|  |  |  |  |
| --- | --- | --- | --- |
| МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені Тараса Шевченка  ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ  **Кафедра програмних систем і технологій**    Дисципліна  **«Ймовірнісні основи програмної інженерії»**  **Лабораторна робота № 3**  **Варіант: 9** | | | |
| **Виконав:** | Мичковський Богдан  Ігорович | **Перевірила**: | Вечерковська А. С. |
| Група | ІПЗ-22 | Дата перевірки |  |
| Форма навчання | денна | Оцінка |  |
| Спеціальність | 121 |
| 2022 | | | |
|  | | | |

**Мета:**

Навчитись використовувати на практиці набуті знання про міри в двовимірній статистиці.

**Завдання:**

1. Намалюйте діаграму розсіювання для даних. Укажіть, чи існує тренд у даних. Якщо так, то вкажіть, чи є це негативним трендом, чи позитивним.

2. Знайдіть центр ваги і коваріацію.

3. Знайти рівняння лініїї регресії y від x.

4. Розрахуйте коефіцієнт кореляції між даними.

5. Зробити висновок про залежності.

Формат вхідних/вихідних даних: Розроблена програма повинна зчитувати вхідні дані з файлу заданого формату та записувати дані у файл. Вхідний файл представляє собою текстовий файл із М+1 рядків. Перший рядок містить число: M, де M —кількість переглядів і-того фільму. До документа завдання також додаються приклад вхідних файлів різної розмірності.

**Постановка задачі:**

Перше задання має на увазі побудову діаграми розсіювання для вибірки, та аналіз тренду.

Друге завдання потребує знайти центр ваги лінії регресії та значення коваріації.

Третє завдання потребує знаходження лінії регресії за значеннями знайденими вище.

Четверте завдання має на увазі знаходження значення коефіцієнту кореляції між даними для аналізу графіку.

П’яте завдання потребує формування звіту на основі результатів виконання програми.

**Побудова математичної моделі:**

Для обрахунку центру ваги, коваріації, рівняння лінії регресії та коефіцієнту кореляції використаємо нище наведені формули.



Graphical user interface, application, Teams

Description automatically generated

Text

Description automatically generated

Graphical user interface, text

Description automatically generated

За допомогою знайдених вище значень дисперсії знайдемо значення параметра b1 та b0 і підставимо їх в рівняння. Результатом буде потрібне нам рівняння регресії.

Коефіцієнт кореляції знайдемо за формулою коваріація поділити на дисперсія по X та Y.

**ПсевдоКод :**

1) def BuildScatter(x,y):

plt.scatter(x, y)

plt.xlabel("X")

plt.ylabel("Y")

plt.show()

2)def Average(list):

average = 0

for i in list:

average += i

return average / len(list)

def Covariacion(x, y, aveX, aveY):

cov = 0

for i in range(len(x)-1):

cov += (x[i] - aveX)\* (y[i] - aveY)

return cov/len(x)

3) def Koeficient(cov, disp):

return cov / disp

b = round(aveY - k \* aveX, 3)

f.write(f"Regression line:\ty = {k} \* x + {b}\n")

4) def Corelation(x, y, aveX, aveY, cov, k):

dispX = Dispercian(x, aveX)

dispY = Dispercian(y, aveY)

return cov / sqrt(dispX \* dispY)

5) def CheckRegression(corelation):

if corelation == -1:

return "Line on points and falling"

elif corelation > -1 and corelation < 0:

return "Line almost on points and falling"

elif corelation > 0 and corelation < 1:

return "Line almost on points and growing"

elif corelation == 1:

return "Line on points and growing"

else: return "Line miss all points"

def Trend(corelation):

if corelation > 0:

return "Trend is positive"

elif corelation < 0:

return "Trend is positive"

else: return "Trend is missing"

**Випробування алгоритму:**

Для перевірки результатів скористаємося вибіркою з меншої кількості елементів, а саме 12. Результат виконання програми матиме поданий нище вигляд.

Text

Description automatically generated with medium confidenceChart, scatter chart

Description automatically generated

Так як центр ваги це середнє значення вибірки, то просто перерахуємо всі значення та поділим на їх кількість. Знайдемо координата X та Y.

**Circle

Description automatically generated with medium confidence A picture containing text, clipart

Description automatically generated**

Результат спів пав, програма працює вірно. Повторимо ці дії щодо наступних чисельних операцій та переконаймося, що програма працює вірно або наближено до результату. Так як програма показала коректні результати в лабораторній роботі номер два, перевіряти дисперсію сенсу немає.

A picture containing text

Description automatically generated A picture containing text

Description automatically generated A picture containing text

Description automatically generated

Відносно значень наведених вище виводимо інформацію про графік та лінію регресії.

**Висновок:**

Під час цієї виконання даної лабораторної роботи було сформовано навички побудови діаграми розсіювання, обрахунків таких параметрів в двомірній статистиці як: центр ваги, коефіцієнт кореляція, лінія регресії, коваріація та тренди. Сформовано навички аналізу графіків та прилежних до них змінних а також навички формування звіту.