

Перевірити гіпотезу про рівність середніх значень світловіддачі ламп, виготовлених за старою та новою технологією. Для перевірки виготовлено  $n = 9$  ламп за старою технологією, які мають значення світловіддачі (лм/Вт): 62, 73, 80, 79, 63, 77, 81, 75, 77, і  $m = 10$  ламп за новою технологією зі світловіддачею:  $y$ : 75, 78, 77, 68, 73, 79, 72, 71, 86, 73.

### Програма роботи

1. Ознайомитися з теоретичними положеннями.
2. Сформулювати і перевірити гіпотезу про рівність середніх значень

світловіддачі ламп, виготовлених за старою та новою технологією.

3. Зробити висновки та оформити звіт про виконану практичну роботу.

### **Порядок виконання роботи**

1. Вибрати та сформулювати гіпотезу  $H_0$ , яку необхідно перевірити ( $H_0$ : середні значення світловіддачі ламп однакові, тобто  $Mx = My$ ).

2. Вибрати рівень значущості ( $\alpha = 0,05$ ).

3. За таблицею знайти, що при  $\nu = 9 + 10 - 2 = 17$ ,  $t_q = 2,1$ .

4. Розрахувати:

$$S_x^2 = \frac{\sum_1^{k_1} (x_i - \bar{x})^2}{k_1 - 1},$$

$$S_y^2 = \frac{\sum_1^{k_2} (y_i - \bar{y})^2}{k_2 - 1},$$

$$t = \frac{(\bar{x} - \bar{y})\sqrt{mn}\sqrt{m+n-2}}{\sqrt{m+n} \cdot \sqrt{(n-1)S_x^2 + (m-1)S_y^2}}.$$

5. Зробити висновок. Якщо критерій не потрапляє у критичну область ( $|t_e| < t_q$ ), гіпотеза приймається, інакше – відхиляється.

### **Оброблення результатів. Зміст звіту**

1. Звіт має містити: номер, тему та мету практичної роботи, порядок виконання роботи з отриманими результатами та висновками.

### **Контрольні запитання**

1. Надайте визначення поняття «розподіл Стюдента».

2. Поясніть суть та надайте алгоритм перевірки гіпотези про рівність математичних очікувань двох випадкових величин.

3. У яких випадках неможливо перевірити сформульовану гіпотезу?

4. Які критерії вибору рівня значущості  $\alpha$ ?



## Практична робота № 4

### Тема. Перевірка гіпотези про закон розподілу випадкової величини

**Мета:** систематизація знань та навичок у галузі статистичного аналізу даних під час планування та організації експерименту, формулювання і перевірки гіпотез.

### Короткі теоретичні відомості

Існує кілька способів перевірки справедливості припущень про закон розподілу випадкової величини. Найбільш поширений та наочний спосіб, у якому як критерій використовується розподіл  $\chi^2$ .

Припустимо, є вибірка  $n$  значень випадкової величини  $x$ .

Об'єм вибірки  $n$  має бути великий – не менше декількох десятків значень. Гіпотеза  $H_0$  полягає в тому, що випадкова величина  $x$  розподілена за деяким певним законом зі щільністю розподілу  $f(x)$ , наприклад, нормально або рівномірно, або як завгодно – безперервно або дискретно, але закон розподілу відомий.

Критерієм перевірки такої гіпотези може слугувати випадкова величина  $\chi^2$ . Закон розподілу величини  $\chi^2$  відомий.

Припустимо, у деяку систему вбудований генератор випадкових чисел, який має генерувати дійсні числа, рівномірно розподілені в інтервалі  $(0,1)$ . Потрібно перевірити, чи відповідає генератор цій вимозі. У табл. 4.1 наведена вибірка, що складається із 100 значень, вироблених генератором.

Таблиця 4.1 – Значення випадкової величини, вироблені генератором

.9995	.5045	.9103	.6868	.4160	.6515	.6885	.5269	.2470	.5298
.0727	.3283	.4865	.0346	.8278	.4777	.8310	.0077	.3654	.9911
.9826	.3911	.7233	.7601	.3541	.2008	.8847	.7012	.3834	.5297
.1665	.0501	.5328	.7665	.6316	.7254	.9478	.5133	.7702	.4121
.9138	.0535	.2693	.9047	.5194	.6793	.5890	.1316	.0475	.4587
.3593	.3842	.8683	.6296	.9304	.3190	.3510	.0159	.5373	.0920
.0908	.9092	.4940	.9347	.8977	.2771	.1254	.7622	.7362	.2662
.8462	.4364	.6789	.9867	.6712	.4154	.0471	.5007	.6539	.7361
.0738	.4175	.2332	.3063	.7556	.0606	.8415	.2190	.2625	.5163
.7227	.7534	.8460	.9410	.4645	.2727	.8886	.5911	.0669	.2370

Гіпотеза  $H_0$ : розподіл рівномірний в інтервалі  $(0,1)$ .

Жодне число з вибірки за межі цього інтервалу не виходить, тому вважаємо, що весь діапазон зміни  $x$  – від 0 до 1. Поділимо цей діапазон на інтервали, кількість їх виберемо так, щоб у середньому на кожен припадало близько 10 елементів вибірки  $i$ , відповідно, щоб теоретична ймовірність попадання у кожен інтервал не була мала (див. табл. 4.2,  $n = 100$ , число інтервалів  $L = 10$ ).

У третій стовпець таблиці введені передбачені гіпотезою імовірності  $p_i$  попадання у  $i$ -й інтервал, а в четвертий стовпець – реальне число елементів вибірки  $m_i$ , що потрапили в цей інтервал.

Тепер розглянемо кожен рядок таблиці окремо. Вироблено всього  $n = 100$  незалежних випробувань, імовірність події, наприклад, попадання в інтервал 2: від 0,1 до 0,2  $p_i = 0,1$ . Число  $m_i$  влучень в цей інтервал – випадкова величина, розподілена за законом Бернуллі.

Таблиця 4.2 – Аналіз функції розподілу за вибіркою

№ інтервалу	Границі інтервалу		$p_i$	$m_i$	$\eta_i^2 = \frac{(m_i - np_i)^2}{np_i(1 - p_i)}$
1	0,0	0,1	0,1	13	9/9
2	0,1	0,2	0,1	3	49/9
3	0,2	0,3	0,1	10	0
4	0,3	0,4	0,1	10	0
5	0,4	0,5	0,1	10	0
6	0,5	0,6	0,1	12	4/9
7	0,6	0,7	0,1	9	1/9
8	0,7	0,8	0,1	12	4/9
9	0,8	0,9	0,1	9	1/9
10	0,9	1,0	0,1	12	4/9

За великого  $n$  величина  $\eta = \frac{m - np}{\sqrt{npq}} \sim N(0,1)$ . Очевидно, що сума квадратів таких незалежних величин повинна мати розподіл  $\chi^2$  і тому може бути критерієм перевірки гіпотези. Число ступенів свободи для критерію береться на 1 менше, ніж число інтервалів ( $\nu = L - 1$ ), оскільки на величини  $\eta_i$ , в даному випадку, накладено один зв'язок:  $\sum m_i = n$ .

У таблиці в п'ятому стовпці поміщені експериментальні величини  $\eta_i$ . Обчислимо  $\chi_e^2 = \sum_{i=1}^L \eta_i^2 = 8$ . Вибравши рівень значущості  $\alpha = 0.05$ , звернемося до таблиць  $\chi^2$ -розподілу, щоб знайти критичну область Q. Вона буде односторонньою. Мала сума  $\chi_e^2$  означає, що  $m_i$  дуже наближене до  $np_i$ , тобто реальне число попадань в інтервал наближене до математичного очікування числа влучень. Це означає справедливність припущень про значення  $p_i$ , тобто гіпотеза  $H_0$  підтверджується. Відхилити її будемо лише у разі великої за абсолютною величиною відмінності  $m_i$  від  $np_i$ . У прикладі  $\nu = 10 - 1 = 9$ ,  $\chi_q^2(\alpha = 0,05) = 16,9$ .

Висновок:  $\chi_e^2 < \chi_q^2$ . Гіпотеза про те, що досліджуваний розподіл рівномірний, перевірена за критерієм  $\chi^2$  на рівні значущості 5 % і прийнята.

Під час перевірки припущення про нормальний закон розподілу можуть виникнути два випадки:

1) параметри цього закону  $M[x]$  і  $D[x]$  (математичне очікування і дисперсія) передбачаються відомими, тобто  $H_0: x \sim N(M[x], D[x])$ , у цьому випадку зручно ввести нову випадкову величину:

$$z \sim \frac{x - M[x]}{\sqrt{D[x]}} \sim N(0,1).$$

Користуючись таблицями нормального розподілу, легко знайти ймовірність попадання  $z$  в інтервали, які вибрані заздалегідь. При цьому необхідно стежити, щоб ймовірність попадання у кожний інтервал не була дуже малою, наприклад, як перший взяти інтервал від  $-\infty$  до  $-2$ . Ймовірність попадання в інтервал обчислюється за допомогою інтеграла ймовірностей  $\Phi(z)$  як різниця значень  $\Phi(z)$  від кінців інтервалу. Отримаємо, наприклад, такі ймовірності, подані в таблиці 4.3.

Тепер можна або перерахувати всі величини вибірки  $x$  у  $z$  за формулою, або перерахувати кінці інтервалів

$$x_i = z\sqrt{D[x]} + M[x],$$

і знайти, скільки реально значень  $x$  потрапляє в цей інтервал.

