

НАВЧАЛЬНО-НАУКОВИЙ КОМПЛЕКС
"ІНСТИТУТ ПРИКЛАДНОГО СИСТЕМНОГО АНАЛІЗУ"
НАЦІОНАЛЬНОГО ТЕХНІЧНОГО УНІВЕРСИТЕТУ УКРАЇНИ
"КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ
СІКОРСЬКОГО"
КАФЕДРА МАТЕМАТИЧНИХ МЕТОДІВ СИСТЕМНОГО АНАЛІЗУ

РОЗРАХУНКОВА РОБОТА
з теми "Регресійний аналіз"

Виконав студент групи
КА-81 Фордуй Нікіта
Перевірила Каніовська І.Ю.

Київ 2020

Зміст

1	Завдання 1	2
1.1	Завдання	2
1.2	Аналіз вибірки та вибір регресійної моделі	3
1.3	Знаходження за МНК оцінки параметрів вибраної моделі .	3
1.4	Перевірка адекватності побудованої моделі	4
1.5	Перевірка гіпотези про значущість найменшого за значенням параметра	4
1.6	Довірчий інтервал для середнього значення відклику та значення відклику в точці	5
1.7	Висновки	6
2	Завдання 2	7
2.1	Завдання	7
2.2	Знаходження за МНК оцінки параметрів вибраної моделі .	8
2.3	Перевірка адекватності побудованої моделі	9
2.4	Перевірка гіпотези про значущість найменшого за значенням параметра	9
2.5	Довірчий інтервал для середнього значення відклику та значення відклику в точці	9
2.6	Висновки	11
3	Завдання 3	12
3.1	Завдання	12
3.2	Знаходження за МНК оцінки параметрів вибраної моделі .	14
3.3	Перевірка адекватності побудованої моделі	14
3.4	Перевірка гіпотези про значущість найменшого за значенням параметра	15
3.5	Довірчий інтервал для середнього значення відклику та значення відклику в точці	15
3.6	Висновки	16

1 Завдання 1

1.1 Завдання

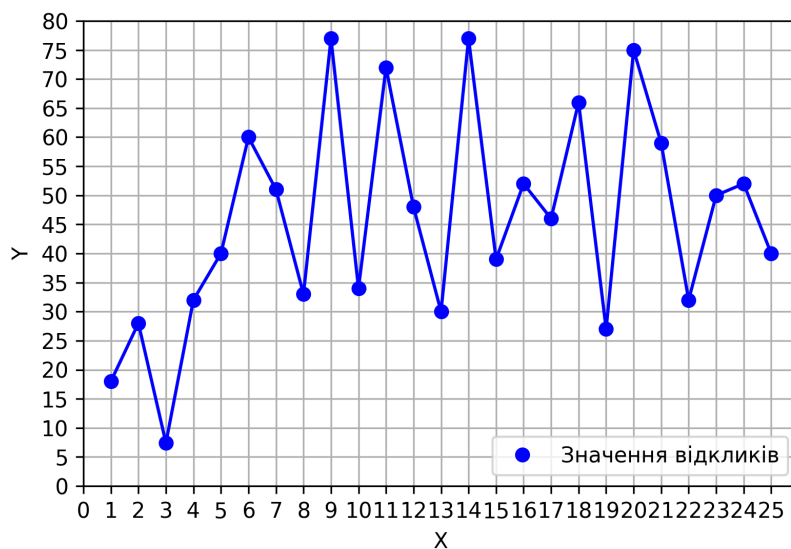
1. Провести аналіз вибірки та вибрати підходящу лінійну регресійну модель.
2. За методом найменших квадратів знайти оцінки параметрів вибраної моделі.
3. На рівні значущості $\alpha = 0,05$ перевірити адекватність побудованої моделі.
4. Для самого малого значення параметра побудованої моделі на рівні значущості $\alpha = 0,05$ перевірити гіпотезу про його значущість.
5. Побудувати прогнозований довірчий інтервал з довірчою ймовірністю $\gamma = 0,95$ для середнього значення відклику та самого значення відклику в точці $x = 16$.
6. Написати висновки.

