Національний технічний університет України «КПІ»

Кафедра автоматизованих систем обробки інформації і управління

Лабораторна робота №5

з дисципліни «ЙМОВІРНІСНІ МОДЕЛІ ТА СТАТИСТИЧНЕ ОЦІНЮВАННЯ В ІНФОРМАЦІЙНО-УПРАВЛЯЮЧИХ СИСТЕМАХ»

на тему: «Парна лінійна регресія»

Виконав:

студент групи ІС-23

Шимків М.В.

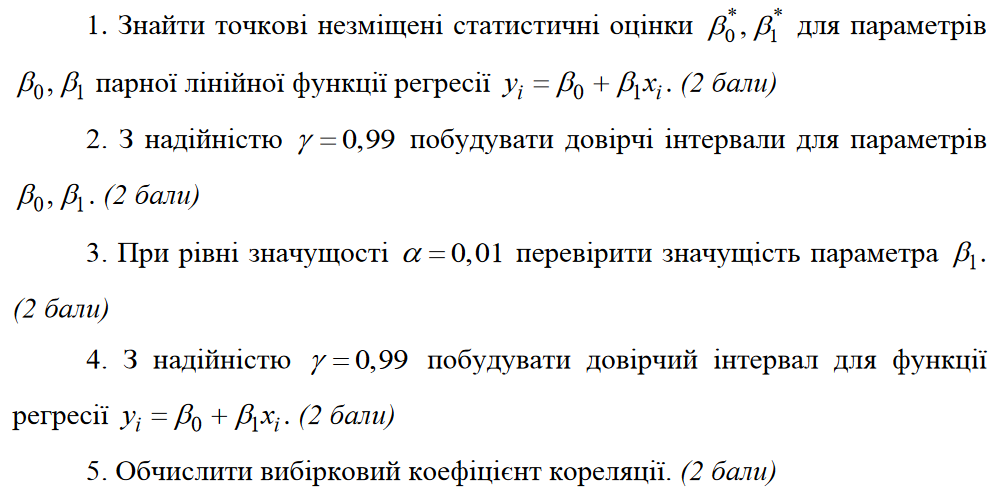
Викладач:

Богданова Н.В

Київ 2024

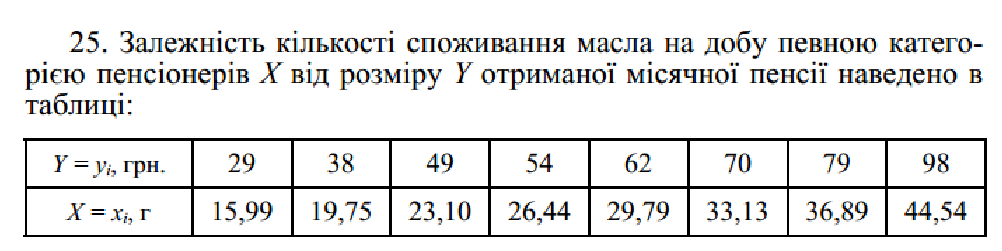
**Мета роботи**: Мета роботи полягає в оцінці параметрів парної лінійної регресії, перевірці їх значущості та побудові довірчих інтервалів для параметрів регресії та самої регресійної функції. Також потрібно обчислити вибірковий коефіцієнт кореляції для визначення сили та напрямку лінійного зв'язку між двома змінними.