Таблиця 4.3 – Інтервали значень випадкової величини  $z \sim N(0,1)$  та ймовірності  $p_i$  попадання у ці інтервали

№ інтервалу	Границі інтервалу		$p_i$
	від	до	
1	$-\infty$	$-2,0$	0,023
2	$-2,0$	$-1,5$	0,044
3	$-1,5$	$-1,0$	0,092
4	$-1,0$	$-0,5$	0,15
5	$-0,5$	$0,0$	0,19
6	$0,0$	$0,5$	0,19
7	$0,5$	$1,0$	0,15
8	$1,0$	$1,5$	0,092
9	$1,5$	$2,0$	0,044
10	$2,0$	$\infty$	0,023

Подальші дії аналогічні описаним у прикладі з рівномірним розподілом. Критерій  $\chi^2$  має  $(L - 1)$  ступенів свободи, де  $L$  – число інтервалів;

2) якщо розподіл передбачається нормальним, але параметри його  $M[x]$  і  $D[x]$  не відомі, то знаходять їх оцінки з тієї ж вибірки за формулами. Гіпотеза  $H_0$ :  $x \sim N(\bar{x}, S_x^2)$ . Подальший алгоритм повністю аналогічний попередньому, тільки для критерію  $\chi^2$  береться  $\nu = L - 3$  ступені свободи, тому що накладені вже 3 зв'язки на величини  $\eta_i$  (оцінки математичного очікування і дисперсії обчислені з елементів тієї ж вибірки).

Інший спосіб перевірки гіпотези про закон розподілу ґрунтується на порівнянні «емпіричної функції розподілу», або «функції розподілу вибірки»  $W_n(x)$ , з передбачуваним законом. Для побудови цієї функції потрібно насамперед розташувати всі елементи вибірки  $(x_1, x_2, \dots, x_k, \dots, x_n)$  у порядку зростання.

Нехай  $x$  – деяка точка на осі  $x$ . Позначимо  $n_x$  – число елементів вибірки, розташованих «лівіше»  $x$ , тобто число елементів, величина яких менша за  $x$ , тоді

$$W_n(x) = \frac{n_x}{n}.$$

$F(x)$  – це ймовірність того, що випадкова величина менша за  $x$ .  $F(x)$  може

бути як безперервною, так і мати розриви.  $W_n(x)$  за визначенням має стрибок у точках  $x$ , рівних елементів вибірки.

Критерієм перевірки гіпотези про те, що випадкова величина має функцію розподілу  $F(x)$ , є вираз:

$$\omega^2 = \int_{-\infty}^{\infty} [W_n(x) - F(x)]^2 dF(x).$$

Очевидно, що чим менше  $\omega^2$ , тим ближче емпіричний розподіл до передбачуваного.

Ураховуючи, що  $W_n(x_1) = 0$  при  $x < x_1$ , ( $x_1$  – найменше значення вибірки):

$$W_n(x) = \frac{k}{n} \quad \text{при } x_k \leq x < x_{k+1},$$

$$W_n(x) = 1 \quad \text{при } x \geq x_n,$$

де  $x_n$  – найбільше значення вибірки.

Інтеграл можна перетворити до вигляду:

$$\omega^2 = \frac{1}{12n^2} + \frac{1}{n} \sum_{k=1}^{k=n} \left[ F(x_k) - \frac{2k-1}{2n} \right]^2.$$

Точний розподіл  $\omega^2$  надто складний, але дослідження показує, що вже при  $n = 40$  розподіл добутку  $n\omega^2$  наближений до деякого граничного розподілу, для якого обчислені таблиці. За цими таблицями визначені критичні значення для величини  $n\omega^2$ .

Розглянемо приклад застосування критерію  $\omega^2$ .

У стовпцях  $x_k$  наведена впорядкована за зростанням вибірка 50 значень випадкової величини (табл. 4.4). У стовпці  $k$  – порядковий номер елемента вибірки.

Гіпотеза  $H_0$ : випадкова величина  $x$  розподілена рівномірно в інтервалі  $(-1,1)$ . У цьому разі:

$$F(x) = \frac{1}{2} \int_{-1}^x dx = \frac{x}{2} + 0.5.$$

Значення  $F(x_k)$  також наведені в таблиці поруч із стовпцями  $(2k-1)/2n$ .

Вибираємо рівень значущості для перевірки гіпотези  $\alpha = 0.05$  і за таблицею знаходимо, що критична область значень  $n\omega^2$  визначається умовою:  $n\omega^2 > 0.461$ .

За формулою і даними табл. 2.5 обчислюємо експериментальне значення  $n\omega_e^2 (n = 50)$ :

$$n\omega_e^2 = \frac{1}{600} + (0.008 - 0.01)^2 + (0.016 - 0.03)^2 + \\ + (0.47 - 0.05)^2 + \dots + (0.991 - 0.99)^2 = 0.116.$$

Таблиця 4.4 – Перевірка гіпотези про закон розподілу за критерієм  $\omega^2$

K	$x_k$	$F(x_k)$	$\frac{2k-1}{2n}$	K	$x_k$	$F(x_k)$	$\frac{2k-1}{2n}$	k	$x_k$	$F(x_k)$	$\frac{2k-1}{2n}$
1	-0.984	0.008	0.01	18	-0.362	0.319	0.35	35	0.303	0.651	0.69
2	-0.968	0.016	0.03	19	-0.298	0.351	0.37	36	0.308	0.654	0.71
3	-0.906	0.047	0.05	20	-0.269	0.365	0.39	37	0.359	0.679	0.73
4	-0.905	0.048	0.07	21	-0.233	0.38	0.41	38	0.377	0.688	0.75
5	-0.879	0.061	0.09	22	-0.176	0.412	0.43	39	0.402	0.701	0.77
6	-0.866	0.067	0.11	23	-0.169	0.415	0.45	40	0.450	0.725	0.79
7	-0.816	0.092	0.13	24	-0.083	0.459	0.47	41	0.472	0.736	0.81
8	-0.749	0.125	0.15	25	-0.045	0.478	0.49	42	0.472	0.736	0.83
9	-0.737	0.132	0.17	26	0.001	0.501	0.51	43	0.524	0.762	0.85
10	-0.598	0.201	0.19	27	0.027	0.513	0.53	44	0.540	0.770	0.87
11	-0.562	0.219	0.21	28	0.033	0.516	0.55	45	0.662	0.831	0.89
12	-0.526	0.237	0.23	29	0.054	0.527	0.57	46	0.683	0.841	0.91
13	-0.506	0.247	0.25	30	0.059	0.529	0.59	47	0.769	0.885	0.93
14	-0.475	0.262	0.27	31	0.060	0.530	0.61	48	0.777	0.889	0.95
15	-0.468	0.266	0.29	32	0.075	0.537	0.63	49	0.896	0.948	0.97
16	-0.455	0.273	0.31	33	0.178	0.589	0.65	50	0.982	0.991	0.99
17	-0.446	0.277	0.33	34	0.182	0.591	0.67				

Тобто,  $n\omega_e^2 < 0.461$ , отже, гіпотеза про передбачуваний закон розподілу перевірена за критерієм  $\omega^2$  на рівні значущості 5 % і прийнята.

### Програма роботи

1. Ознайомитися з теоретичними положеннями.
2. За вказівкою викладача, експериментально отримати таблицю випадкової величини.
3. Сформулювати і перевірити гіпотезу про особливості закону розподілу запропонованої вибірки.
4. Зробити висновки та оформити звіт про виконану практичну роботу.

### **Порядок виконання роботи**

1. За вказівкою викладача, експериментально отримати таблицю випадкової величини.
2. Згідно із запропонованими методиками, сформулювати і перевірити гіпотезу про особливості закону розподілу отриманої вибірки.
3. Зробити висновок.

### **Обробка результатів. Зміст звіту**

1. Звіт має містити: номер, тему та мету практичної роботи, порядок виконання роботи з отриманими результатами та висновками.

### **Контрольні питання**

1. Надайте визначення поняття «нормальний розподіл».
2. Які існують форми розподілу, окрім нормального?
3. Чому саме нормальний розподіл використовують частіше за інші?
4. Поясніть суть та надайте алгоритм перевірки гіпотези про закон розподілу.
5. У яких випадках неможливо перевірити сформульовану гіпотезу? Що робити у цьому разі?
6. Які критерії вибору рівня значущості  $\alpha$ ?

## **Практична робота № 5**

### **Тема. Основи кореляційного та регресійного аналізу**

**Мета:** систематизація знань та навичок у галузі статистичного аналізу даних під час планування та організації експерименту із застосуванням кореляційного та регресійного аналізу.

### **Короткі теоретичні відомості**

Під час регресійного аналізу ставиться завдання на підставі даних експерименту відшукати функціональну залежність математичного очікування відгуку  $Mu$  від значень одного чи декількох факторів:  $Mu = f(x_1)$  або  $Mu = f(x_1, x_2)$  і т. д.

