

ІНДИВІДУАЛЬНЕ ЗАВДАННЯ №1 З КУРСУ

«МАТЕМАТИЧНА СТАТИСТИКА»

Кравець Ольга ПМО-21

Постановка задачі

Згенерувати вибірку заданого об'єму (не менше 50) з вказаного проміжку для дискретної статистичної змінної.

1. На підставі отриманих вибірових даних: побудувати варіаційний ряд та частотну таблицю; представити графічно статистичний матеріал, побудувати емпіричну функцію розподілу; обчислити числові характеристики дискретного розподілу.
2. Для цієї вибірки утворити інтервальний розподіл, побудувати гістограму розподілу, обчислити числові характеристики для згрупованих даних.

Теоретичні відомості

Ряд $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ називається **варіаційним рядом** для спостереження $x_{(1)}, \dots, x_{(n)}$ над одновимірною неперервною мінливою величиною.

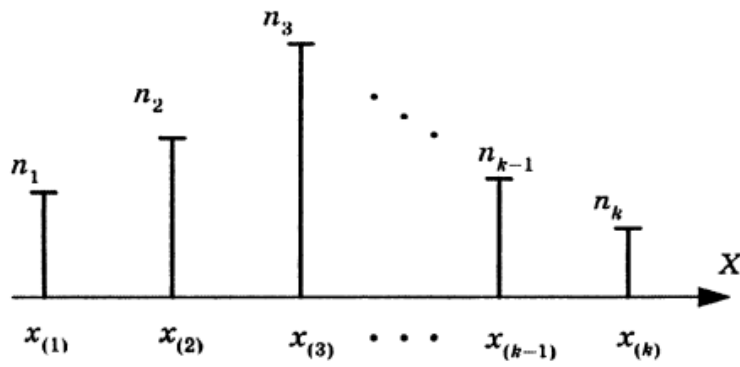
Таблицю

X	$x_{(1)}$	$x_{(1)}$...	$x_{(k)}$
v	n_1	n_2	...	n_k

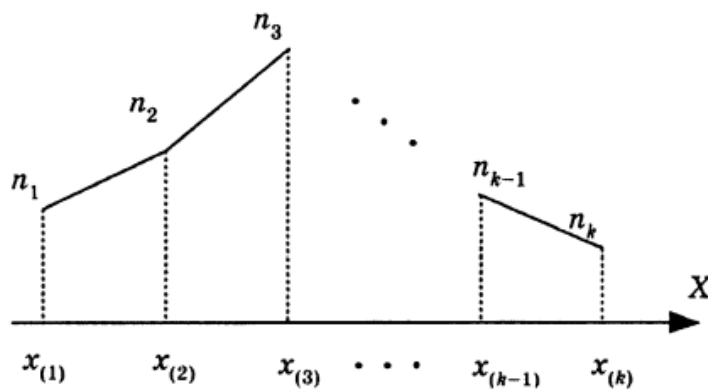
називають **частотною таблицею** або **статистичним розподілом дискретної варіанти X**.

Діаграма частот:

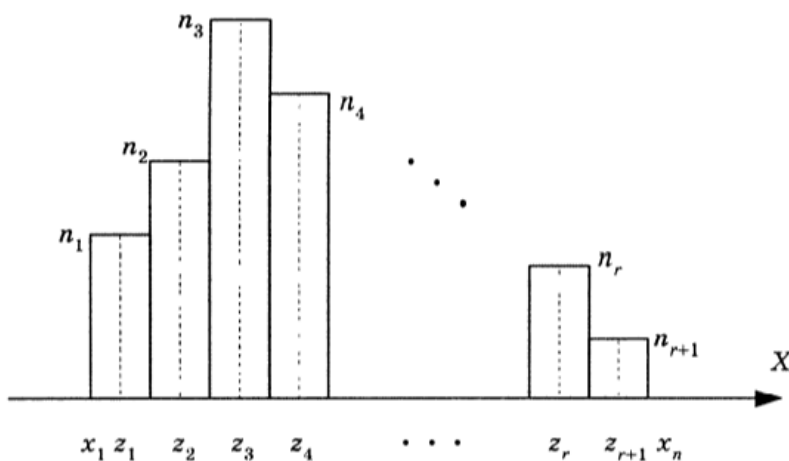
на вісі абсцис – можливі значення дискретної величини і в цих точках відповідні частоти n_i



Якщо з'єднати відрізками сусідні пункти $(x_{(i)}, n_i)$, то отримаємо **полігон частот**.