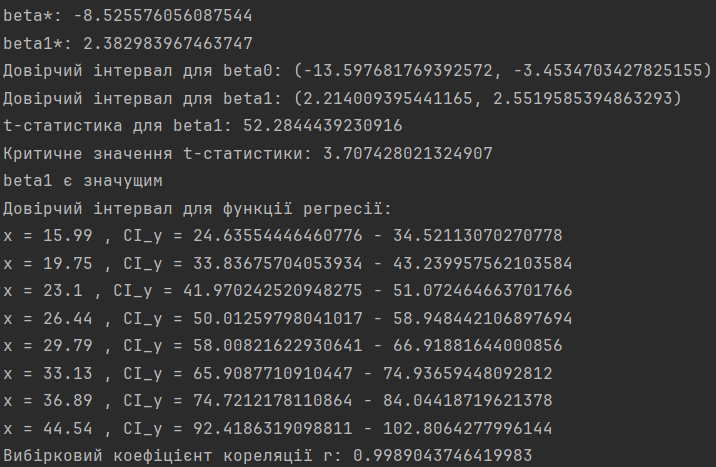
**Завдання до роботи:**

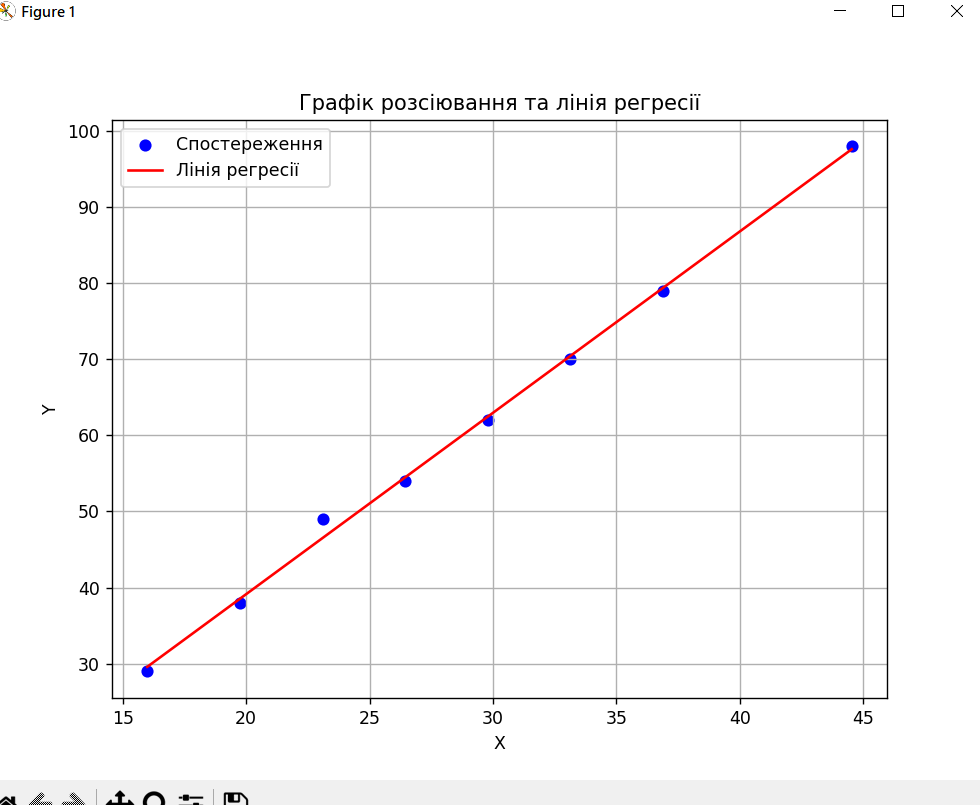


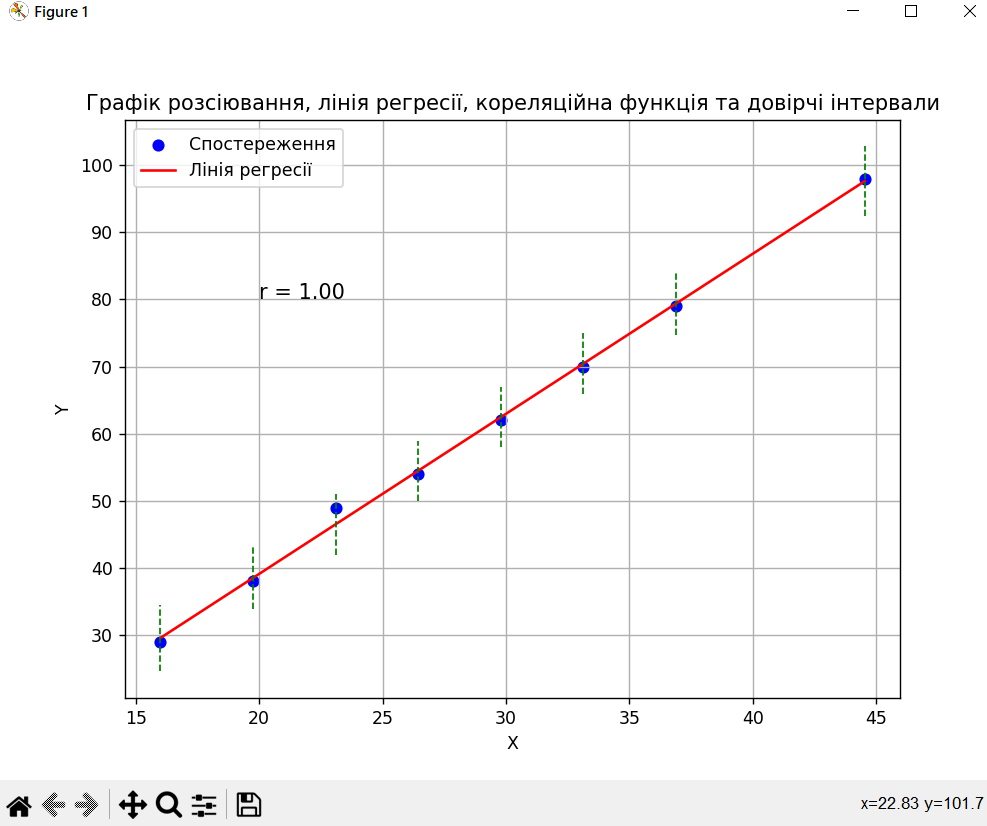
**Виконання:**

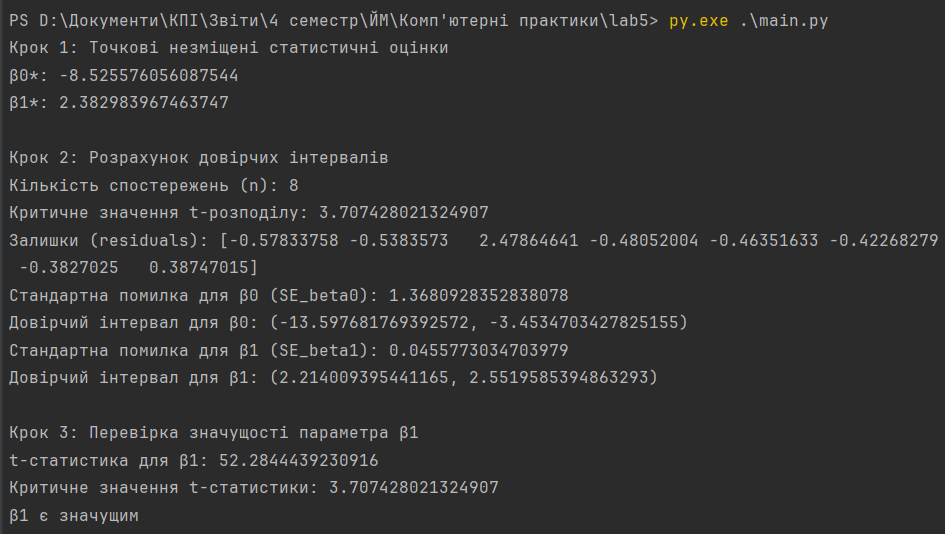
**Варіант 25**

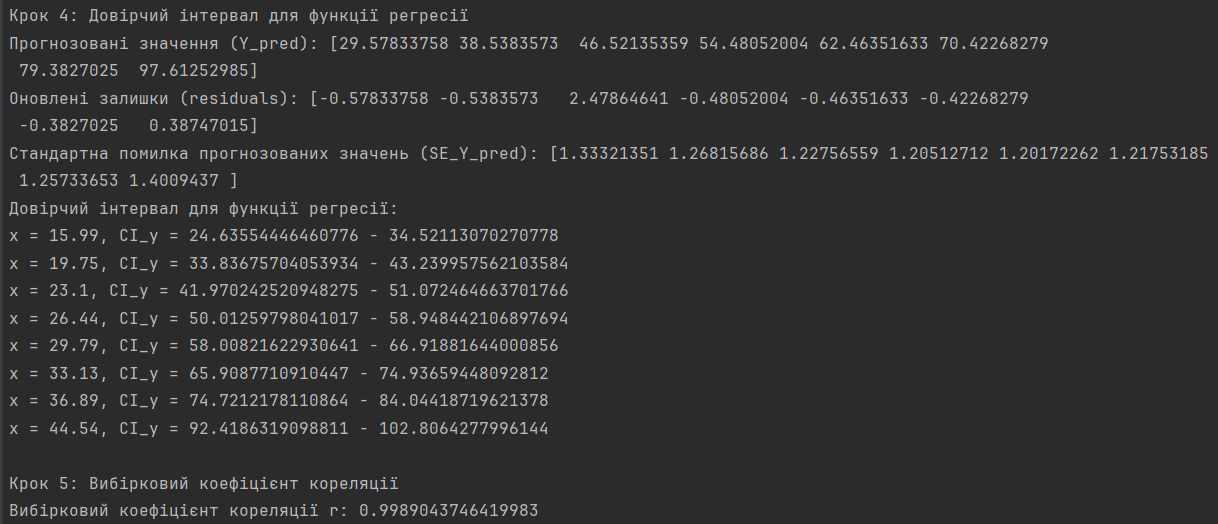


****

****

****

****

****

**Висновки:**

1. **Оцінки параметрів регресії:**
   * Точкова незміщена оцінка для β0β0​: −8.53−8.53
   * Точкова незміщена оцінка для β1β1​: 2.382.38
2. **Довірчі інтервали для параметрів:**
   * Довірчий інтервал для β0β0​ з надійністю 0.990.99: (−13.60,−3.45)(−13.60,−3.45)
   * Довірчий інтервал для β1β1​ з надійністю 0.990.99: (2.21,2.55)(2.21,2.55)
3. **Перевірка значущості параметра β1β1​:**
   * t-статистика для β1β1​: 52.2852.28
   * Критичне значення t-статистики при рівні значущості α=0.01α=0.01: 3.713.71
   * Оскільки t-статистика значно перевищує критичне значення, параметр β1β1​ є значущим.
4. **Довірчі інтервали для функції регресії з надійністю 0.990.99:**
   * Для різних значень xx, довірчі інтервали для прогнозованого значення yy наведені в діапазоні, що підтверджує високу точність прогнозів.
5. **Вибірковий коефіцієнт кореляції rr:**
   * Вибірковий коефіцієнт кореляції: 0.99890.9989
   * Це свідчить про дуже сильний позитивний лінійний зв'язок між змінними xx та yy.