X	Y
1	18
2	28
3	7.4
4	32
5	40
6	60
7	51
8	33
9	77
10	34
11	72
12	48
13	30
14	77
15	39
16	52
17	46

X	Y
18	66
19	27
20	75
21	59
22	32
23	50
24	52
25	40

1.2 Аналіз вибірки та вибір регресійної моделі

Розглянемо розташування значень відкликів Y на площині:



За розташуванням значень відкликів на площині стає зрозуміло, що залежність є нелінійна і схожа на квадратичну. Тому розглянемо модель $f(x) = \beta_0 + \beta_1 x + \beta_2 x^2$

1.3 Знаходження за МНК оцінки параметрів вибраної моделі

Матриця плану для вибраної моделі матиме вигляд:

$$F = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & 25 \\ 1^2 & 2^2 & \dots & 25^2 \end{pmatrix}^T \quad (1)$$

Після цього використовуємо матрицю F для знаходження інформаційної матриці Фішера $A = FF^T$ та дисперсійну матрицю Фішера A^{-1} ; обидві з них мають розмір 3×3 . Обчислені матриці наведені у додатку 1.

Знайдемо значення оцінок параметрів регресійної моделі за формулою:

$$\vec{\beta}_{\text{знач.}}^* = A^{-1} F^T \vec{\eta}_{\text{знач.}}, \quad (2)$$

де $\vec{\eta}_{\text{знач.}} = (18, 28, \dots, 40)^T$ - вектор значень відкликів. Отримаємо $\vec{\beta}_{\text{знач.}}^* \approx (14.87, 5.24, -0.17)^T$. Тоді сама модель матиме вигляд:

$$f^*(x) = 14.87 + 5.24x - 0.17x^2 \quad (3)$$

1.4 Перевірка адекватності побудованої моделі

Перевіряємо на рівні значущості $\alpha = 0.05$ адекватність побудованої моделі. Задля перевірки побудованої регресійної моделі скористаємось тим факт, що випадкова величина

$$\gamma = \frac{(\sigma^2)^{**}}{D_{\eta}^{**}} \sim F(n - m, n - 1) \quad (4)$$

У нашому випадку $n = 25, m = 3 \Rightarrow$ статистика γ буде розподілена за законом Фішера-Снедекора із 22 та 24 ступенями вільності відповідно. Критична область при саме такому виборі статистики критерію буде правостороння, тому знайдемо межу критичної області з виразу $\mathbb{P}(\gamma > t) = \alpha$, $t_{\text{кр.}} = 2.03$; $\gamma_{\text{знач.}} \approx 0.01 < t_{\text{кр.}}$. Таким чином гіпотеза про адекватність побудованої моделі на рівні значущості 0.05 не протирічить дослідним даним.

1.5 Перевірка гіпотези про значущість найменшого за значенням параметра

Перевіримо гіпотезу про значущість параметра для $(\beta_2^*)_{\text{знач.}} = -0.17$ на рівні значущості $\alpha = 0.05$. Основна гіпотеза $H_0 : \beta_2 = 0$ проти лівосторонньої альтернативи $H_1 : \beta_2 < 0$, відповідно, критична область - лівостороння. Для перевірки гіпотези скористаємося статистикою:

$$\gamma = \frac{\beta_2^*}{\sqrt{(\sigma^2)^{**} a_{22}}} \sim St_{n-m} \quad (5)$$

З таблиці розподілу Стюдента з $n - m = 22$ ступенями вільності отримуємо критичне значення $t_{\text{кр.}} = 1.717$. Знаходимо $(\sigma^2)^{**} \approx 3.48$, $a_{22} = (A^{-1})_{22} \approx 1.858 \cdot 10^{-5}$. Отримуємо $\gamma_{\text{знач.}} \approx -20.933$. Т.я. $\gamma_{\text{знач.}} < t_{\text{кр.}}$, то основна гіпотеза H_0 відхиляється на користь альтернативної. Таким чином, параметр β_2 є значущим, залишаємо побудовану модель незмінною.