Починають зазвичай з лінійного регресійного аналізу. Передбачається, що спостережуване в досвіді значення відгуку  $y$  можна розділити на дві частини: одна з них закономірно залежить від  $x$  (тобто є функцією  $x$ ), інша частина – випадкова відносно до  $x$ . Відгук можна подати у вигляді:

$$y = f(x) + \varepsilon,$$

де  $\varepsilon$  – деяка випадкова величина. Іноді  $\varepsilon$  називають помилкою експерименту, пов'язуючи її з недосконалістю методу вимірювання. Виникає задача про відшукування функції відгуку  $f(x)$ . Клас функцій, у якому шукається функція  $f(x)$ , часто визначається фізичною суттю процесів. Це можуть бути експоненти, статичні функції, але починають підбір і аналіз, зазвичай, з лінійної функції:

$$f(x) = a + bx.$$

Перш ніж переходити до визначення коефіцієнтів  $a$  і  $b$ , проводять обчислення коефіцієнта кореляції, величина якого указує на обґрунтованість прийняття лінійної моделі.

Коефіцієнтом кореляції називають величину:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}.$$

Коефіцієнт кореляції бере значення в інтервалі  $[-1,1]$ . Чим більша абсолютна величина коефіцієнта кореляції, тим з більшою підставою можна припускати наявність лінійної залежності між змінними. У цьому разі можна переходити до визначення коефіцієнтів лінійної залежності, наприклад, методом найменших квадратів.

Якщо передбачається нелінійна залежність  $y$  від  $x$ , то за допомогою перетворень нерідко вдається привести залежність до лінійної та скористатися потім інструментами визначення коефіцієнтів у лінійному рівнянні регресії.

У пакеті Microsoft Excel функції обчислення коефіцієнта кореляції та побудови лінійної регресії є вбудованими.

### **Програма роботи**

1. Ознайомитися з теоретичними положеннями.



2. За вказівкою викладача, експериментально отримати таблицю випадкової величини.

3. Провести кореляційний аналіз.

4. Зробити висновки та оформити звіт про виконану практичну роботу.

### **Порядок виконання роботи**

1. За вказівкою викладача, експериментально отримати таблицю випадкової величини (приблизно 20 вимірювань незалежної змінної  $x$  і залежної змінної  $y$ ).

2. У табличному редакторі MS EXCEL сформувати таблицю за зразком на рис. 5.1а.

3. Обчислити коефіцієнт кореляції між змінними  $x$  і  $y$ . Для цього використати вбудовану в EXCEL функцію **КОРРЕЛ**. Чим ближче значення коефіцієнта кореляції до 1, тим більша залежність між  $x$  і  $y$ .

4. Отримати значення коефіцієнтів регресії. Для цього використати вбудовану в EXCEL функцію **ЛИНЕЙН**. Оскільки ця функція повертає два числа: коефіцієнти  $A$  і  $B$  рівняння регресії, спочатку потрібно виділити дві комірки та ввести вираз у фігурних дужках (рис. 5.1б), або натиснути Ctrl + Shift + Enter (так звана змінна типу МАСИВ у EXCEL).

5. Побудувати діаграму з зображенням графіка вихідних значень  $y$  і прямої регресії (рис. 5.2).

Первинний аналіз		
результатів експеримента		
Побудова рівняння регресії		
№	x	y
1	10	11,08
2	20	12,18
3	30	12,28
4	40	16,38
5	50	17,48
6	60	19,58
7	70	20,68
8	80	25,78
9	90	26,88
10	100	27,98
11	110	30,08
12	120	35,18
13	130	33,28
14	140	39,38
15	150	40,48
16	160	39,58
17	170	42,68
18	180	44,78
19	190	49,88
20	200	47,98

fx {=ЛИНЕЙН(Таблица3[y];Таблица3[x])}				
	D	E	F	G
	Коефіцієнти рівняння регресії			
	y = Ax + B	A	B	
		0,214662	7,490526	

а)

б)

Рисунок 5.1 – Отримання коефіцієнтів регресії

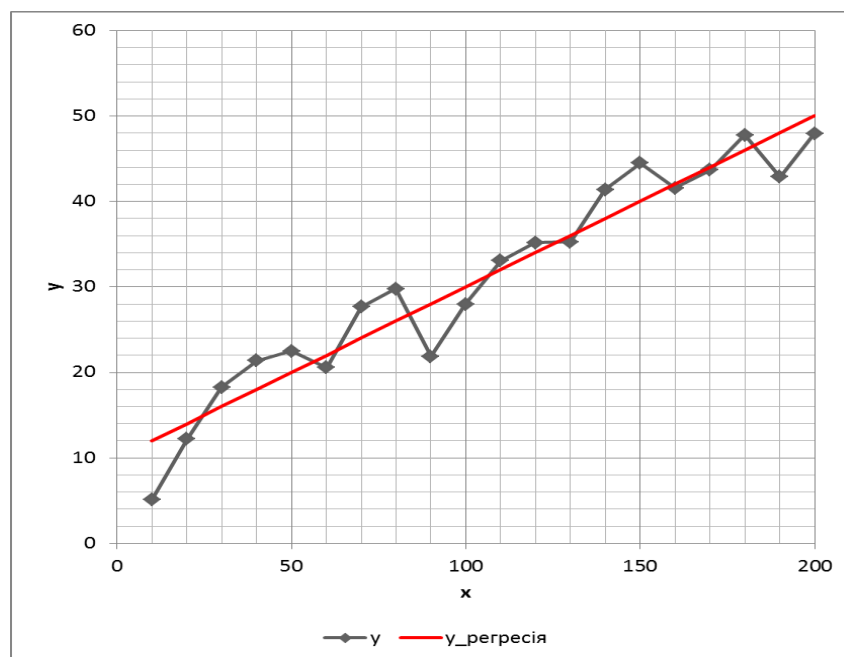


Рисунок 5.2 – Графік результатів регресійного аналізу

## **Обробка результатів. Зміст звіту**

1. Звіт має містити: номер, тему та мету практичної роботи, порядок виконання роботи з отриманими результатами та висновками.

### **Контрольні питання**

1. Надайте визначення понять: «кореляція», «коефіцієнт кореляції», «регресія».
2. Поясніть суть та надайте алгоритм регресійного аналізу.
3. У яких випадках неможливо побудувати рівняння регресії?

## **Практична робота № 6**

### **Тема. Однофакторний дисперсійний аналіз**

**Мета:** ознайомлення з призначенням і передумовами застосування методу однофакторного дисперсійного аналізу, з його ідеєю та обчислювальним алгоритмом для обробки експериментальної інформації.

### **Короткі теоретичні відомості**

У багатьох галузях практичної діяльності є об'єкти дослідження, стан яких визначається входними змінними або факторами, що не мають кількісного опису. Такими факторами можуть бути некеровані та керовані зміни, що з будь-якої причини не дозволяють здійснювати їхнє вимірювання у експерименті; а також ті неконтрольовані змінні, випадкові рівні варіювання яких можна довільно вибирати і фіксувати в часі.

Для вивчення впливу подібних факторів на вихідний параметр об'єкта, для загального оцінювання впливу факторів, для ранжування і виділення серед них важливих непридатні методи аналізу, які передбачають вимірювання рівнів факторів, які досліджуються.

Приклади:

1) під час обробки деталей паралельно на декількох верстатах автоматичної лінії потрібно з'ясувати, чи однотипні середні розміри деталей, що виготовляються на різних верстатах, тобто оцінити, чи значно впливає

фактор індивідуальності верстата на процес обробки;

2) під час використання радіодеталей з декількох партій треба визначити, чи суттєво відрізняються параметри деталей різноманітних партій.

Постановка задачі в загальному вигляді.

Дано:

а) вихідний параметр  $y$  може залежати (з фізичних міркувань) від  $n$  незалежних факторів  $x_1 \dots x_n$ , що не мають кількісного опису, і їхніх парних взаємодій (використовуємо єдине позначення для змінних: керованих  $x$ , некерованих  $z$ , неконтрольованих  $w$ );

б) кожен фактор  $x$  може варіюватися на  $u_1$  рівнях;

в) повний факторний експеримент складається з  $N = u_1 \cdot u_2 \cdot \dots \cdot u_n$  серій незалежних спостережень за числом усіх можливих сполучень (які не повторюються) рівнів  $n$  факторів;

г) будь-яка  $j$ -а серія містить  $m_j$  спостережень  $y_{j1} \dots y_{jm_j}$  дослідів, які дублюються.

Потрібно: визначити, наскільки суттєвим на фоні випадкових похибок є вплив того чи іншого фактора  $x_i$  або комбінації (взаємодії) таких факторів на вихідний параметр  $y$ , провести порівняння з іншими факторами і виділити найбільш важливі.

Припущення, на яких ґрунтується дисперсійний аналіз.

Припущення 1. Величина  $y$  – нормально розподілена випадкова величина з центром розподілу  $M[y] = \varphi(x_1 \dots x_n) = const$  і з дисперсією  $\sigma[y] = const$ . Отже, фактори визначають усереднену величину  $y$ , залишаючи простір для випадкових помилок спостережень, які підпорядковуються нормальному закону.

Припущення 2. Дисперсія одиничного спостереження  $\sigma^2$ , обумовлена випадковими помилками, постійна в усіх дослідженнях і не залежить від  $x_1 \dots x_n$ , тобто дисперсії  $\sigma_j^2[y]$  будуть дорівнювати одна одній при  $j = 1 \dots N$ , а їх вибіркові оцінки  $S_i^2[y]$  будуть однорідні – умова відтворення дослідів.

З даних задачі та зазначених припущень зрозуміло, що чим більший

вплив деякого фактора  $x$  на вихідний параметр  $y$ , тим більше розходження між собою середніх арифметичних  $\bar{y}_j$  серій паралельних спостережень, проведених при різних сполученнях рівнів варіювання досліджуваних факторів. Статистичний аналог такого розходження вказує на суттєвий вплив факторів.

