

Міністерство освіти і науки України  
Національний технічний університет України  
«Київський політехнічний інститут ім. Ігоря Сікорського»  
Факультет інформатики та обчислювальної техніки  
Кафедра обчислювальної техніки

## **ЛАБОРАТОРНА РОБОТА № 2**

з дисципліни «Методи оптимізації та планування експерименту» на тему  
«Проведення двофакторного експерименту з використанням лінійного  
рівняння регресії»

Виконав:  
студент II курсу ФІОТ  
групи ІО-93  
Бриль Владислав  
Залікова – 9303  
Номер у списку: 2

ПЕРЕВІРИВ:  
Асистент Регіда П. Г.

**Мета:** провести двофакторний експеримент, перевірити однорідність дисперсії за критерієм Романовського, отримати коефіцієнти рівняння регресії, провести натуралізацію рівняння регресії.

Введемо такі позначення:

N – кількість точок плану (рядків матриці планування)

k – кількість факторів(кількість x)

m – кількість дослідів у за однієї і тієї ж комбінації факторів (test)

x - нормовані значення факторів( 1, )

### Завдання на лабораторну роботу:

1. Записати лінійне рівняння регресії.
2. Обрати тип двофакторного експерименту і скласти матрицю планування для нього з використанням додаткового нульового фактору (x<sub>0</sub>=1).
3. Провести експеримент в усіх точках повного факторного простору (знайти значення функції відгуку у). Значення функції відгуку задати випадковим чином у відповідності до варіанту у діапазоні у<sub>min</sub> ÷ у<sub>max</sub>

у<sub>max</sub> = (30 - Nваріанту)\*10,

у<sub>min</sub> = (20 - Nваріанту)\*10.

### Варіант:

302	20	70	-15	45
-----	----	----	-----	----

### Програмний код:

```
import random
import math
n_variant = 302
m = 6
y_max = (30 - n_variant) * 10
y_min = (20 - n_variant) * 10

x1_min, x1_max, x2_min, x2_max = 20, 70, -15, 45

xn = [[-1, -1], [-1, 1], [1, -1]]

# середнє значення y
def average_value_y(c_list):
    average_l_y = []
    for i in range(len(c_list)):
        sum_el = 0
        for j in c_list[i]:
            sum_el += j
        average_l_y.append(sum_el / len(c_list[i]))
    return average_l_y
```

```

# дисперсія
def dispersion(c_list):
    d = []
    for i in range(len(c_list)):
        sum_of_y = 0
        for k in c_list[i]:
            sum_of_y += (k - average_value_y(c_list)[i]) ** 2
        d.append(sum_of_y / len(c_list[i]))
    return d

# перевірка для кожної пари комбінації
def f_uv(u, v):
    if u >= v:
        return u / v
    else:
        return v / u

# визначник
def determ(x11, x12, x13, x21, x22, x23, x31, x32, x33):
    det = x11 * x22 * x33 + x12 * x23 * x31 + x32 * x21 * x13 - x13 * x22 *
x31 - x32 * x23 * x11 - x12 * x21 * x33
    return det

# генерація y
y = [[random.randint(y_min, y_max) for j in range(6)] for i in range(3)]
average_y = average_value_y(y)

# основне відхилення
sigma_tet = math.sqrt((2 * (2 * m - 2)) / (m * (m - 4)))

fuv = []
teta = []
ruv = []

fuv.append(f_uv(dispersion(y)[0], dispersion(y)[1]))
fuv.append(f_uv(dispersion(y)[2], dispersion(y)[0]))
fuv.append(f_uv(dispersion(y)[2], dispersion(y)[1]))

teta.append(((m - 2) / m) * fuv[0])
teta.append(((m - 2) / m) * fuv[1])
teta.append(((m - 2) / m) * fuv[2])

# експериментальне значення критерію Романовського
ruv.append(abs(teta[0] - 1) / sigma_tet)
ruv.append(abs(teta[1] - 1) / sigma_tet)
ruv.append(abs(teta[2] - 1) / sigma_tet)
# візьмемо значення з таблиці при m=6
r_kr = 2

# перевірка гіпотези про однорідність дисперсій
for i in range(len(ruv)):
    if ruv[i] > r_kr:
        print("Дисперсія - неоднорідна, повторіть експеримент")

# позначення для системи рівнянь з коефіцієнтами для лінійної регресії
mx1 = (xn[0][0] + xn[1][0] + xn[2][0]) / 3
mx2 = (xn[0][1] + xn[1][1] + xn[2][1]) / 3
my = (average_y[0] + average_y[1] + average_y[2]) / 3
# коефіцієнти регресії
a1 = (xn[0][0] ** 2 + xn[1][0] ** 2 + xn[2][0] ** 2) / 3
a2 = (xn[0][0] * xn[0][1] + xn[1][0] * xn[1][1] + xn[2][0] * xn[2][1]) / 3
a3 = (xn[0][1] ** 2 + xn[1][1] ** 2 + xn[2][1] ** 2) / 3
a11 = (xn[0][0] * average_y[0] + xn[1][0] * average_y[1] + xn[2][0] *
average_y[2]) / 3