1.6 Довірчий інтервал для середнього значення відклику та значення відклику в точці

Будуватимемо довірчий інтервал для середнього значення відклику та самого значення відклику в точці $x = 26$.

Спочатку побудуємо прогнозований довірчий інтервал для середнього значення відклику з довірчою ймовірністю $\gamma = 0.95$. Для цього скористаємось статистикою :

$$\zeta = \frac{f^*(x) - f(x)}{\sqrt{(\sigma^2)^{**}(\vec{x})^T A^{-1} \vec{x}}} \sim St_{n-m} = St_{22} \quad (6)$$

Довірчий інтервал для середнього значення відклику матиме вигляд :

$$\left(f^*(x) - t\sqrt{(\sigma^2)^{**}(x)^T A^{-1} \vec{x}}, f^*(x) + t\sqrt{(\sigma^2)^{**}(\vec{x})^T A^{-1} \vec{x}} \right) \quad (7)$$

Величина t знаходиться з рівняння $\mathbb{P}(|\zeta| < t) = 0.95$; $t_{\text{зн.}} = 2.074$; $(\vec{x})^T A^{-1} \vec{x} \approx 0.42$.

Таким чином, отримаємо довірчий інтервал для середнього значення відклику в точці $x = 26$ з надійністю $\gamma = 0.95$:

$$(34.85, 39.89) \quad (8)$$

Тепер побудуємо довірчий інтервал для самого значення відклику в точці $x = 26$ з надійністю $\gamma = 0.95$. Для цього скористаємось статистикою:

$$\zeta = \frac{\eta - f^*(x)}{\sqrt{(\sigma^2)^{**}(1 + (\vec{x})^T A^{-1} \vec{x})}} \quad (9)$$

Довірчий інтервал матиме вигляд:

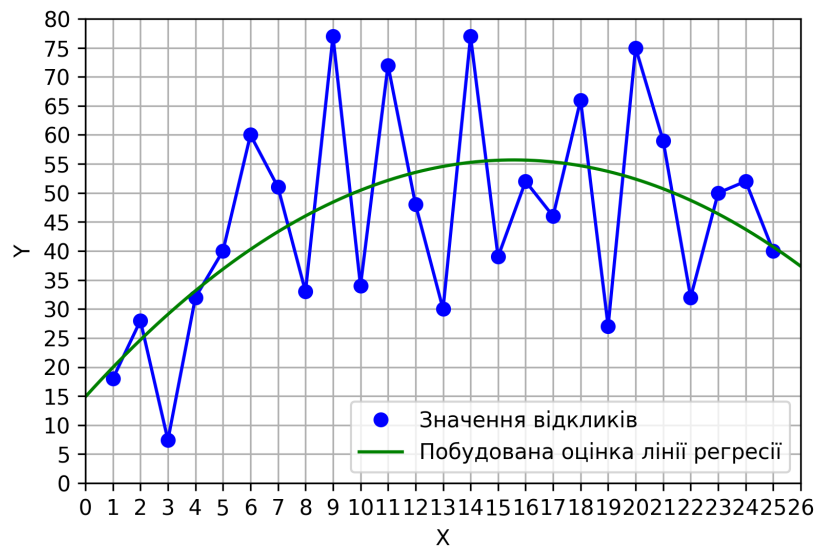
$$\left(f^*(x) - t\sqrt{(\sigma^2)^{**}(1 + (\vec{x})^T A^{-1} \vec{x})}, f^*(x) + t\sqrt{(\sigma^2)^{**}(1 + (\vec{x})^T A^{-1} \vec{x})} \right) \quad (10)$$