За двома серіями спостережень порівняння середніх і перевірка нуль-гіпотези здійснюються за допомогою  $t$ -критерію методом Стюдента. У сформульованій задачі вимагається одночасно довільно порівняти велику кількість середніх і на підставі цього зробити висновок про значущість впливу того чи іншого фактора.

Ідея дисперсійного аналізу. Щоб мати можливість оцінити вплив кожного фактора на вихідний параметр і порівняти вплив різних факторів, визначимо деякий показник цього впливу.

Нехай за відсутності помилок досліду ( $\sigma^2 = 0$ ) під час варіювання фактора  $x$  на різних рівнях отримані фактичні значення  $y_j$  вихідного параметра  $y$ . Тоді за показник впливу фактора  $x$  беруть величину, яку називають дисперсією фактора  $x$ :

$$\sigma_x^2 = \frac{1}{u} \sum_{j=1}^u (y_j - \bar{y})^2,$$
$$\bar{y} = \frac{1}{u} \sum_{j=1}^u y_j.$$

При цьому треба мати на увазі, що числа  $y_j$  не є випадковими, і тому, дисперсія  $\sigma_x^2$  не пов'язана ні з якою випадковою величиною, оскільки припускається, що  $\sigma_x^2 = 0$ .

Визначати вплив факторів за величинами їх дисперсій зручно, оскільки це простіша ступінь розсіювання і, до того ж, аналогічна ступеню впливу фактора випадкових причин, тобто аналогічна дисперсії одиничного спостереження (відтворення)  $\sigma^2$ . Завдяки цьому є можливість порівнювати вплив будь-якого досліджуваного та випадкового факторів. Таке дослідження факторів за їх дисперсіями називають дисперсійним аналізом.

Розглянемо ідею дисперсійного аналізу на прикладі вивчення впливу одного фактора на фоні випадкових похибок, якщо дисперсія відтворення  $\sigma^2$  відома. Під час варіювання фактора  $x$  на  $u$  рівнях як результат спостережень отримаємо значення  $y_1 \dots y_u$ , розсіювання яких можна характеризувати вибірковою дисперсією:

$$s^2 = \frac{1}{u-1} \sum_{i=1}^u (y_i - \bar{y})^2,$$

з числом ступенів свободи  $\nu = u - 1$ .

Якщо різниця між  $s^2$  і  $\sigma^2$  незначна, то з цього випливає, що розкид спостережень, який зумовлений цією різницею, пов'язаний лише з випадковими причинами, тому вплив фактора  $x$  незначний. Якщо різниця між  $s^2$  і  $\sigma^2$  значна, підвищений розкид спостережень викликаний не лише випадковими причинами, а і впливом фактора  $x$ , який необхідно визнати суттєвим.

Оскільки в останньому випадку додаються впливи двох незалежних факторів – випадкових причин (з дисперсією  $\sigma^2$ ) і фактора  $x$  (з дисперсією  $\sigma_x^2$ ), що призводить до загального розсіювання спостережень, то загальна дисперсія буде сумою ( $\sigma^2 + \sigma_x^2$ ), а її оцінка буде:

$$s^2 \approx \sigma^2 + \sigma_x^2.$$

Тому дисперсія фактора визначається за виразом:

$$\sigma_x^2 \approx s^2 - \sigma^2.$$

Загалом, якщо дисперсія відтворення  $\sigma^2$  невідома, схема дисперсійного аналізу має дозволити знайти її оцінку поряд з оцінками дисперсій досліджуваних факторів. Для цього планується проведення серій дослідів дублювання, але кожного з усіх можливих сполучень рівнів досліджуваних факторів.

Отже, основна ідея дисперсійного аналізу полягає в розкладенні загальної дисперсії  $s^2$  на складові, які залежать від випадкових причин, від кожного з розглянутих факторів і від їх взаємодії, а також в оцінюванні статистичної

значимості дисперсій останніх з урахуванням похибки відтворення досліду.

### **Завдання на практичну роботу**

Дослідіть, чи залежить довговічність  $y$  (вимірюється у годинах) електричних машин від технології та матеріалу виготовлення (фактор  $x$ ). Припустимо, що виконуються допущення дисперсійного аналізу, тобто довговічність має нормальний розподіл і вплив роду матеріалу та технології виготовлення електричних машин не має впливу на дисперсію  $\sigma^2$  величини  $y$ , але викликає розбіжність середніх значень. Для досліджень були відібрані нерівнокількісні серії зразків з чотирьох партій продукції ( $u = 4$ ).

### **Програма роботи**

1. Ознайомитися з теоретичними положеннями.
2. За вказівкою викладача, експериментально отримати таблицю даних для дисперсного аналізу.
3. На підставі розрахунків перевірити значущість впливу досліджуваного фактора.
4. Зробити висновки та оформити звіт про виконану практичну роботу.

### **Порядок виконання роботи**

1. За вказівкою викладача, експериментально отримати таблицю випадкової величини (приклад в табл. 6.1).

Таблиця 6.1 – Приклад експериментальних даних

X	Y							
	1	2	3	4	5	6	7	8
1	1600	1610	1650	1680	1700	1700	1800	–
2	1580	1640	1640	1700	1750	–	–	–
3	1460	1550	1600	1620	1640	1660	1740	1820
4	1510	1520	1530	1570	1600	1680	–	–

2. Для зручності розрахунків зменшити всі результати вимірювань на однакову величину (наприклад, 1500 год).
3. Піднести отримані результати ( $y_{ji}$ ) до квадрата ( $y_{ji}^2$ ).

4. Знайти спочатку суму ( $\sum_{j=1}^u y_{j1}$ ) перших результатів вимірювань з кожної партії, потім других і т. д.

5. Знайти суму результатів (підсумок,  $Y_j$ ) вимірювань у кожній спробі (партії).

6. Піднести до квадрата отримані підсумки ( $Y_j^2$ ).

7. Поділити отримані (для кожної партії) результати на кількість зразків у партії ( $\frac{1}{m_j} Y_j^2$ ).

8. Знайти суми піднесених до квадрата результатів вимірювань для кожної партії окремо ( $\sum_{l=1}^{m_j} y_{jl}^2$ ).

9. Знайти суму результатів усіх вимірювань ( $\sum_{j=1}^u Y_j$ ).

10. Знайти суму квадратів підсумків за серіями, поділеними на кількості спостережень у серіях:

$$Q_2 = \sum_{j=1}^u \frac{Y_j^2}{m_j}.$$

11. Знайти суму піднесених до квадрата результатів усіх вимірювань:

$$Q_1 = \sum_{j=1}^u \sum_{l=1}^{m_j} y_{jl}^2.$$

12. Результати розрахунків у пунктах 2...11 занести до табл. 6.2.

Таблиця 6.2 – Розрахункова таблиця

X	Y								
	1		...		M		$Y_j$	$\frac{1}{m_j} Y_j^2$	$\sum_{l=1}^{m_j} y_{jl}^2$
	$y_{j1}$	$y_{j1}^2$	...	...	$y_{jm}$	$y_{jm}^2$			
1									
...									
4									
$\sum_{j=1}^u$									

13. Знайти загальну кількість вимірювань (M).

14. Знайти загальне середнє для усіх вимірів:

$$\bar{y} = \frac{1}{M} \sum_{j=1}^u Y_j.$$

15. Знайти квадрат загального підсумку, поділений на кількість усіх



спостережень:

$$Q_3 = \frac{1}{M} \left( \sum_{j=1}^u Y_j \right)^2.$$

16. Знайти загальну суму квадратів відхилень:

$$S = Q_1 - Q_3.$$

17. Знайти суму квадратів відхилень по середині серій:

$$S_0 = Q_1 - Q_2.$$

18. Знайти суму квадратів відхилень між серіями вимірювань:

$$S_X = Q_2 - Q_3 = S - S_0.$$

19. Знайти кількість ступенів свободи по середині серій:

$$v_0 = N - u.$$

20. Знайти кількість ступенів свободи між серіями:

$$v_X = u - 1.$$

21. Знайти загальну кількість ступенів свободи:

$$v = M - 1.$$

22. Знайти загальну вибірккову дисперсію:

$$S^2 = S/v.$$

23. Знайти вибірккову дисперсію по середині серій:

$$S_0^2 = S_0/v_0.$$

24. Знайти вибірккову дисперсію між серіями:

$$S_X^2 = S_X/v_X.$$

25. Виконати перевірку значущості впливу фактора  $x$ , знайшовши дисперсійне відношення (розрахункове значення критерію Фішера):

$$F = S_X^2/S_0^2.$$

26. Використовуючи таблицю з додатка, визначити значення F-критерію для п'ятивідсоткового рівня значущості та кількості ступенів свободи  $v_1$  і  $v_2$   $F_{0,05}(v_1; v_2)$ .

27. Зробити висновок. Якщо  $F < F_{0,05}(v_1; v_2)$ , вплив роду матеріалу та технології виготовлення на довговічність електричних машин у партіях, що розглядалися, можна вважати незначним, інакше – значним.

## **Оброблення результатів. Зміст звіту**

1. Звіт має містити: номер, тему та мету практичної роботи, порядок виконання роботи з отриманими результатами та висновками.

### **Контрольні питання**

1. Надайте визначення понять: «дисперсія», «дисперсійний аналіз».
2. Поясніть суть та надайте алгоритм дисперсійного аналізу.
3. У яких випадках неможливо виконати дисперсійний аналіз?