Абсциси – центри інтервалів, в точці z_i ставимо ординату n_i . Якщо над інтервалом з центром в точці z_i поставити прямокутник висотою n_i , то одержимо **гістограму частот**.



Функцію $F_n(x)$ називається **емпіричною функцією розподілу**.

$$F_n(x) = \begin{cases} 0, & x < x_{(1)}, \\ \frac{m_i}{n}, & x_{(i)} \leq x < x_{(i+1)}, \quad (i = \overline{1, n-1}), \\ 1, & x_{(n)} \leq x, \end{cases}$$

де, m_i – кількість елементів вибірки, які не більші за x .

Числові характеристики дискретного розподілу:

Середнім арифметичним називається сума всіх елементів статистичного матеріалу, поділена на обсяг статистичного матеріалу, позначається \tilde{x} .

$$\tilde{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Модою називається цей елемент статистичного матеріалу, який найчастіше зустрічається, позначається M_0 (може бути декілька).

Медіаною називається цей елемент статистичного матеріалу, який ділить відповідний варіаційний ряд $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ на дві рівні за обсягом частини, позначається M_e .

Якщо обсяг статистичного матеріалу n є непарним числом, тобто $n = 2 \cdot k + 1$, де k — натуральне число, то медіана $Me = x_{k+1}$:

$$\underbrace{x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(k)}}_{k\text{-елементів}} \leq x_{(k+1)} \leq \underbrace{x_{(k+2)} \leq \dots \leq x_{(n)}}_{k\text{-елементів}}.$$

Якщо обсяг статистичного матеріалу є парним числом, тобто $n = 2 \cdot k$, де k — натуральне число, то медіаною вважаємо величину² $Me = \frac{x_{(k)} + x_{(k+1)}}{2}$:

$$\underbrace{x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(k)}}_{k\text{-елементів}} \leq \underbrace{x_{(k+1)} \leq \dots \leq x_{(n)}}_{k\text{-елементів}}.$$

Сума квадратів відхилень елементів статистичного матеріалу від її середнього називається **девіацією**, позначається **dev**.

$$\text{dev} = \sum_{i=1}^n n_i (x_i - \tilde{x})^2$$

Варіансою називається девіація поділена на обсяг статистичного матеріалу без одного, позначається s^2 .

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

Дисперсія

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

Між варіансою та вибірковою дисперсією є простий зв'язок:

$$s^2 = \frac{n}{n-1} D$$

Стандартом називають арифметичний корінь з варіанси, позначається s .

$$s = +\sqrt{s^2}$$

Варіацією називають відношення стандарту вибірки до середнього арифметичного.

$$v = \frac{s}{\bar{x}}$$

Розмахом називається різниця між найбільшим і найменшим елементами статистичного матеріалу, позначається ρ .

$$\rho = \max(x_{(1)}, x_{(n)}) - \min(x_{(1)}, x_{(n)}) = x_{(n)} - x_{(1)}$$

Середньоквадратичним відхиленням називається арифметичний корінь з дисперсії, позначається $\tilde{\delta}$.

$$\tilde{\delta} = \sqrt{D}$$

Програмна реалізація виконана на мові Python.