### Загальні висновки:

Результати аналізу вказують на те, що парна лінійна регресія добре підходить для моделювання даних. Параметр β1β1​ є значущим, а високий коефіцієнт кореляції свідчить про сильний лінійний зв'язок між змінними. Довірчі інтервали для параметрів та прогнозованих значень функції регресії показують високу точність оцінок та прогнозів.

**Додаток з кодом:**

import numpy as np  
import scipy.stats as stats  
# Набір даних  
Y = np.array([29, 38, 49, 54, 62, 70, 79, 98])  
X = np.array([15.99, 19.75, 23.10, 26.44, 29.79, 33.13, 36.89, 44.54])  
  
# Крок 1: Знайти точкові незміщені статистичні оцінки в0\*, 61\*  
beta1 = np.cov(X, Y, bias=True)[0, 1] / np.var(X)  
beta0 = np.mean(Y) - beta1 \* np.mean(X)  
print("beta\*:", beta0)  
print("beta1\*:", beta1)  
# Крок 2: Побудувати довірчі інтервали для параметрів b0\*, b1\* з надійністю у = 0.99  
n = len(X)  
gamma = 0.99  
t\_value = stats.t.ppf(1 - (1 - gamma) / 2, n - 2) # двобічний довірчий інтервал  
  
residuals = Y - (beta0 + beta1 \* X)  
SE\_beta0 = np.sqrt(np.sum(residuals \*\* 2) / (n - 2) \* (1 / n + np.mean(X) \*\* 2 / np.sum((X - np.mean(X)) \*\* 2)))  
CI\_beta0 = (beta0 - t\_value \* SE\_beta0, beta0 + t\_value \* SE\_beta0)  
  
SE\_beta1 = np.sqrt(np.sum(residuals \*\* 2) / ((n - 2) \* np.sum((X - np.mean(X)) \*\* 2)))  