де t знаходимо з рівняння $\mathbb{P}(|\zeta| < t) = 0.95$. Всі необхідні нам значення були обчислені. Після усіх розрахунків отримаємо довірчий інтервал:

$$(32.753, 41.989) \quad (11)$$

1.7 Висновки

В ході цієї роботи було побудовано лінійну регресійну модель вигляду $f(x) = \beta_0 + \beta_1 x + \beta_2 x^2$. За допомогою методу найменших квадратів було знайдено оцінки параметрів цієї моделі. Була перевірена адекватність побудованої моделі на рівні значущості 0.05 та на тому ж самому рівні значущості була перевірена гіпотеза про значущість параметру з найменшим значенням його оцінки. Так як обидві гіпотези не протирічали дослідним даним, то модель залишилася незмінною. В подальшому був побудований довірчий інтервал для значення відклику та середнього значення відклику для значення фактору $x = 26$. Обидва інтервали були побудовані з довірчою ймовірністю 0.95. Необхідні обчислення були виконані за допомогою бібліотек Pandas, Matplotlib та Numpy мови програмування Python. Нижче на графіку зображена побудована лінія регресії:



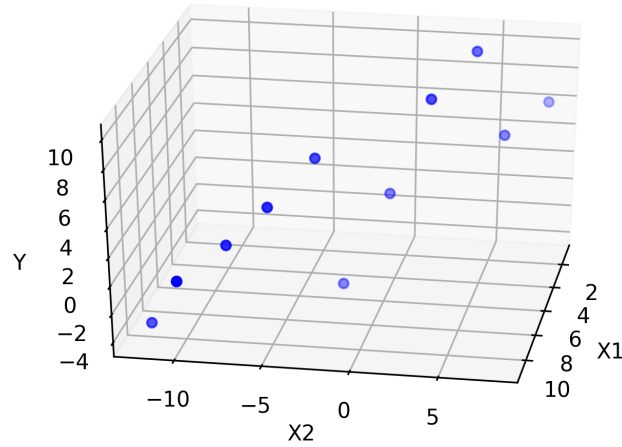
2 Завдання 2

2.1 Завдання

1. За методом найменших квадратів знайти оцінки параметрів двофакторної регресійної моделі.
2. На рівні значущості $\alpha = 0.05$ перевірити адекватність побудованої моделі.
3. Для самого малого значення параметра побудованої моделі на рівні значущості $\alpha = 0.05$ перевірити гіпотезу про його значущість.
4. Побудувати прогнозований довірчий інтервал з довірчою ймовірністю $\gamma = 0.95$ для середнього значення відклику та самого значення відклику в деякій точці.
5. Написати висновки.

№	X_1	X_2	Y
1	1	8	6
2	4	2	8
3	9	-8	1
4	11	-10	0
5	3	6	5
6	8	-6	3
7	5	0	2
8	10	-12	-4
9	2	4	10
10	7	-2	-3
11	6	-4	5

На малюнку нижче зображені значення відкликів у просторі.



2.2 Знаходження за МНК оцінки параметрів вибраної моделі

Матриця плану для вибраної моделі матиме вигляд:

$$F = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 4 & 9 & \dots & 6 \\ 8 & 2 & -8 & \dots & -4 \end{pmatrix}^T \quad (12)$$

Після цього використовуємо матрицю F для знаходження інформаційної матриці Фішера $A = FF^T$ та дисперсійну матрицю Фішера A^{-1} ; обидві з них мають розмір 3×3 . Обчислені матриці наведені у додатку 1. Знайдемо значення оцінок параметрів регресійної моделі за формулою:

$$\vec{\beta}_{\text{знач.}}^* = A^{-1}F^T\vec{\eta}_{\text{знач.}}, \quad (13)$$

де $\vec{\eta}_{\text{знач.}} = (6, 8, \dots, 5)^T$ - вектор значень відкликів. Отримаємо $\vec{\beta}_{\text{знач.}}^* \approx (14, -2, -0.5)^T$. Тоді сама модель матиме вигляд:

$$f^*(x) = 14 - 2x_1 - 0.5x_2 \quad (14)$$

2.3 Перевірка адекватності побудованої моделі

Перевіряємо на рівні значущості $\alpha = 0.05$ адекватність побудованої моделі. Задля перевірки побудованої регресійної моделі скористаємось тим факт, що випадкова величина

$$\gamma = \frac{(\sigma^2)^{**}}{D_{\eta}^{**}} \sim F(n - m, n - 1) \quad (15)$$

У нашому випадку $n = 11, m = 3 \Rightarrow$ статистика γ буде розподілена за законом Фішера-Снедекора із 8 та 10 ступенями вільності відповідно. Критична область при саме такому виборі статистики критерію буде правостороння, тому знайдемо межу критичної області з виразу $\mathbb{P}(\gamma > t) = \alpha$, $t_{\text{кр.}} = 3.07$; $\gamma_{\text{знач.}} \approx 0.021 < t_{\text{кр.}}$. Таким чином гіпотеза про адекватність побудованої моделі на рівні значущості 0.05 не протирічить дослідним даним.

