Чернівецький національний університет імені Юрія Федьковича

Інститут фізико-технічних та комп’ютерних наук

Кафедра Математичних проблем управління і кібернетики

Лабораторна робота № 11

“Тема: Двофакторний дисперсійний аналіз”

з дисципліни

“Інтелектуальний аналіз”

Варіант - 2

Виконав:

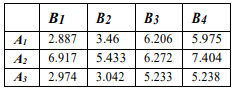
ст. гр. 341СК Гаваньо Дмитро

Прийняв:

професор Дрінь Я. М.

Чернівці – 2025

**Мета:** Ознайомитися з методикою проведення двофакторного дисперсійного аналізу



**Завдання 11.1:** Проведіть двофакторний дисперсійний аналіз даних із таблиці.

1. Визначте матрицю спостережень і введіть її елементи; визначте і введіть число рівнів кожного фактору й обчисліть обсяг вибірки.
2. Обчисліть групові і генеральну середні.
3. Обчислите оцінки дисперсії.
4. Визначте для першого фактору величину 𝐹𝐴.
5. Знайдіть для першого фактору величину 𝑞𝐹(𝑙 − 𝛼, 𝑘,𝑙) (розв’язок рівняння 𝐹(𝑥) = 1 − 𝛼 для розподілу Фішера 𝐹 з відповідним числомступенів вільності).
6. Порівняйте значення 𝐹𝐴 зі знайденим значенням 𝑥𝛼 і сформулюйте висновок про справедливість гіпотези 𝐻𝐴.
7. Визначте для другого фактору величину 𝐹𝐵.
8. Знайдіть для другого фактору величину 𝑞𝐹(𝑙 − 𝛼, 𝑘,𝑙) (розв’язок рівняння 𝐹(𝑥) = 1 − 𝛼 для розподілу Фішера 𝐹 з відповідним числом ступенів вільності).
9. Якщо гіпотеза 𝐻𝐴 відхиляється, обчисліть коефіцієнт детермінації для першого фактору.
10. Порівняйте значення 𝐹𝐵 зі значенням 𝑞𝐹(𝑙 − 𝛼, 𝑘, 𝑙) і сформулюйте висновок про справедливість гіпотези 𝐻𝐵 .
11. Якщо гіпотеза 𝐻𝐵 відхиляється, обчисліть коефіцієнт детермінації для другого фактору.
12. Обчисліть оцінки параметрів розподілу випадкової величини, яка досліджується.

**Хід роботи:**

**Завдання 1:**

Програмний код:

setwd("D:/Study/R/lab11")

data <- read.csv("selection.csv", header = FALSE)

data\_matrix <- as.matrix(data)

alpha <- 0.05

m\_a <- nrow(data\_matrix)

m\_b <- ncol(data\_matrix)

n <- m\_a \* m\_b

row\_dot\_means <- rowMeans(data\_matrix)

col\_dot\_means <- colMeans(data\_matrix)

global\_mean <- mean(data\_matrix)

cat("Середні значення:\n")

cat("  Середнє по рядках:", row\_dot\_means, "\n")

cat("  Середнє по стовпцях:", col\_dot\_means, "\n")

cat("  Загальне середнє:", global\_mean, "\n")

dx\_a <- (m\_b / n) \* sum((row\_dot\_means - global\_mean)^2)

dx\_b <- (m\_a / n) \* sum((col\_dot\_means - global\_mean)^2)

dx\_0 <- sum((data\_matrix - global\_mean)^2) / n

dx <- dx\_0 + dx\_a + dx\_b

cat("Суми квадратів:\n")

cat("  Сума квадратів фактору A:", dx\_a, "\n")

cat("  Сума квадратів фактору B:", dx\_b, "\n")

cat("  Сума квадратів помилки:", dx\_0, "\n")

cat("  Загальна сума квадратів:", dx, "\n")

F\_a <- dx\_a / dx\_0

F\_critical\_a <- qf(1 - alpha, m\_a - 1, (m\_a - 1) \* (m\_b - 1))

cat("--- Гіпотеза про вплив фактору A за рівнем значущості", alpha,"---\n")

if (F\_a > F\_critical\_a) {

  cat("  --- Відхиляємо  гіпотезу", F\_a, ">", F\_critical\_a, "\n")

  det\_coef\_a <- (dx\_a) / (dx\_a + dx\_0)

  cat("  --- Коефіцієнт детермінації для фактору A:", det\_coef\_a, "\n")

} else {

  cat("  --- Не відхиляємо гіпотезу", F\_a, "<=", F\_critical\_a, "\n")

}

F\_b <- dx\_b / dx\_0

F\_critical\_b <- qf(1 - alpha, m\_b - 1, (m\_a - 1) \* (m\_b - 1))

cat("--- Гіпотеза про вплив фактору B за рівнем значущості", alpha,"---\n")

if (F\_b > F\_critical\_b) {

  cat("  --- Відхиляємо  гіпотезу", F\_b, ">", F\_critical\_b, "\n")

  det\_coef\_b <- (dx\_b) / (dx\_b + dx\_0)

  cat("  --- Коефіцієнт детермінації для фактору B:", det\_coef\_b, "\n")

} else {

  cat("  --- Не відхиляємо гіпотезу", F\_b, "<=", F\_critical\_b, "\n")

}

sigma\_sq <- dx\_0 / (n - m\_a - m\_b + 1)

cat("Параметри моделі:\n")

cat("  Незсунена дисперсія:", sigma\_sq , "\n")

cat("  Середнє значення:", global\_mean, "\n")

cat("  Вплив фактору A:", row\_dot\_means - global\_mean, "\n")

cat("  Вплив фактору B:", col\_dot\_means - global\_mean, "\n")

Результат виконання:

Середні значення:

Середнє по рядках: 4.632 6.5065 4.12175

Середнє по стовпцях: 4.259333 3.978333 5.903667 6.205667

Загальне середнє: 5.08675

Суми квадратів:

Сума квадратів фактору A: 1.051238

Сума квадратів фактору B: 0.9581333

Сума квадратів помилки: 2.371673

Загальна сума квадратів: 4.381044

--- Гіпотеза про вплив фактору A за рівнем значущості 0.05 ---

--- Не відхиляємо гіпотезу 0.4432473 <= 5.143253

--- Гіпотеза про вплив фактору B за рівнем значущості 0.05 ---

--- Не відхиляємо гіпотезу 0.4039905 <= 4.757063

Параметри моделі:

Незсунена дисперсія: 0.3952788

Середнє значення: 5.08675

Вплив фактору A: -0.45475 1.41975 -0.965

Вплив фактору B: -0.8274167 -1.108417 0.8169167 1.118917

**Висновок:** яознайомився з методикою проведення двофакторного дисперсійного аналізу