

ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені ІВАНА ФРАНКА

Факультет прикладної математики та інформатики

Кафедра дискретного аналізу

Звіт з дисципліни
“Теорія ймовірності та математична статистика”

ІНДИВІДУАЛЬНЕ ЗАВДАННЯ № 1

Виконала:

A blue rectangular redaction mark covering the signature of the student.

Оцінка

Перевірила:

Квасниця Г. А.

Постановка задачі

Згенерувати вибірку заданого об'єму з певного проміжку для дискретної статистичної змінної. На підставі отриманих вибірових даних:

1. Побудувати варіаційний ряд;
2. Побудувати частотну таблицю;
3. Побудувати емпіричну функцію розподілу;
4. Представити графічно статистичний матеріал;
5. Обчислити числові характеристики.

Згенерувати вибірку заданого об'єму з певного проміжку для неперервної статистичної змінної. На підставі отриманих вибірових даних:

1. Побудувати варіаційний ряд;
2. Побудувати частотну таблицю;
3. Представити графічно статистичний матеріал;
4. Обчислити числові характеристики.

Короткі теоретичні відомості

Вибірковим середнім статистичного розподілу вибірки називається середнє арифметичне її варіант з урахуванням їх частот.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k n_i x_i = \frac{n_1 x_1 + n_2 x_2 + \dots + n_k x_k}{n}.$$

Модою дискретного статистичного розподілу вибірки називається те значення варіанти, якому відповідає найбільша частота. Іншими словами, мода – це варіанта, яка найчастіше зустрічається у вибірці.

Медіаною дискретного статистичного розподілу вибірки називається варіанта, з обох сторін від якої міститься не більше ніж половина варіант, враховуючи їх частоту.

Розмахом вибірки називають різницю між найбільшим і найменшим значеннями її варіант.

Девіація - сума квадратів відхилень елементів статистичного матеріалу від середнього арифметичного

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

Варіансою називається девіація, поділена на обсяг статистичного матеріалу без одного.

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

Стандартом (флюктуацією) називається арифметичний корінь з варіанси і позначається

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

Варіацією називається відношення стандарту до середнього арифметичного.

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

Вибірковою дисперсією статистичного розподілу вибірки називають середнє арифметичне квадратів відхилень його варіант від вибіркового середнього

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

Вибірковим середнім квадратичним відхиленням називається число, яке обчислюється за формулою

$$\bar{\sigma} = \sqrt{\bar{D}}$$

Квантилем порядку, якщо він існує, називається той елемент статистичного матеріалу (відповідного варіаційного ряду), до якого включно маємо $\alpha\%$ елементів статистичного матеріалу (відповідного варіаційного ряду).

При $\alpha < \beta$, різницю між квантилем порядку α і квантилем порядку β називають інтерквантильною шириною порядку $\beta - \alpha$.

Асиметрією або **скошеністю** статистичного матеріалу називається відношення 3-го центрального моменту до 2-го центрального моменту в степені півтора.

$$A_s = \gamma_1 = \frac{\mu_3}{\mu_2^{3/2}}$$

Якщо асиметрія > 0 , то статистичний матеріал скошений вправо і має додатну асиметрію.

Якщо асиметрія < 0 , то статистичний матеріал скошений вліво і має від'ємну асиметрію.

Якщо асиметрія $= 0$, то статистичний матеріал симетричний.

Ексцесом (крутістю, сплюсненістю) статистичного матеріалу називається відношення 4-го центрального моменту до 2-го центрального моменту в квадраті мінус три.

$$E_k = \gamma_2 = \frac{\mu_4}{\mu_2^2} - 3$$

При ексцесі > 0 , то статистичний матеріал високовершинний.

При ексцесі < 0 , то статистичний матеріал низьковершинний.

При ексцесі $= 0$, то статистичний матеріал нормальновершинний.

Програмна реалізація

Реалізувала програму мовою Python в середовищі PyCharm. Згенерувала вибірку, використовуючи бібліотеку `random`, задаючи межі та розмір дискретної статистичної змінної. На основі вибірки побудувала варіаційний ряд, гістограму, полігон частот, емпіричну функцію розподілу та обчислила числові характеристики. Таким же чином згенерувала та обчислила дані для неперервної статистичної змінної, поділивши вибірку на 5 класів (інтервалів). Для реалізації використовувала бібліотеки `random` та `matplotlib.pyplot`.

Отримані результати та їхній аналіз

Дискретна величина.

Дані в межах від 0 до 9 розміром 100.