## **Практична робота № 7**

### **Тема. Основи планування повного факторного експерименту**

**Мета:** вивчення основ планування факторних експериментів, технології планування повного факторного експерименту.

### **Короткі теоретичні відомості**

У повному факторному експерименті реалізуються усі можливі сполучення рівнів факторів. Якщо число рівнів кожного фактора дорівнює двом (верхній та нижній), то маємо повний факторний експеримент (ПФЕ) типу  $2^m$ , де  $m$  – число факторів. Число дослідів ПФЕ  $n = 2^m$ , а кодовані значення факторів будуть рівні:  $+1$  (верхній рівень);  $-1$  (нижній рівень). У матрицях планування іноді позначають замість  $(+1)$  просто  $(+)$ , замість  $(-1)$  просто  $(-)$ .

Побудова матриць плану для числа факторів більше 2 ґрунтується у більшості випадків на трьох прийомах.

1. Перші два фактори заповнюються як для двофакторного експерименту, а третій фактор у перших чотирьох дослідів тільки на верхньому рівні  $(+)$ , у останніх чотирьох дослідів – на нижньому рівні  $(-)$ .

2. Другий прийом заснований на поелементному перемножуванні вектор-стовпців матриці для третього фактора для перших чотирьох дослідів. Для останніх чотирьох дослідів знаки третього фактора змінюються на зворотні.

3. Третій прийом ґрунтується на правилі чергування знаків: у першому стовпці знаки змінюються по чергову, у другому стовпці чергуються через 2, у

третьому – через 4 і т. д.

ПФЕ має такі особливості.

1. Дослідні точки розташовуються оптимально з точки зору точності математичного опису.

2. Усі фактори і, відповідно, коефіцієнти рівняння оцінюються незалежно один від одного, що забезпечується незалежністю та ортогональністю матриці планування.

3. Якщо вплив на відгук у квадратичних членів  $x_2$  стає значущим і необхідно оцінити коефіцієнти при цих членах, то за допомогою ПФЕ не вдається оцінити їх коефіцієнти, тому що відповідні стовпці матриці планування ідентичні між собою і, отже, матриця стає не ортогональною, із залежними стовпцями.

4. Вплив квадратичних членів рівняння змішується з вільним членом.

Під час проведення експерименту необхідно враховувати, що значення відгуку ( $y$ ) через наявність неконтрольованих і не врахованих факторів матиме випадковий характер. Для оцінки дисперсії  $y$  можна дублювати досліди в кожній точці плану експерименту. Число паралельних дослідів зазвичай вибирають рівним 2...4, тоді  $i$ -те середнє значення  $\bar{y}_i$  у матриці планування дорівнюватиме

$$\bar{y}_i = \frac{1}{l} \sum_{j=1}^l y_{ij},$$

де  $l$  – число паралельних дослідів у  $i$ -тому рядку матриці;  $j$  – номер дослідів.

Для забезпечення випадкового порядку проведення експерименту (для виключення систематичної помилки), за якого всі неконтрольовані фактори були б рандомізовані, перед реалізацією плану експерименту послідовність серій дослідів визначається за допомогою таблиці випадкових чисел.

Статистична обробка результатів ПФЕ виконується за таким алгоритмом.

1. Якщо є повтори (паралельні досліди), то визначаються середні значення параметра оптимізації (відгуку)  $\bar{y}_i$  для кожного рядка плану.

2. Визначаються оцінки дисперсій за рядками:

$$S_i^2 = \sum_{j=1}^l (y_{ij} - \bar{y}_i)^2 / (l - 1).$$

3. Проводиться контроль відтворюваності результатів дослідження, тобто перевірка рівноточності вимірювань. Така перевірка необхідна з малим числом дослідів, тому що в цьому разі навіть одна груба помилка може дуже спотворити результати. З рівним числом повторів перевірка відтворюваності проводиться за допомогою критерію Кохрена, згідно з яким обчислюється сума дисперсій рядків:

$$S_{\Sigma}^2 = \sum_{i=1}^n S_i^2.$$

Для перевірки рівноточності необхідно вибрати найбільшу з порядкових дисперсій  $S_{imax}^2$  і обчислити  $G$  – критерій:

$$G = \frac{S_{imax}^2}{S_{\Sigma}^2}.$$

Якщо  $G < G_T$ , де  $G_T$  – табличне значення критерію з числом ступенів свободи  $\nu_1 = l - 1$  і  $\nu_2 = n$ , то дослідів рівноточні. За нерівноточних дослідів потрібно збільшити число паралельних дослідів у рядку.

4. Для ортогональної матриці планування виконуються такі співвідношення:

$$\sum_{i=1}^n f_{i0} = n; \quad \sum_{i=1}^n f_{ik} f_{ik} = n; \quad \sum_{i=1}^n f_{ik} f_{ij} = 0.$$

Коефіцієнти рівняння регресії знаходять за виразами:

$$\theta_0 = \frac{1}{n} \sum_{i=1}^n y_i; \quad \theta_k = \frac{1}{n} \sum_{i=1}^n f_{ik} y_i.$$

Будь-яка матриця планування з числом рівнів більше двох не буде ортогональною. Але такий експеримент можна провести і визначити коефіцієнти рівняння як у звичайному регресійному аналізі.

5. Проводиться оцінювання дисперсії відтворюваності:

$$S_e^2 = \frac{1}{n} \sum_{i=1}^n S_i^2.$$

Якщо в кожній точці плану проводиться тільки один дослід, то дисперсія відтворюваності може бути визначена в незалежних від матриці планування дослідах або за паралельними дослідом в одній точці плану експерименту.

Оцінка дисперсії відтворюваності для середнього значення відгуку дорівнює:

$$S^2(\bar{y}) = S_e^2/l,$$

а середньоквадратичне відхилення середнього значення  $y$ :

$$S(\bar{y}) = \sqrt{S^2(\bar{y})}.$$

Для перевірки значущості отриманих коефіцієнтів використовується критерій Стюдента  $t$  з рівнем значущості  $\alpha$  (зазвичай  $\alpha$  вибирається рівним 0,05) і числом ступенів свободи:

$$\nu = (l - 1)n.$$

Для ортогонального планування оцінки дисперсії коефіцієнтів рівняння регресії рівні між собою та визначаються за формулою:

$$S_{\theta}^2 = S^2(\bar{y})/n.$$

Коефіцієнт рівняння статистично значущий, якщо:

$$|\theta_i| > t_{\alpha\nu} \sqrt{S_{\theta}^2}.$$

Статистична незначущість коефіцієнта  $\theta_i$  може бути викликана такими обставинами:

1. Базовий рівень цієї змінної  $x_{0i}$  або добуток змінних наближені до точки приватного екстремуму.
2. Крок зміни змінної вибраний занадто малим.
3. Велика помилка експерименту через наявність некерованих і неконтрольованих змінних.
4. Цей фактор або взаємодія факторів справді не роблять істотного впливу на значення параметра відгуку  $y$ .

Оскільки за ортогонального планування коефіцієнти рівняння регресії оцінюються незалежно один від одного, то за наявності незначних коефіцієнтів перерахунку не потрібно.

Відкинувши (прирівнявши до нуля) незначущі коефіцієнти, отримаємо рівняння зв'язку між відгуком і факторами.

Для перевірки адекватності отриманої математичної моделі проводиться оцінювання дисперсії адекватності:

$$S_{ад}^2 = \frac{1}{n - d} \sum_{i=1}^n (\bar{y} - \bar{y}_{розр})^2,$$

де  $d$  – число значущих коефіцієнтів у рівнянні.

Підраховується критерій Фішера:

$$F = \frac{S_{ад}^2}{S^2(\bar{y})}.$$

За таблицею знаходиться табличне значення  $F$  для обраного рівня значущості та числа свободи чисельника  $\nu_1 = n - d$ , знаменника  $\nu_2 = n(l - 1)$ . Якщо  $S_e^2(y)$  визначається в окремому досліді, то  $\nu_2 = l - 1$ , де  $l$  – число паралельних дослідів, за якими обчислена  $S_e^2(y)$ .

### **Програма роботи**

1. Ознайомитися з теоретичними положеннями.
2. Скласти план і провести експеримент.

### **Порядок виконання роботи**

1. Скласти план повного факторного експерименту для трьох факторів.
2. За вказівкою викладача, експериментально отримати таблицю даних відповідно до створеного плану.
3. Виконати розрахунки згідно з методикою ПФЕ.
4. Зробити висновки та оформити звіт про виконану практичну роботу.

### **Обробка результатів. Зміст звіту**

1. Звіт має містити: номер, тему та мету практичної роботи, порядок виконання роботи з отриманими результатами та висновками.

### **Контрольні запитання**

1. Надайте визначення понять: «фактор», «повний факторний експеримент».
2. Поясніть суть та надайте алгоритм ПФЕ.
3. У яких випадках недоцільно виконувати ПФЕ?
4. Яка кількість дослідів необхідна для експерименту з кількістю факторів: 2, 3, 4, 5?
5. Як можна зменшити час проведення експерименту з великою кількістю факторів?
6. Які висновки можна зробити з результатів повного факторного експерименту з кількістю факторів – 2.
7. У чому полягає суть обробки результатів повного факторного експерименту?
8. Як можна зменшити вплив випадкових і невідомих на початку експерименту факторів на результати?
9. Наведіть приклади повного факторного експерименту з різною кількістю факторів.
10. Чи можна провести ПФЕ з неповною кількістю факторів?