2.4 Перевірка гіпотези про значущість найменшого за значенням параметра

Перевіримо гіпотезу про значущість параметра для $(\beta_2^*)_{\text{знач.}} = -0.5$ на рівні значущості $\alpha = 0.05$. Основна гіпотеза $H_0 : \beta_2 = 0$ проти лівосторонньої альтернативи $H_1 : \beta_2 < 0$, відповідно, критична область - лівостороння. Для перевірки гіпотези скористаємося статистикою:

$$\gamma = \frac{\beta_2^*}{\sqrt{(\sigma^2)^{**} a_{22}}} \sim St_{n-m} \quad (16)$$

З таблиці розподілу Стьюдента з $n - m = 8$ ступенями вільності отримуємо критичне значення $t_{\text{кр.}} = 1.86$. Знаходимо $(\sigma^2)^{**} \approx 0.374$, $a_{22} = (A^{-1})_{22} \approx 0.042$. Отримуємо $\gamma_{\text{знач.}} \approx -3.973$. Т.я. $\gamma_{\text{знач.}} < t_{\text{кр.}}$, то основна гіпотеза H_0 відхиляється на користь альтернативної. Таким чином, параметр β_2 є значущим, залишаємо побудовану модель незмінною.

2.5 Довірчий інтервал для середнього значення відклику та значення відклику в точці

Будуватимемо довірчий інтервал для середнього значення відклику та самого значення відклику в точці $(10, -9)$.

Спочатку побудуємо прогнозований довірчий інтервал для середнього значення відклику з довірчою ймовірністю $\gamma = 0.95$. Для цього скористаємось статистикою :

$$\zeta = \frac{f^*(\vec{x}) - f(\vec{x})}{\sqrt{(\sigma^2)^{**}(\vec{x})^T A^{-1} \vec{x}}} \sim St_{n-m} = St_8 \quad (17)$$

Довірчий інтервал для середнього значення відклику матиме вигляд :

$$\left(f^*(\vec{x}) - t\sqrt{(\sigma^2)^{**}(\vec{x})^T A^{-1} \vec{x}}, f^*(\vec{x}) + t\sqrt{(\sigma^2)^{**}(\vec{x})^T A^{-1} \vec{x}} \right) \quad (18)$$

Величина t знаходиться з рівняння $\mathbb{P}(|\zeta| < t) = 0.95$; $t_{\text{зн.}} = 2.306$; $(\vec{x})^T A^{-1} \vec{x} \approx 0.262$.

Таким чином, отримаємо довірчий інтервал для середнього значення відклику в точці $(10, -9)$ з надійністю $\gamma = 0.95$:

$$(-2.22, -0.78) \quad (19)$$

Тепер побудуємо довірчий інтервал для самого значення відклику в точці $(10, -9)$ з надійністю $\gamma = 0.95$. Для цього скористаємось статистикою:

$$\zeta = \frac{\eta - f^*(\vec{x})}{\sqrt{(\sigma^2)^{**}(1 + (\vec{x})^T A^{-1} \vec{x})}} \sim St_{n-m} = St_8 \quad (20)$$

Довірчий інтервал матиме вигляд:

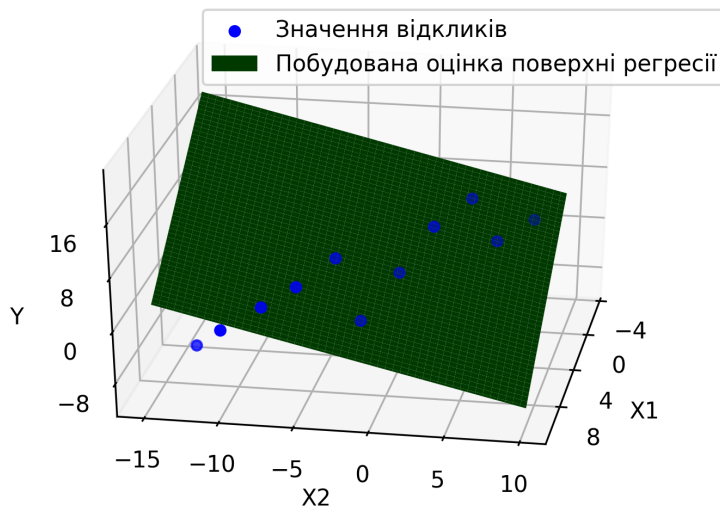
$$\left(f^*(\vec{x}) - t\sqrt{(\sigma^2)^{**}(1 + (\vec{x})^T A^{-1} \vec{x})}, f^*(\vec{x}) + t\sqrt{(\sigma^2)^{**}(1 + (\vec{x})^T A^{-1} \vec{x})} \right) \quad (21)$$

де t знаходимо з рівняння $\mathbb{P}(|\zeta| < t) = 0.95$. Всі необхідні нам значення були обчислені. Після усіх розрахунків отримаємо довірчий інтервал:

$$(-3.09, 0.09) \quad (22)$$

2.6 Висновки

В ході цієї роботи було побудовано двофакторну лінійну регресійну модель вигляду $f(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$. За допомогою методу найменших квадратів було знайдено оцінки параметрів цієї моделі. Була перевірена адекватність побудованої моделі на рівні значущості 0.05 та на тому ж самому рівні значущості була перевірена гіпотеза про значущість параметру з найменшим значенням його оцінки. Так як обидві гіпотези не протирічали дослідним даним, то модель залишилася незмінною. В подальшому був побудований довірчий інтервал для значення відклику та середнього значення відклику в точці (10, -9). Обидва інтервали були побудовані з довірчою ймовірністю 0.95. Необхідні обчислення були виконані за допомогою бібліотек Pandas, Matplotlib та Numpy мови програмування Python. Нижче на графіку зображена побудована поверхня регресії:



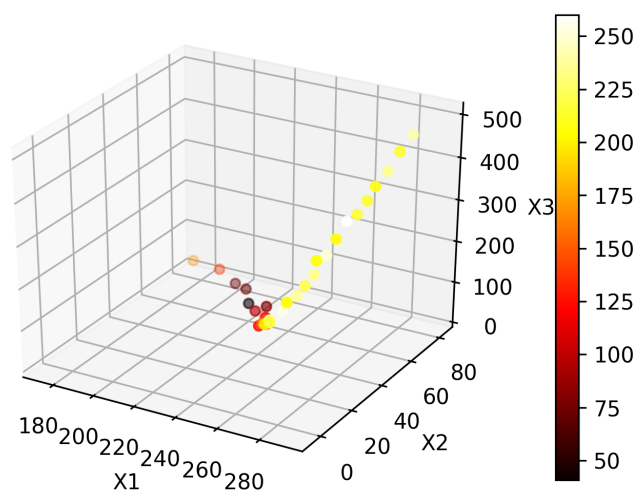
3 Завдання 3

3.1 Завдання

1. За методом найменших квадратів знайти оцінки параметрів трьохфакторної регресійної моделі.
2. На рівні значущості $\alpha = 0.05$ перевірити адекватність побудованої моделі.
3. Для самого малого значення параметра побудованої моделі на рівні значущості $\alpha = 0.05$ перевірити гіпотезу про його значущість.
4. Побудувати прогнозований довірчий інтервал з довірчою ймовірністю $\gamma = 0.95$ для середнього значення відклику та самого значення відклику в деякій точці (точку вибирайте самі).
5. Написати висновки.