```

```

a22 = (xn[0][1] * average_y[0] + xn[1][1] * average_y[1] + xn[2][1] *
average_y[2]) / 3

# рішення системи з коефіцієнтами регресії методом Крамера
b0 = determ(my, mx1, mx2, a11, a1, a2, a22, a2, a3) / determ(1, mx1, mx2,
mx1, a1, a2, mx2, a2, a3)
b1 = determ(1, my, mx2, mx1, a11, a2, mx2, a22, a3) / determ(1, mx1, mx2,
mx1, a1, a2, mx2, a2, a3)
b2 = determ(1, mx1, my, mx1, a1, a11, mx2, a2, a22) / determ(1, mx1, mx2,
mx1, a1, a2, mx2, a2, a3)
# лінійна регресія (практичне)
y_pr1 = b0 + b1 * xn[0][0] + b2 * xn[0][1]
y_pr2 = b0 + b1 * xn[1][0] + b2 * xn[1][1]
y_pr3 = b0 + b1 * xn[2][0] + b2 * xn[2][1]

# натуралізація плану
dx1 = abs(x1_max - x1_min) / 2
dx2 = abs(x2_max - x2_min) / 2
x10 = (x1_max + x1_min) / 2
x20 = (x2_max + x2_min) / 2
# обчислення натуралізованих коефіцієнтів
a_0 = b0 - (b1 * x10 / dx1) - (b2 * x20 / dx2)
a_1 = b1 / dx1
a_2 = b2 / dx2
# натуралізоване рівняння регресії
yP1 = a_0 + a_1 * x1_min + a_2 * x2_min
yP2 = a_0 + a_1 * x1_max + a_2 * x2_min
yP3 = a_0 + a_1 * x1_min + a_2 * x2_max

print('Матриця планування для m =', m)
for i in range(3):
    print(y[i])

print('Експериментальні значення критерію Романовського:')
for i in range(3):
    print(ruv[i])

print('Натуралізовані коефіцієнти: \na0 =', round(a_0, 5), 'a1 =', round(a_1,
5), 'a2 =', round(a_2, 5))
print('У практичний ', round(y_pr1, 5), round(y_pr2, 5), round(y_pr3, 5),
'\nУ середній', round(average_y[0], 5), round(average_y[1], 5),
round(average_y[2], 5))
print('У практичний нормалізований', round(yP1, 5), round(yP2, 5), round(yP3,
5))

