

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 № 1

Виконав:	Якубець М. В.	Перевірила:	Вечерковська А. С.
Група	ІПЗ-22	Дата перевірки	
Форма навчання	денна	Оцінка	
Спеціальність	121		
2022			

Назва: Центральні тенденції та міра дисперсії.

Мета: Навчитись використовувати на практиці набуті знання про центральні тенденції та міри.

Постановка задачі:

1. Побудувати таблицю частот та сукупних частот для переглянутих фільмів.
Визначити фільм, який був переглянутий частіше за інші.
2. Знайти Моду та Медіану даної вибірки.
3. Порахувати Дисперсію та Середнє квадратичне відхилення розподілу.
4. Побудувати гістограму частот для даного розподілу.
5. Зробити висновок з вигляду гістограми, про закон розподілу.

Математична модель:

Mean Median Mode Formula



$$\text{Mean} = \frac{\text{Sum of Observations}}{\text{Total Number of Observations}}$$

$$\text{If 'n' is odd: Median} = \left(\frac{n+1}{2} \right)^{\text{th}} \text{ term}$$

$$\text{If 'n' is even: Median} = \frac{\left(\frac{n}{2} \right)^{\text{th}} \text{ term} + \left(\frac{n}{2} + 1 \right)^{\text{th}} \text{ term}}{2}$$

$$\text{Mode} = L + h \frac{(f_m - f_1)}{(f_m - f_1) + (f_m - f_2)}$$

$$D = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$$

Дисперсію = _____

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}.$$

Середнє квадратичне відхилення розподілу =

Код алгоритму:

```

1  from functools import reduce
2  from math import sqrt
3
4
5  def get_frequencies(list_: list[int]) -> dict[int, int]:
6      return {value: list_.count(value) for value in list_}
7
8
9  def get_cumulative_frequencies(frequencies: dict[int, int]) -> dict[int, int]:
10     cumulative_frequencies = {}
11     previous = 0
12     for key, frequency in frequencies.items():
13         cumulative_frequencies[key] = previous + frequency
14     return cumulative_frequencies
15
16
17 def get_max_values(list_: list[int]) -> list[tuple[int, int]]:
18     max_value = max(list_)
19     return [(index, value) for index, value in enumerate(list_, 1) if value == max_value]
20
21
22 def calculate_modes(list_: list[int]) -> list[int]:
23     frequencies = get_frequencies(list_)
24     max_value = max(frequencies.values())
25     return [value for value, frequency in frequencies.items() if frequency == max_value]
26
27
28 def calculate_median(list_: list[int]) -> float:
29     list_ = sorted(list_)
30     half_length = len(list_) / 2
31     return (list_[int(half_length - 0.5)] + list_[int(half_length)]) / 2
32
33
34 def calculate_mean(list_: list[int]) -> float:
35     movies_views_length = len(list_)
36     return reduce(lambda x, y: x + y, list_) / movies_views_length
37
38
39 def _calculate_numerator_dispersion(list_: list[int]) -> float:
40     mean = calculate_mean(list_)
41     result = 0
42     for movie_views in list_:
43         result += (movie_views - mean) ** 2
44     return result
45
46
47 def calculate_dispersion(list_: list[int]) -> float:
48     numerator = _calculate_numerator_dispersion(list_)

```

```

49     return numerator / (len(list_) - 1)
50
51
52     def calculate_standard_deviation(list_: list[int]) -> float:
53         numerator = _calculate_numerator_dispersion(list_)
54         return sqrt(numerator / len(list_))
55

```

Випробування алгоритму:

Завдання №1:

Таблиця частот:

Перегляди Частота

75	1
10	1
97	1
66	2
12	2
1	3

Таблиця сукупних частот:

Перегляди Сукупна частота

75	1
10	2
97	3
66	5
12	7
1	10

Фільм, переглянутий найчастіше: 8 -- 97

Завдання №2:

Мода: 1

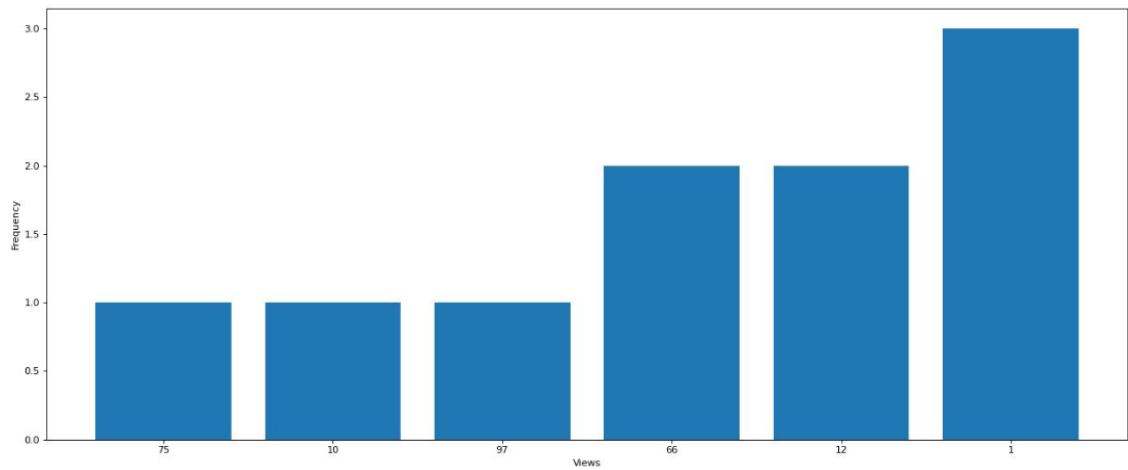
Медіана: 12.0

Завдання №3:

Дисперсія: 1389.88

Середнє квадратичне відхилення розподілу: 35.37

Завдання №4:



Медиана набора чисел

1 66 75 1 1 12 10 97 12 66

Разделитель пробел

Вычислить

1 66 75 1 1 12 10 97 12 66

Медіана = 12

Дисперсия (s^2)

1389.88

Стандартное отклонение (s)	37.2811
Считать (n)	10
Сумма (Σx)	341
Иметь в виду (\bar{x})	34.1
Коэффициент дисперсии	1.0933
Сумма площадей (SS)	12508.9

Дисперсія =

Середнє квадратичне відхилення розподілу =

Standard Deviation Calculator

Result

Standard Deviation, σ : **35.367923320433**

Count, N: 10
Sum, Σx : 341
Mean, μ : 34.1
Variance, σ^2 : 1250.89

Steps

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}.$$

$$\begin{aligned}\sigma^2 &= \frac{\sum (x_i - \mu)^2}{N} \\ &= \frac{(1 - 34.1)^2 + \dots + (66 - 34.1)^2}{10} \\ &= \frac{12508.9}{10} \\ &= 1250.89 \\ \sigma &= \sqrt{1250.89} \\ &= 35.367923320433\end{aligned}$$

Висновок: Навчився використовувати на практиці набуті знання про центральні тенденції та міри. Побачив та довів зв'язок моди, медіани, дисперсії та середньоквадратичного відхилення розподілу з початковими даними. Отримав досвід роботи з гістограмою.