## Практична робота № 8

### Тема. Основи планування повного факторного експерименту

**Мета роботи:** отримання практичних навичок побудови планів дробового факторного експерименту та обробки його результатів методами регресійного аналізу даних з метою отримання математичної моделі, що адекватно описує досліджуваний процес.

### Теоретичні відомості

Дробовий факторний експеримент (ДФЕ) Відрізняється від ПФЕ тим, що дозволяє суттєво зменшити кількість дослідів. В теорії планування експериментів є методи побудови частин планів ПФЕ, які мають властивості планів ПФЕ.

Такі плани називаються *дробовими*. Якщо використовується половина плану  $\frac{1}{2}$  ПФЕ  $2^k$ , то такий план називається ДФЕ  $2^{k-1}$ , називається *напівреплікою* ПФЕ. Якщо –  $\frac{1}{4}$  ДФЕ  $2^{k-2}$  називається *четвертною реплікою* ПФЕ.

Будь які дробові репліки повинні мати властивості ортогональності, нормування, симетрія та ротатабельність.

При створенні планів ДФЕ спочатку будується план ПФЕ, але для меншої кількості факторів. А потім несуттєве: взаємодія факторів представляє відсутній фактор.

На прикладі побудови ДФЕ  $2^{4-1}$  будуюмо план для меншої кількості ПФЕ  $2^3$  (табл. 8.1).

В результаті обробки експерименту отримаємо рівняння регресії:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{123}x_1x_2x_3.$$

Таблиця 8.1. План експерименту з меншою кількістю факторів – ПФЕ  $2^3$ .

Номер дослідів	$x_1$	$x_2$	$x_3$	$x_1x_2$	$x_1x_3$	$x_2x_3$	$x_1x_2x_3$
1	-1	-1	-1	+1	+1	+1	-1
2	+1	-1	-1	-1	-1	+1	+1
3	-1	+1	-1	-1	+1	-1	+1
4	+1	+1	-1	+1	-1	-1	-1
5	-1	-1	+1	+1	-1	-1	+1
6	+1	-1	+1	-1	+1	-1	-1
7	-1	+1	+1	-1	-1	+1	-1
8	+1	+1	+1	+1	+1	+1	+1



Нехай взаємодія першого і третього факторів буде не суттєвою, тобто  $x_1x_3$  приблизно дорівнює нулю. Тоді для реалізації плану  $ДФЕ 2^{4-1}$  замінюємо відсутній фактор  $x_4$  цією несуттєвою взаємодією –  $x_4 = x_1x_3$ . В результаті цього отримаємо такий план проведення експерименту, який матиме назву дробовий факторний експеримент  $ДФЕ 2^{4-1}$ , а вираз –  $x_4 = x_1x_3$  називають його генеруючим співвідношенням.

План його проведення  $ДФЕ 2^{4-1}$  наведено в табл.8.2.

Таблиця 8.2. План проведення експериментів  $ДФЕ 2^{4-1}$ .

Номер дослід	$x_0$	$x_1$	$x_2$	$x_3$	$x_4 = x_1x_2x_3$	$x_1x_2$	$x_1x_3$	$x_1x_4$	$x_2x_3$	$x_2x_4$	$x_3x_4$	$x_1x_2x_3x_4$
1	+1	-1	-1	-1	-1	+1	+1	+1	+1	+1	+1	-1
2	+1	+1	-1	-1	+1	-1	-1	+1	+1	-1	-1	+1
3	+1	-1	+1	-1	+1	-1	+1	-1	-1	+1	-1	+1
4	+1	+1	+1	-1	-1	+1	-1	-1	-1	-1	+1	-1
5	+1	-1	-1	+1	+1	+1	-1	-1	-1	-1	+1	+1
6	+1	+1	-1	+1	-1	-1	+1	-1	-1	+1	-1	-1
7	+1	-1	+1	+1	-1	-1	-1	+1	+1	-1	-1	-1
8	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1

Тоді математична модель буде представлятись рівнянням регресії такого виду:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{123}x_1x_2x_3$$

Недоліком цього плану є те, що деякі із стовпчиків співпадають, тому їх треба викреслити, а це означає, наприклад, що  $b_1$  є оцінкою коефіцієнтів  $a_1 + a_{34}$ :

$$b_3 \rightarrow a_1 + a_{14}$$

$$b_1 \rightarrow a_4 + a_{13}$$

Для виявлення цього ефекту змішування оцінок введено поняття визначального контрасту, який отримано шляхом множення лівої та правої частин генеруючого співвідношення на введену змінну (в нашому випадку це  $x_4$ ).

Тоді визначальний контраст матиме такий вигляд:

$$1 = x_1x_3x_4$$

Визначальний контраст дозволяє визначити всі змішування оцінок коефіцієнтів математичної моделі шляхом множення лівої та правої частин його виразу на відповідні змінні або їх комбінації, наприклад:

$$x_3 = x_1 x_4$$

$$x_2 = x_1 x_2 x_3 x_4$$

$$x_1 = x_3 x_4$$

Тобто можна визначити змішування оцінок, що для наведених виразів має такий вигляд:

$$b_1 \rightarrow a_1 + a_{34}$$

$$b_2 \rightarrow a_2 + a_{1234}$$

$$b_3 \rightarrow a_3 + a_{14}$$

$$x_1 x_2 = x_2 x_3 x_4$$

$$b_3 \rightarrow a_{12} + a_{234}$$

Дробові репліки дозволяють значно зменшити кількість дослідів. Можна зменшити кількість дослідів до такого рівня, коли вона буде дорівнювати кількості факторів.

План в якому кількість дослідів дорівнює кількості факторів називається *насиченим*.

### Завдання для виконання практичної роботи

Для набору вихідних даних, що отримані в результаті проведення експериментальних досліджень (варіант отримати у викладача), побудувати план ПФЕ $2^{k-1}$  при  $k=4$  та  $k=5$ , які є значеннями функції відгуку проведених експериментальних досліджень, та виконати послідовну їх обробку методами регресійного аналізу. Для цього необхідно виконати таку послідовність дій.

1. Побудувати план проведення дробового факторного експерименту для наданої кількості факторів з врахуванням генеруючого співвідношення.

2. Видалити з плану проведення експерименту дублювання стовпців та визначити взаємний вплив факторів на коефіцієнти рівняння регресії.

3. Виконати переведення початкових факторів із натуральної в кодовану шкалу вимірювання.

4. Реалізувати процес статистичної обробки результатів планового експерименту в такій послідовності:

– виконати перевірку значень вихідної величини на наявність грубих похибок за  $t$  - критерієм Стьюдента;

– виконати перевірку отриманих результатів експерименту на однорідність вимірів за  $G$ - критерієм Кохрена;

– виконати перевірку гіпотези про однорідність дисперсій;

– визначити коефіцієнти рівняння регресії для заданого виду математичної моделі;

– виконати перевірку статистичної значимості отриманих коефіцієнтів рівняння регресії ;за  $t$  - критерієм Стьюдента;

– виконати перевірку отриманої математичної моделі на адекватність опису досліджуваного об'єкта за F- критерієм Фішера.

5. Навести результати обчислень за кожним етапом та зробити відповідні висновки.

6. Представити отриману математичну модель, де початкові фактори надано в натуральній шкалі вимірювання.

7. Зробити висновки за результатами виконання практичної роботи.

### **Контрольні запитання**

1. Який експеримент називається дробовим?
2. Що таке напіврепліка, четвертна репліка тощо?
3. Які властивості мають плани дробового факторного експерименту (ДФЕ)?
4. Як будується план ДФЕ?
5. Що називається генеруючим співвідношенням?
6. Як вибирається генеруюче співвідношення?
7. Що таке визначаючий контраст, на що він вказує?
8. Як визначити змішування оцінок коефіцієнтів у моделі, що одержана в результаті ДФЕ?
9. Який план ДФЕ називається насиченим?
10. Які переваги та недоліки має ДФЕ і коли його рекомендується використовувати?

## Практична робота № 9

### Тема. Використання методу парних порівнянь та аналізу ієрархій

**Мета роботи:** вивчити основні правила використання методу парних порівнянь та методу аналізу ієрархій у різних сферах застосування.

#### Теоретичні відомості

Проведення процедури виявлення пріоритетів серед альтернатив, об'єктів, критеріїв, чинників, характеристик (далі – об'єктів) тощо може бути здійснено методом парних порівнянь. Цей метод є важливим компонентом методу аналізу ієрархій, коли об'єкти упорядковуються і підпорядковуються за рівнями ієрархії. Метод парних порівнянь має і самостійне значення.

Парні порівняння виконує експерт або група експертів під керівництвом керівника (модератора). «Ваги» об'єктів є суб'єктивними думками і потребують оцінювання узгодженості. У табл. 9.1 наведена шкала відносної важливості об'єктів для їх парного порівняння. При цьому порівняння здійснюється відносно до деякого вищого за ієрархією критерію. Шкали бувають трибальні, п'ятибальні та найпоширеніші – дев'ятибальні (шкала Сааті).