```

## Вивід програми:

```

C:\Users\Владислав\AppData\Local\Programs\Python\Python38\python.exe C:/Users/Владислав/PycharmProjects/MOPE_LAB2/MOPE_Lab2.py
Матриця планування для m = 6
[-2776, -2732, -2779, -2813, -2781, -2801]
[-2802, -2765, -2781, -2758, -2768, -2762]
[-2743, -2754, -2729, -2803, -2772, -2784]
Експериментальні значення критерію Романовського:
0.7103361192795219
0.24436250400909953
0.6715870076524593
Натуралізовані коефіцієнти:
a0 = -2784.88333 a1 = 0.32333 a2 = 0.12778
У практичний -2780.33333 -2772.66667 -2764.16667
У середній -2780.33333 -2772.66667 -2764.16667
У практичний нормалізований -2780.33333 -2764.16667 -2772.66667

Process finished with exit code 0

```

## Контрольні питання:

1)

В теорії планування експерименту найважливішою частиною є оцінка результатів вимірів. При цьому використовують апроксимуючі поліноми, за допомогою яких ми можемо описати нашу функцію. В ТПЕ ці поліноми отримали спеціальну назву - регресійні поліноми, а їх знаходження та аналіз - регресійний аналіз. Найчастіше в якості базисної функції використовується ряд Тейлора, який має скінченну кількість членів.

$$F(x) = F(a) + \frac{x-a}{1!} F'(a) + \frac{(x-a)^2}{2!} F''(a) + \dots + \frac{(x-a)^N}{N!} F^{(N)}(a)$$

Але при використанні апроксимуючого полінома Тейлора в його початковому вигляді виникає ряд проблем, пов'язаних із знаходженням похідних, оскільки нам невідома функція, а відомий лише ряд її значень. Тому ми замінюємо поліном Тейлора аналогічним йому рівнянням регресії:

$$\hat{y} = b_0 + \sum_{i=1}^k b_i x_i + \sum_{i,j=1}^k b_{i,j} x_i x_j + \sum_{i=1}^k b_{i,i} x_i^2 + \sum_{i,j,n=1}^k b_{i,j,n} x_i x_j x_n + \dots$$

де  $k$  – кількість факторів (кількість  $x$ )

Мета даної роботи – дослідити лінійну регресійну модель

$$\hat{y} = b_0 + \sum_{i=1}^k b_i x_i$$

2)

Дисперсія – це сума квадратів відхилень величин  $y_{jg}$  від середнього значення  $\bar{y}_j$ .

Дисперсія обчислюється для кожного рядка за формулою:

$$\sigma^2 \{y_j\} = \frac{1}{m} \sum_{g=1}^m (y_{jg} - \bar{y}_j)^2, (j = \overline{1, N}, g = \overline{1, m}), \bar{y}_j = \frac{1}{m} \sum_{g=1}^m y_{jg}$$

Якщо  $y_{jg}$  – нормально розподілена величина, і кількості дослідів  $m$  є достатньою, то дисперсії розподілів  $y_{jg}$  для кожної комбінації повинні бути **рівними**.

Тобто, для будь-яких  $i$  та  $j$  ( $i = \overline{1, N}; j = \overline{1, N}$ ):  $\sigma^2 \{y_i\} = \sigma^2 \{y_j\}$ ;

Груба похибка (промах) – це похибка результату окремого виміру, що входить в ряд вимірів, котра для даних умов різко відрізняється від інших результатів цього ряду. Причинами появи грубих похибок є різкі зміни умов вимірювання і помилки, допущені оператором. Наявність грубих похибок говорить про неоднорідності дисперсії.

Якщо хоча б для одної пари  $u, v$  має місце  $R_{uv} > R_{кр}$ , то гіпотеза про однорідність дисперсій не підтверджується. В цьому випадку розбіжність між дисперсіями експериментальних значень  $u$ -ї та  $v$ -ї комбінацій є значною. Необхідно збільшити кількість дослідів  $m=m+1$ , провести нові досліді і заново перевірити критерій.

3) ПФЕ – повний факторний експеримент, - це коли використовуються усі можливі комбінації рівнів факторів; при ПФЕ кількість комбінацій  $Nn=rk$ .