№	X_1	X_2	X_3	Y	№	X_1	X_2	X_3	Y
1	173.50	83.00	22.0	169.91	16	281.69	0.00	259.0	254.34
2	194.05	74.25	48.0	138.22	17	283.58	0.00	282.0	205.54
3	209.37	65.00	56.0	69.01	18	285.21	4.00	288.0	238.28
4	221.52	56.25	81.0	72.50	19	286.62	7.00	304.0	222.52
5	231.49	45.00	89.0	40.83	20	287.81	11.00	320.0	235.60
6	239.88	38.25	100.0	96.49	21	288.81	11.00	353.0	208.35
7	247.04	36.00	126.0	82.22	22	289.63	16.00	353.0	249.94
8	253.22	27.25	132.0	115.50	23	290.28	21.00	377.0	208.17
9	258.59	16.00	151.0	126.87	24	290.78	27.00	402.0	260.02
10	263.29	12.25	174.0	160.31	25	291.13	33.00	402.0	217.63
11	267.42	9.00	186.0	159.74	26	291.35	39.00	418.0	216.33
12	271.05	6.25	203.0	215.98	27	291.43	44.00	438.0	208.81
13	274.24	-1.00	223.0	196.75	28	291.40	52.00	452.0	234.50
14	277.06	-1.75	238.0	237.24	29	291.25	60.00	477.0	212.37
15	279.53	-5.00	246.0	217.10	30	291.00	69.00	493.0	241.34

На малюнку нижче зображені значення відкликів у просторі.



Кольорами позначені значення відкликів при відповідних значеннях наших факторів.

3.2 Знаходження за МНК оцінки параметрів вибраної моделі

Матриця плану для вибраної моделі матиме вигляд:

$$F = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 173.5 & 194.05 & 209.37 & \dots & 291 \\ 83 & 74.25 & 65 & \dots & 69 \\ 22 & 48 & 56 & \dots & 493 \end{pmatrix}^T \quad (23)$$

Після цього використовуємо матрицю F для знаходження інформаційної матриці Фішера $A = FF^T$ та дисперсійну матрицю Фішера A^{-1} ; обидві з них мають розмір 4×4 . Обчислені матриці наведені у додатку 1. Знайдемо значення оцінок параметрів регресійної моделі за формулою:

$$\vec{\beta}_{\text{знач.}}^* = A^{-1}F^T\vec{\eta}_{\text{знач.}}, \quad (24)$$

де $\vec{\eta}_{\text{знач.}} = (169.91, 138.22, \dots, 241.34)^T$ - вектор значень відкликів. Отримаємо $\vec{\beta}_{\text{знач.}}^* \approx (1314.989, -5.069, -3.686, 1.265)^T$. Тоді сама модель матиме вигляд:

$$f^*(x) = 1314.989 - 5.069x_1 - 3.686x_2 + 1.265x_3 \quad (25)$$

3.3 Перевірка адекватності побудованої моделі

Перевіряємо на рівні значущості $\alpha = 0.05$ адекватність побудованої моделі. Задля перевірки побудованої регресійної моделі скористаємось тим факт, що випадкова величина

$$\gamma = \frac{(\sigma^2)^{**}}{D_{\eta}^{**}} \sim F(n - m, n - 1) \quad (26)$$

У нашому випадку $n = 30, m = 4 \Rightarrow$ статистика γ буде розподілена за законом Фішера-Снедекора із 26 та 29 ступенями вільності відповідно. Критична область при саме такому виборі статистики критерію буде правостороння, тому знайдемо межу критичної області з виразу $\mathbb{P}(\gamma > t) = \alpha$, $t_{\text{кр.}} = 1.84$; $\gamma_{\text{знач.}} \approx 0.0014 < t_{\text{кр.}}$. Таким чином гіпотеза про адекватність побудованої моделі на рівні значущості 0.05 не протирічить дослідним даним.