Таблиця 9.1. – Шкала відносної важливості для парного порівняння

Ступінь переваги	Визначення
0	Незалежні
1	Однакова важливість (значущість)
2	Деяка (слаба) перевага значущості
3	Значна перевага значущості
4	Надзвичайна перевага значущості
5	Абсолютна перевага значущості

Матриця парних порівнянь об'єктів. Основним елементом для зображення інтенсивності взаємовпливу об'єктів є матриця парних порівнянь. Об'єкти, що знаходяться на одному рівні ієрархії мають однакові набори показників. Значення цих показників для кожного об'єкта різні. Кінцева мета порівняння об'єктів – з'ясувати їх рейтинг серед розглянутої множини, причому, рейтинг прагнуть отримати у вигляді кількісної індивідуальної оцінки. Рішення завдання здійснюється віднизу доверху. Спочатку розглядають об'єкти, що знаходяться на найнижчому рівні ієрархії (експерти, альтернативи, критерії, чинники тощо), і попарно порівнюють один з одним. Порівнюючи пари об'єктів, експерт прагне встановити, наскільки один об'єкт кращий (гірший) за інший, що виражається визначенням кількісної оцінки. Переглянувши всі поєднання можливих пар

об'єктів і встановивши між ними оцінки взаємного впливу, експерт отримує матрицю парних порівнянь.

Метод парних порівнянь полягає в порівнянні досліджуваних об'єктів між собою. Об'єкти порівнюються попарно відносно їх впливу («ваги», або «інтенсивності») на загальну для них (вищу в ієрархії) характеристику.

Позначимо через  $A_1, A_2, \dots, A_n$  основні чинники, що визначають склад об'єкта. Тоді для визначення структури об'єкта заповнюється матриця парних порівнянь. Якщо позначити частку чинника  $A_i$  через  $W_j$  (оцінка, яку виставляє експерт, відповідно до вибраної шкали), то елемент матриці  $a_{ij} = W_j / W_i$  подано в табл. 9.2. У запропонованому варіанті застосування методу парних порівнянь визначаються не величини різниць значень чинників, а їх відношення  $a_{ij} = 1 / a_{ji}$ . Матриця парних порівнянь є позитивно спрямованою, назад- симетричною матрицею. Особливістю симетричних матриць парних порівнянь є:

- на головній діагоналі завжди має бути оцінка, що дорівнює 1;
- завжди має витримуватися співвідношення, яке відповідає умові: якщо під час порівняння  $i$ -го об'єкта з  $j$ -м об'єктом ставиться оцінка  $a_{ij}$ , то під час порівняння  $j$ -го об'єкта з  $i$ -м оцінка  $a_{ji}$  має бути протилежною.

Таблиця 9.2. – Матриця парних порівнянь

Загальне уявлення				
	$A_1$	$A_2$		$A_n$
$A_1$	1	$a_{12}$		$a_{1n}$
$A_2$	$a_{21}$	1		$a_{2n}$
$A_n$	$a_{n1}$	$a_{n2}$		1

Робота експертів полягає в тому, що, здійснюючи парне порівняння чинників  $A_1, A_n$ , необхідно заповнити таблицю парних порівнянь. Парні порівняння елементів можуть здійснюватися з використанням суб'єктивних суджень, чисельно оцінюються за шкалою, а потім визначається компонент  $W$ . Отже, рішення задачі полягає в знаходженні вектора  $(W_1, W_2, W_n)$ . Існує кілька різних способів пошуку подібного вектора пріоритетів. Одним з таких способів є перевірка на узгодженість.

Визначивши вектор пріоритетів, необхідно знайти головне власне значення матриці парних порівнянь  $\lambda_{max}$ . Воно використовується для цінування узгодженості експертних оцінок. Чим ближче воно до розмірності матриці, тим більше узгоджені оцінки. Якщо відомі точні значення порівнюваних об'єктів (наприклад, довжини, відстані тощо), то головне власне значення матриці

парних порівнянь дорівнює розмірності матриці. Відхилення (неузгодженість з ідеальним значенням) оцінок парних порівнянь від ідеального значення обчислюється як добуток нормалізованої оцінки для кожного об'єкта на сумарне значення оцінок для цього об'єкта, виставлених експертом (сума оцінок за стовпцем матриці парних порівнянь).

Під час парного порівняння об'єктів виставляється оцінка, яка показує величину – наскільки один об'єкт переважає над іншим. Експерт може помилитися, порівнюючи іншу пару об'єктів, що призведе до суперечливості результатів.

Для виявлення суперечливості результатів, які запропонував експерт під час заповнення матриці парних порівнянь, використовується кількісна оцінка – індекс узгодженості ( $I_U$ ). Якщо відхилення від узгодженості перевищують установлені межі, то необхідно скорегувати матрицю. Відхилення від узгодженості може бути виражено величиною, що дорівнює відношенню різниці  $\lambda_{\max}$  і  $n$  до  $n-1$ :

$$I_U = (\lambda_{\max} - n) / (n - 1) \quad (9.1)$$

Використовується дискретна шкала відносної важливості, що призводить до неузгодженості реальних оцінок з ідеальними оцінками. Для остаточного з'ясування узгодженості результатів парних порівнянь обчислюється кількісна оцінка відносної узгодженості ( $B_U$ ). Це відношення індексу узгодженості ( $I_U$ ) до середньостатистичного значення індексу узгодженості ( $I_U$ ) у разі випадкового вибору коефіцієнтів матриці порівнянь. Відносна узгодженість для системи в цілому характеризує зважене середнє значення відносної узгодженості за всіма матрицями порівнянь. Дані можна вважати майже несуперечливими (досить узгодженими), якщо значення відносної узгодженості менше, ніж 0,1. Як коригувальний коефіцієнт для остаточного з'ясування узгодженості оцінок у матриці парних порівнянь використовується середнє значення випадкового індексу ( $B_I$ ), згідно з табл. 9.3.

Таблиця 9.3. – Середнє значення випадкового індексу

Середнє значення випадкового індексу узгодженості										
Розмір матриці	1	2	3	4	5	6	7	8	9	10
Випадкова узгодженість ( $B_U$ )	0	0	0,58	0,90	1,12	1,24	1,32	1,41	1,45	1,49

Індекс узгодженості не залежить від шкал порівнянь, але залежить від кількості парних порівнянь. Індекс узгодженості – позитивне число. Чим менше протиріч у порівняннях, тим менше значення індексу узгодженості. З використанням методу порівнянь з еталоном значення індексу узгодженості дорівнює нулю. Ідеальним порівнянням відповідають нульовий індекс узгодженості і, відповідно, нульове значення відносної узгодженості.

**Приклад:** Обрати оптимальний варіант джерела електропостачання з врахуванням аспектів вартості, надійності та якості електричної енергії.

1. Формальна постановка задачі.

Для аналізу цієї проблеми пропонуємо три альтернативи:  $A_1$ ,  $A_2$  та  $A_3$ .

Запропоновано використати наступні критерії:  $K_1$  – вартість електричної енергії;  $K_2$  – надійність електричної енергії;  $K_3$  – якість електричної енергії.

2. Складаємо матрицю попарних порівнянь (див.9.4).

Таблиця 9.4. – Матриця попарних порівнянь альтернатив

$K_1$	$A_1$	$A_2$	$A_3$
$A_1$	1	1/5	4
$A_2$	5	1	9
$A_3$	1/4	1/9	1

Для кожного рядка матриці визначаємо добуток її членів та візьмемо корінь 3-го степеня з добутку:

$$a_1 = \sqrt[3]{1 \cdot 1/5 \cdot 4} = 0,9283$$

$$a_2 = \sqrt[3]{5 \cdot 1 \cdot 9} = 3,5568$$

$$a_3 = \sqrt[3]{1/4 \cdot 1/9 \cdot 1} = 0,3029$$

Далі визначаємо суму всіх координат:

$$a_{\Sigma} = a_1 + a_2 + a_3 = 0,9283 + 3,5568 + 0,3029 = 4,788.$$

Складемо нормований вектор, для цього розділимо кожену координату  $a_i$  на суму всіх координат:

$$S_1^{K1} = \frac{a_1}{a_{\Sigma}} = \frac{0,9283}{4,788} \approx 0,1939.$$

$$S_2^{K1} = \frac{a_2}{a_\Sigma} = \frac{3,5568}{4,788} \approx 0,7429;$$

$$S_3^{K1} = \frac{a_3}{a_\Sigma} = \frac{0,3029}{4,788} \approx 0,0633.$$

Вектор  $S_2(0,7429)$  є найкращим для критерію  $K_1$ .

3. Перевірка узгодженості експертних оцінок. Для кожного стовпця матриці визначимо суму його елементів, а потім визначимо суму попарних добутків:

$$b_1 = 1 + 5 + 1/4 = 6,25.$$

$$b_2 = 1/5 + 1 + 1/9 = 1,31.$$

$$b_3 = 4 + 9 + 1 = 14.$$

Отримаємо вектор  $b(6,25; 1,31; 14)$ .

Покоординатно перемножимо вектори та додаємо отримані добутки. Отримаємо максимальне власне число експериментальної матриці парних порівнянь ( $\lambda_{\max}$ ):

$$\begin{aligned} \lambda_{\max} &= b_1 \cdot S_1 + b_2 \cdot S_2 + b_3 \cdot S_3 = 0,1939 \cdot 6,25 + 0,7429 \cdot 1,31 + 0,0633 \cdot 14 = \\ &= 1,2119 + 0,9732 + 0,8862 = 3,0713. \end{aligned}$$

Визначаємо індекс узгодженості ( $IU$ ):

$$IU = \frac{\lambda_{\max} - n}{n - 1}.$$

$$IU = \frac{3,0713 - 3}{3 - 1} \approx 0,0357.$$

Тепер визначаємо відношення узгодженості ( $BU^*$ ):

$$BU^* = \frac{IU}{BU},$$

де  $BU = 0,58$  - відповідно до таблиці 9.3, а отже:

$$BU^* = \frac{0,0357}{0,58} \approx 0,0615.$$

Відношення узгодженості ( $BU$ ) перебуває в межах від 0 до 0,15, отже, складена експериментальна матриця парних порівнянь узгоджена і отриманим вектором пріоритетів  $S(0,1939; 0,7429; 0,0633)$  можна користуватися.