CI\_beta1 = (beta1 - t\_value \* SE\_beta1, beta1 + t\_value \* SE\_beta1)  
  
print("Довiрчий iнтервал для bеta0:", CI\_beta0)  
print("Довiрчий iнтервал для bеta1:", CI\_beta1)  
  
# Крок 3: Перевірити значущість параметра b1 при рівні значущості а = 0.01  
t\_statistic\_beta1 = beta1 / SE\_beta1  
critical\_t\_value = stats.t.ppf(0.995, n - 2)  
print("t-статистика для bеtа1:", t\_statistic\_beta1)  
print("Критичне значення t-статистики:", critical\_t\_value)  
  
if np.abs(t\_statistic\_beta1) > critical\_t\_value:  
 print("beta1 є значущим")  
else:  
 print("beta1 не є значущим")  
  
# Крок 4: Побудувати довірчий інтервал для функції регресії уі = b0 + b1хі з надійністю у = 0.99  
Y\_pred = beta0 + beta1 \* X  
residuals = Y - Y\_pred  
SE\_Y\_pred = np.sqrt((np.sum(residuals \*\* 2) / (n - 2)) \* (1 + 1 / n + (X - np.mean(X)) \*\* 2 / np.sum((X - np.mean(X)) \*\* 2)))  
CI\_Y\_pred = (Y\_pred - t\_value \* SE\_Y\_pred, Y\_pred + t\_value \* SE\_Y\_pred)  
print("Довірчий інтервал для функції регресії:")  
  
for i in range(len(X)):  
 print("x =", X[i], ", CI\_y =", CI\_Y\_pred[0][i], "-", CI\_Y\_pred[1][i])  
  
# Крок 5: Обчислити вибірковий коефіцієнт кореляції  
r = np.corrcoef(X, Y)[0, 1]  
print("Вибірковий коефіцієнт кореляції r:", r)