3.4 Перевірка гіпотези про значущість найменшого за значенням параметра

Перевіримо гіпотезу про значущість параметра для $(\beta_3^*)_{\text{знач.}} = 1.265$ на рівні значущості $\alpha = 0.05$. Основна гіпотеза $H_0 : \beta_3 = 0$ проти правосторонньої альтернативи $H_1 : \beta_3 > 0$, відповідно, критична область - правостороння. Для перевірки гіпотези скористаємося статистикою:

$$\gamma = \frac{\beta_3^*}{\sqrt{(\sigma^2)^{**} a_{33}}} \sim St_{n-m} \quad (27)$$

З таблиці розподілу Стюдента з $n - m = 26$ ступенями вільності отримуємо критичне значення $t_{\text{кр.}} = 1.706$. Знаходимо $(\sigma^2)^{**} \approx 5.589$, $a_{33} = (A^{-1})_{33} \approx 4.26 \cdot 10^{-5}$. Отримуємо $\gamma_{\text{знач.}} \approx 81.96$. Т.я. $\gamma_{\text{знач.}} > t_{\text{кр.}}$, то основна гіпотеза H_0 відхиляється на користь альтернативної. Таким чином, параметр β_3 є значущим, залишаємо побудовану модель незмінною.

3.5 Довірчий інтервал для середнього значення відклику та значення відклику в точці

Будуватимемо довірчий інтервал для середнього значення відклику та самого значення відклику в точці (230, 46, 90).

Спочатку побудуємо прогнозований довірчий інтервал для середнього значення відклику з довірчою ймовірністю $\gamma = 0.95$. Для цього скористаємось статистикою :

$$\zeta = \frac{f^*(\vec{x}) - f(\vec{x})}{\sqrt{(\sigma^2)^{**}(\vec{x})^T A^{-1} \vec{x}}} \sim St_{n-m} = St_{26} \quad (28)$$

Довірчий інтервал для середнього значення відклику матиме вигляд :

$$\left(f^*(\vec{x}) - t\sqrt{(\sigma^2)^{**}(\vec{x})^T A^{-1} \vec{x}}, f^*(\vec{x}) + t\sqrt{(\sigma^2)^{**}(\vec{x})^T A^{-1} \vec{x}} \right) \quad (29)$$

Величина t знаходиться з рівняння $\mathbb{P}(|\zeta| < t) = 0.95$; $t_{\text{зн.}} = 2.056$; $(\vec{x})^T A^{-1} \vec{x} \approx 0.117$.

Таким чином, отримаємо довірчий інтервал для середнього значення відклику в точці (230, 46, 90) з надійністю $\gamma = 0.95$:

$$(91.676, 95) \quad (30)$$

Тепер побудуємо довірчий інтервал для самого значення відклику в точці (230, 46, 90) з надійністю $\gamma = 0.95$. Для цього скористаємось статистикою:

$$\zeta = \frac{\eta - f^*(\vec{x})}{\sqrt{(\sigma^2)^{**}(1 + (\vec{x})^T A^{-1} \vec{x})}} \sim St_{n-m} = St_{26} \quad (31)$$

Довірчий інтервал матиме вигляд:

$$\left(f^*(\vec{x}) - t\sqrt{(\sigma^2)^{**}(1 + (\vec{x})^T A^{-1} \vec{x})}, f^*(\vec{x}) + t\sqrt{(\sigma^2)^{**}(1 + (\vec{x})^T A^{-1} \vec{x})} \right) \quad (32)$$

де t знаходимо з рівняння $\mathbb{P}(|\zeta| < t) = 0.95$. Всі необхідні нам значення були обчислені. Після усіх розрахунків отримаємо довірчий інтервал:

$$(88.201, 98.476) \quad (33)$$

3.6 Висновки

В ході цієї роботи було побудовано трьохфакторну лінійну регресійну модель вигляду $f(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$. За допомогою методу найменших квадратів було знайдено оцінки параметрів цієї моделі. Була перевірена адекватність побудованої моделі на рівні значущості 0.05 та на тому ж самому рівні значущості була перевірена гіпотеза про значущість параметру з найменшим значенням його оцінки. Так як обидві гіпотези не протирічали дослідним даним, то модель залишилася незмінною. В подальшому був побудований довірчий інтервал для значення відклику та середнього значення відклику в точці (230, 46, 90). Обидва інтервали були побудовані з довірчою ймовірністю 0.95. Необхідні обчислення були виконані за допомогою бібліотек Pandas, Matplotlib та Numpy мови програмування Python.