Отримані результати

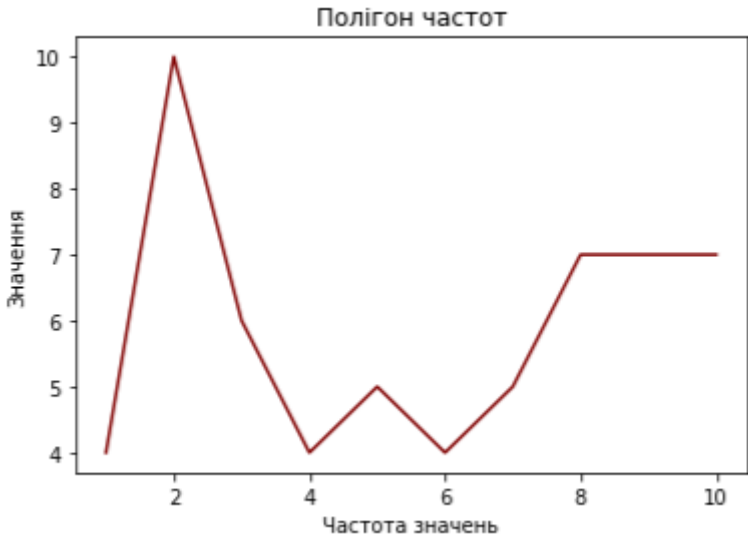
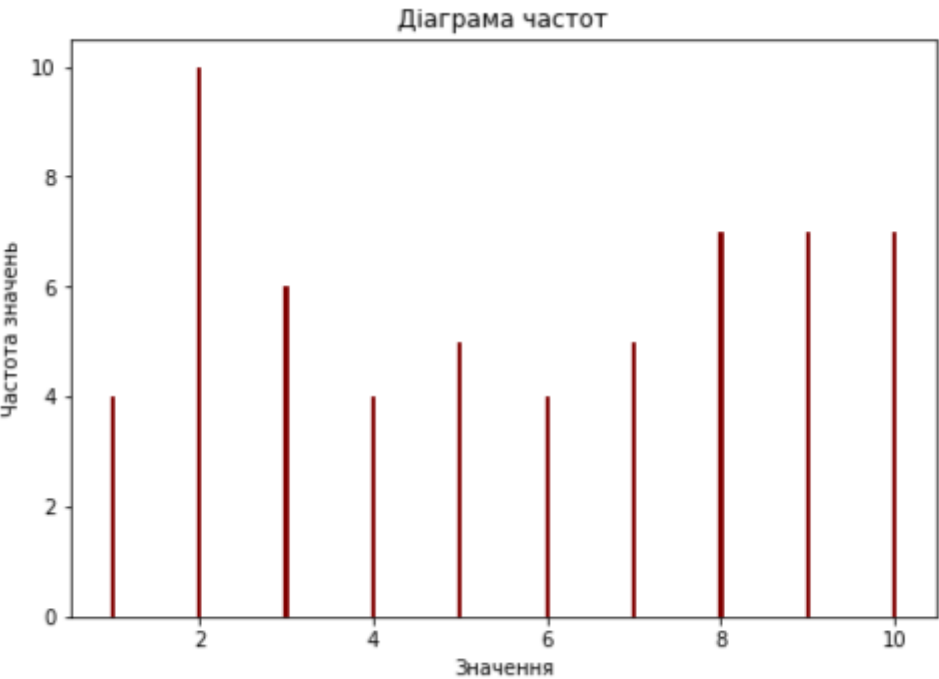
Введіть a: 1
Введіть b: 11
Введіть n: 59

Умова: [10, 2, 5, 7, 4, 8, 5, 2, 8, 4, 6, 2, 2, 8, 3, 5, 10, 8, 8, 6, 9, 8, 7, 10, 7, 5, 3, 3, 9, 3, 2, 4, 5, 9, 10, 9, 1, 3, 1, 10, 2, 1, 8, 10, 1, 7, 10, 7, 2, 6, 2, 9, 4, 2, 2, 9, 6, 3, 9]

Варіаційний ряд: [1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 7, 7, 7, 7, 7, 8, 8, 8, 8, 8, 8, 8, 8, 8, 9, 9, 9, 9, 9, 9, 9, 9, 9, 10, 10, 10, 10, 10, 10, 10, 10]

Частотна таблиця: значення-частота

1	4
2	10
3	6
4	4
5	5
6	4
7	5
8	7
9	7
10	7



```
---Числові характеристики---
Середнє арифметичне = 5.610169491525424
Мода = 2
Медіана = 6
Девіація = 534.0338983050847
Варіанса = 9.20748100526008
Дисперсія = 9.051422005170927
Стандарт = 3.034383134223508
Розмах = 9
Варіація = 0.5408719181848549
Середньоквадратичне відхилення = 3.008558127271422
Емпірична функція Fn(x)
1      0.067797
2      0.237288
3      0.338983
4      0.406780
5      0.491525
6      0.559322
7      0.644068
8      0.762712
9      0.881356
10     1.000000
```

```
Введіть m: 3
Довжина класу = 3.0
[[1, 4.0], [4.0, 7.0], [7.0, 10.0]]
```

Висновки. Під час виконання індивідуального завдання я навчилася генерувати вибірку заданого об'єму певного проміжку для неперервної та дискретної статистичної змінної. Також на основі отриманих вибірових даних будувати варіаційний ряд, частотну таблицю, представляти статистичний матеріал графічно, будувати емпіричну функцію розподілу та її графік і обчислювати числові характеристики: моду, медіану, середнє вибірове, девіацію, варіансу, стандарт, варіацію, розмах, вибірову дисперсію, квадратичне відхилення.