Вибірка:

0 7 8 4 5 7 1 3 1 5 5 8 6 9 6 8 3 0 9 7 5 4 8 7 8 7 8 6 8 2 9 5 6 8 0 0 9 1 4 9 9 2 5 3 4 3 0 0 9
8 7 3 9 7 2 0 6 0 9 9 4 1 0 6 6 8 3 9 5 9 7 8 7 1 8 7 2 4 2 9 9 1 5 8 2 2 5 8 9 2 7 1 4 9 6 5 6 9
4 1

Варіаційний ряд:

0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 3 3 3 3 3 4 4 4 4 4 4 4 5 5 5 5 5 5 5 5 5 5
6 6 6 6 6 6 6 6 6 7 7 7 7 7 7 7 7 7 8 8 8 8 8 8 8 8 8 8 8 8 9 9 9 9 9 9 9 9 9 9 9 9 9 9
9 9

Частотна таблиця:

x_i	0	1	2	3	4	5	6	7	8	9
n_i	9	8	8	6	8	10	9	11	14	17

Таблиця ймовірностей:

x_i	0	1	2	3	4	5	6	7	8	9
p_i	0.09	0.08	0.08	0.06	0.08	0.1	0.09	0.11	0.14	0.17

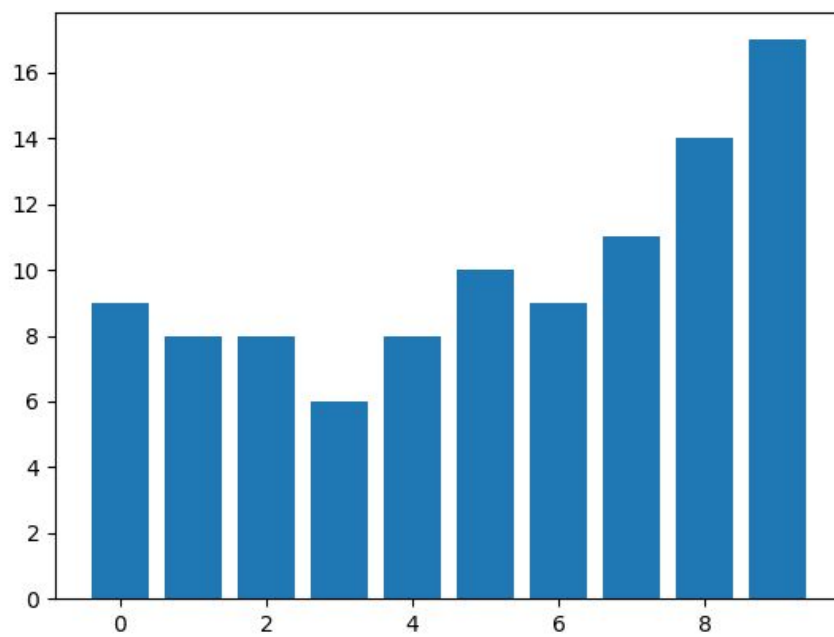
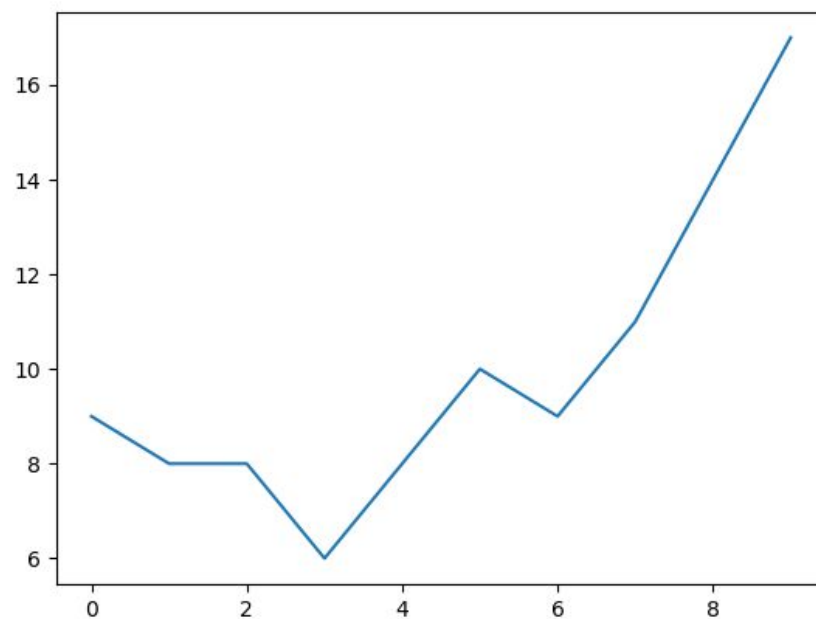
Емпірична функція розподілу:

0	$x < 0$
0.09	$0 \leq x < 1$
0.17	$1 \leq x < 2$
0.25	$2 \leq x < 3$
0.31	$3 \leq x < 4$
0.39	$4 \leq x < 5$
0.49	$5 \leq x < 6$
0.58	$6 \leq x < 7$
0.69	$7 \leq x < 8$

0.83 $8 \leq x < 9$

1 $x > 9$

Графічне представлення



Числові характеристики

Мода: 9
Медіана: 6.0
Середнє вибіркове: 5.2
Девіація: 904.0
Варіанса: 9.131
Стандарт: 3.022
Варіація: 0.581
Розмах: 9
Вибіркова дисперсія: 9.04
Квадратичне відхилення: 3.007
Асиметрія: -0.024
Екссес: -1.215
Квартилі: 2 5 8
Широта: 6
Децилі: 0 2 3 4 5 7 7 8 9
Широта: 9

Неперервна величина.

Дані в межах від 1.01 до 1.66 розміром 50.

Вибірка:

1.22 1.14 1.52 1.06 1.5 1.35 1.52 1.16 1.54 1.39 1.2 1.09 1.63 1.02 1.03 1.64 1.21 1.08 1.27
1.65 1.39 1.2 1.08 1.36 1.43 1.43 1.45 1.42 1.37 1.53 1.66 1.62 1.41 1.53 1.47 1.27 1.46
1.49 1.65 1.06 1.49 1.1 1.57 1.11 1.59 1.17 1.42 1.6 1.53 1.1

Варіаційний ряд:

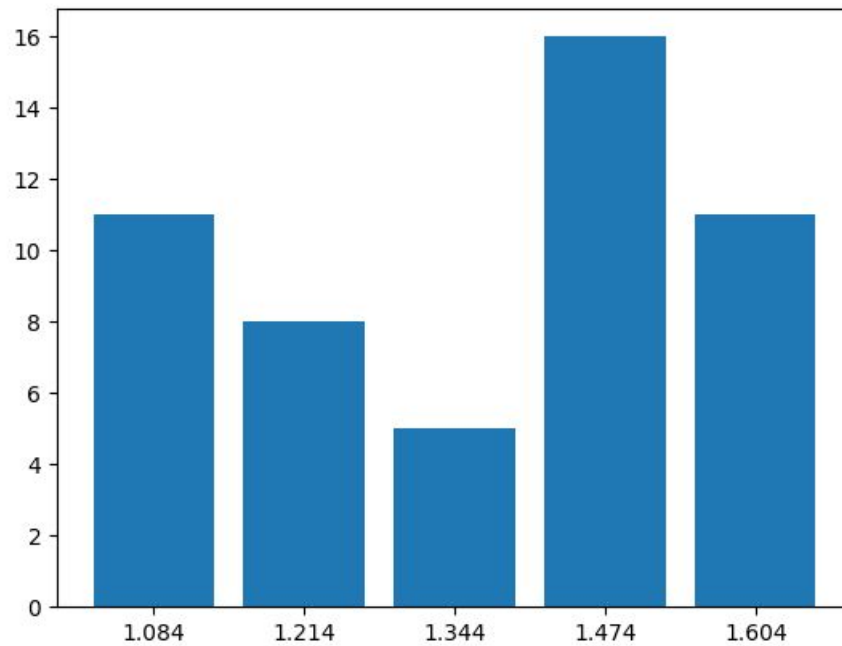
1.02 1.03 1.06 1.06 1.08 1.08 1.09 1.1 1.1 1.11 1.14 1.16 1.17 1.2 1.2 1.21 1.22 1.27 1.27
1.35 1.36 1.37 1.39 1.39 1.41 1.42 1.42 1.43 1.43 1.45 1.46 1.47 1.49 1.49 1.5 1.52 1.52
1.53 1.53 1.53 1.54 1.57 1.59 1.6 1.62 1.63 1.64 1.65 1.65 1.66

Частотна таблиця:

1.02 - 1.15	1.15 - 1.28	1.28 - 1.41	1.41 - 1.54	1.54 - 1.67
11	8	5	16	11

Середні значення інтервалів: 1.084, 1.214, 1.344, 1.474, 1.604

Графічне представлення



Інтервальні характеристики.

Мода: 1.529

Медіана: 1.418

Характеристики для середини відрізків.

Мода: 1.474

Медіана: 1.415

Середнє вибіркове: 1.392

Девіація: 1.91

Варіанса: 0.039

Стандарт: 0.197

Варіація: 0.142

Розмах: 0.52

Вибіркова дисперсія: 0.038

Квадратичне відхилення: 0.195

Асиметрія: -180.669

Ексцес: -1.222

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, квантилі (квартилі та децилі), асиметрію та ексцес.