За аналогічним підходом отримуємо вагові коефіцієнти за другим  $S(0,1312; 0,7928; 0,076)$  і за третім  $S(0,1782; 0,7514; 0,0704)$  критеріями

4. Метод аналізу ієрархій.

Виконуємо порівняння на наступному рівні ієрархії (див.табл.9.5)



Таблиця 9.5. – Матриця порівняння критеріїв

$K$	$K_1$	$K_2$	$K_3$
$K_1$	1	1/4	5
$K_2$	4	1	7
$K_3$	1/5	1/7	1

Наступні кроки знаходження вагових коефіцієнтів важливості ті ж самі – $p$  (0,71471; 0,06679; 0,21849). Кожну матрицю обов’язково додатково перевіряємо на узгодженість.

#### 5. Розрахунок глобальних пріоритетів альтернатив.

Визначаємо глобальні пріоритети альтернатив:

$$R_1 = S_1^{K1} \cdot p_1 + S_1^{K2} \cdot p_2 + S_1^{K3} \cdot p_3 = 0,0473 + 0,0901 + 0,0123 = 0,1497.$$

$$R_2 = S_2^{K1} \cdot p_1 + S_2^{K2} \cdot p_2 + S_2^{K3} \cdot p_3 = 0,1810 + 0,5447 + 0,0049 = 0,7777.$$

$$R_3 = S_3^{K1} \cdot p_1 + S_3^{K2} \cdot p_2 + S_3^{K3} \cdot p_3 = 0,154 + 0,0522 + 0,0049 = 0,0725.$$

Отже, оптимальним варіантом є джерело електропостачання  $A_3$ .

#### Завдання для виконання практичної роботи:

1. Ознайомитись з теоретичними положеннями.
2. Отримати варіант завдання у викладача.
3. Визначити оптимальний варіант об’єкта за вашими пріоритетами згідно наданого завдання.

#### Контрольні запитання:

1. У чому відмінність/схожість методів парних порівнянь та аналізу ієрархій?
2. Розкажіть послідовність дій під час застосування методу парних порівнянь і методу аналізу ієрархій.
3. Для чого виконують перевірку на узгодженість матриці? У яких випадках імовірність похибки більша?

## ПЕРЕЛІК РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ

1. Навчальний посібник / А. В. Журахівський, А. Я. Яцейко, Н. Б. Дьяченко. Львів: Видавництво Львівської політехніки, 2012. - 380 с.
2. Навчальний посібник / Х. С. Соболев, Н. І. Петровська, О. М. Гуняк. Львів : Видавництво Львівської політехніки, 2018. - 92 с.
3. Основи наукових досліджень [Електронний ресурс]: навч. посіб. для студ. спеціальності 141 «Електроенергетика, електротехніка та електромеханіка» / КПІ ім. Ігоря Сікорського; уклад.: Г. Г. Стрелкова, М. М. Федосенко, А. І. Замулко, О. С. Іщенко. – Електронні текстові дані (1 файл: 500 Кбайт). – Київ: КПІ ім. Ігоря Сікорського, 2019. – 120 с.
4. Основи наукових досліджень: навч. посіб. / за заг. ред. Т.В. Гончарук. – Тернопіль, 2014. – 272 с. Детальніше за посиланням: <http://surl.li/tjhn>
5. Чмиленко Ф.О. Посібник до вивчення дисципліни «Методологія та організація наукових досліджень» / Ф.О. Чмиленко, Л.П. Жук. – Д.: РВВ ДНУ, 2014. – 48 с. Детальніше за посиланням: <http://surl.li/tjns>
6. Методика та організація наукових досліджень: Навчальний посібник / С. Е. Важинський, Т. І. Щербак. – Суми : Вид-во СумДПУ імені А. С. Макаренка, 2016. – 260 с.
7. Методологія та організація наукових досліджень: навчальний посібник / Б. І. Мокін, О. Б. Мокін. – 2-е вид., змін. та доп. – Вінниця : ВНТУ, 2015. – 317 с. Детальніше за посиланням: <http://surl.li/ijglv>
8. Грабченко А.І., Федорович В.О., Гаращенко Я.М. Методи наукових досліджень: Навч. посібник. – Х.: НТУ "ХПІ", 2009. – 142 с. Детальніше за посиланням: <http://surl.li/ijgme>
9. Основи методології та організації наукових досліджень: Навч. посіб. для студентів, курсантів, аспірантів і ад'юнтів / за ред. А. Є. Конверського. – К.: Центр учбової літератури, 2010. – 352 с.
10. Данильян О. Г. Організація та методологія наукових досліджень: навч. посіб. / О. Г. Данильян, О. П. Дзьобань. – Харків : Право, 2017. – 448 с.
11. Betz F. Managing Science: Methodology and Organization of Research. - New York, NY [u.a.] : Springer, 2011. – 388 p. DOI 10.1007/978-1-4419-7488-4
12. Valsiner J. From Methodology to Methods in Human Psychology. Springer, Cham, 2017. 115 p.
13. Бірта Г. О. Методологія і організація наукових досліджень / Г. О. Бірта, Ю.Г. Бургу. – К. : «Центр учбової літератури», 2014. – 142 с.
14. Методи планування та організації експериментальних досліджень: методичні вказівки до виконання практичних робіт для студентів спеціальності 141 "Електроенергетика, електротехніка та електромеханіка" / [уклад.: Савеленко

I.B., Петрова К.Г., Зінзура В.В.], Центральноукр. нац. техн. ун-т. – Кропивницький: ЦНТУ, 2017. – 42 с. . Детальніше за посиланням: <http://surl.li/fufam>

15. Методи планування та організації експериментальних досліджень: методичні вказівки для виконання самостійних робіт з дисципліни «Методи планування та організації експериментальних досліджень» для студентів, що навчаються за спеціальністю 141 «Електроенергетика, електротехніка та електромеханіка» / [уклад.: Савеленко I.B., Котиш А.I., Сіріков О.I.].- Кропивницький: ЦНТУ, 2017 – 48 с

16. Васильковський О.М., Лещенко С.М., Васильковська К.В., Петренко Д.I. Основи наукових досліджень. Перші наукові кроки. Навчальний посібник для студентів технічних спеціальностей. Харків. 2019. – 164 с. Детальніше за посиланням: <http://dspace.kntu.kr.ua/jspui/handle/123456789/10486>.

17. Методологія наукових досліджень : навч. посібник / В. Є. Юринець. – Львів: ЛНУ імені Івана Франка, 2011. – 178 с. Детальніше за посиланням: <http://surl.li/bigte>

18. Основи наукових досліджень та інженерної творчості // Навчальний посібник для студентів напрямів підготовки 144 «Теплоенергетика». – К.: КПІ ім. Ігоря Сікорського, 2016. – 270 с. Детальніше за посиланням: <http://surl.li/aszal>

19. Методичні вказівки до виконання кваліфікаційної роботи для здобувачів вищої освіти другого (магістерського) рівня зі спеціальності 141 «Електроенергетика, електротехніка та електромеханіка» за освітньо-професійними програмами «Електротехнічні системи електроспоживання» та «Енергетичний менеджмент» / [П. Г. Плешков, Н. Ю. Гарасьова, А. І. Котиш] ; М-во освіти і науки України, Центральноукраїн. нац. техн. ун-т. – Кропивницький: ЦНТУ, 2017 - 48 с.

20. Кушлик-Дивульська О.I., Кушлик Б.Р. Основи теорії прийняття рішень. – К., 2014. – 94 с.

21. Методичні вказівки щодо виконання лаб. робіт з навчальної дисципліни «Організація та планування наукового експерименту» для студентів денної та заочної форм навчання за ОП «Енергетичний менеджмент» зі спеціальності 141 – «Електроенергетика, електротехніка та електромеханіка» другого (магістерського) освітнього рівня / [уклад. : Якимець С. М., Усатюк О. О.] ; М-во освіти і науки України, Кременчуц. нац. ун-т. ім. М. Остроградського - Кременчук, 2018. - 51 с.

**ДОДАТОК. Титульний аркуш звіту з практичної роботи**  
Центральноукраїнський національний технічний університет  
Кафедра «Електротехнічні системи та енергетичний менеджмент»

**ЗВІТ**

з практичної роботи № 9

з курсу: *«Методи планування та організації наукових  
досліджень в енергетиці»*

на тему: *«Використання методу парних порівнянь та аналізу ієрархій»*

Виконав здобувач вищої освіти  
групи ЕЕ-2\_ М спеціальності 141 «Електроенергетика,  
електротехніка та електромеханіка» ОПП  
«Електротехнічні системи  
електроспоживання»/«Енергетичний менеджмент»

---

Перевірила  
к.т.н., доц. К. Петрова

Національна шкала \_\_\_\_\_  
Кількість балів: \_\_\_\_\_

м. Кропивницький – 20\_\_ рік