

МІНІСТЕРСТВО ОСВІТИ ТА НАУКИ УКРАЇНИ  
ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені ІВАНА ФРАНКА

Кафедра дискретного аналізу  
та інтелектуальних систем

**Індивідуальне завдання №2**  
з курсу "Теорія ймовірності та математична статистика"

Виконав:  
студент групи ПМі-21  
Урдейчук Ростислав Ігорович

Оцінка

Перевірила:  
доц. Квасниця Г.А.

Львів 2024

### Постановка задачі:

1. Зчитати дані з текстового файлу, побудувати полігон або гістограму частот;
2. На основі графічного представлення сформулювати гіпотезу про закон розподілу досліджуваної ознаки генеральної сукупності;
3. Передбачити можливість користувачу задати параметри розподілу вручну або оцінити на основі даних вибірки;
4. Для заданого користувачем рівня значущості перевірити сформульовану гіпотезу за критерієм  $\chi^2$ .

### Варіант 16

**ЗАДАЧА 3 (варіанти 12-16).** Для вдосконалення організації праці на підприємствах торгівлі були зібрані дані про реалізацію за місяць товарів у магазинах міста. Розподіл кількості  $n_i$  магазинів залежно від обсягу реалізації  $X$  наведено в таблиці

$X$ , ум.од.	28-30	30-32	32-34	34-36	36-38	38-40	40-42	42-44	44-46	46-48
$n_i$ (варіант 12)	1	2	10	51	88	85	45	15	3	1
$n_i$ (варіант 13)	1	2	10	48	88	91	45	13	3	1
$n_i$ (варіант 14)	1	2	8	51	94	85	42	15	3	1
$n_i$ (варіант 15)	1	2	12	51	82	85	48	15	3	1
$n_i$ (варіант 16)	1	2	10	54	88	79	45	17	3	1

**ЗАДАЧА 9 (варіанти 16-20)** Для покращення обслуговування пасажирів реєстрували час очікування автобусів на зупинці. Розподіл за день кількості пасажирів  $n_i$  залежно від часу  $T$ , який вони очікували на автобус, відображено в таблиці.

$T$ , час очікування (хв)	0-1,5	1,5-3	3-4,5	4,5-6	6-7,5	7,5-9	9-10,5	10,5-12	12-13,5	13,5-15
$n_i$ (варіант 13)	143	138	139	142	143	138	139	143	142	140
$n_i$ (варіант 14)	119	122	131	118	122	123	118	117	132	124
$n_i$ (варіант 15)	108	115	112	107	108	111	113	114	116	104
$n_i$ (варіант 16)	142	141	138	136	137	142	143	154	144	135
$n_i$ (варіант 20)	138	159	147	143	148	175	190	147	123	134

## Короткі теоретичні відомості:

Перевірка гіпотез про закон розподілу випадкової величини — це процес, в ході якого статистично перевіряється відповідність спостережуваних даних певному теоретичному розподілу.

Він проходить у кілька етапів:

1. *Формулювання гіпотез*: висувається гіпотеза  $H_0$  про закон розподілу, яким керується вибірка
2. *Визначення рівня значущості*: ймовірність відхилення нульової гіпотези, виражена у відсотках
3. *Обчислення тестової статистики*: цей процес залежить від того, який метод ми обираємо. Наприклад, за критерієм Пірсона, ми обчислюємо величину  $\chi^2$ -квадратичне
4. *Прийняття рішення*: порівнюючи отримані значення з критичним значенням, ми приймаємо або відхиляємо початкову гіпотезу

Для критерія Пірсона ми обчислюємо величину  $\chi^2$ -критичне, використовуючи рівень значущості та ступені свободи.

## Програмна реалізація:

Для написання коду я використав мову програмування Python, середовище Jupyter Notebook і наступні бібліотеки: pandas, numpy, matplotlib, scipy

Спочатку я зчитую дані вибірки з файлу за допомогою методів pandas, показую їх у вигляді таблиці та будує гістограму за допомогою бібліотеки matplotlib. Потім показую висунуту гіпотезу. Користувач може або ввести дані розподілу самостійно, або вибрати, щоб вони були оцінені. Він також вводить значення рівня значущості.

Потім обчислюю ймовірності для кожного проміжку в окремій функції. Для першої задачі, з неперервним випадком, я використовую функцію лапласа; для другої, з експонентним випадком, я використовую  $F(x)$  для експонентного закону розподілу.

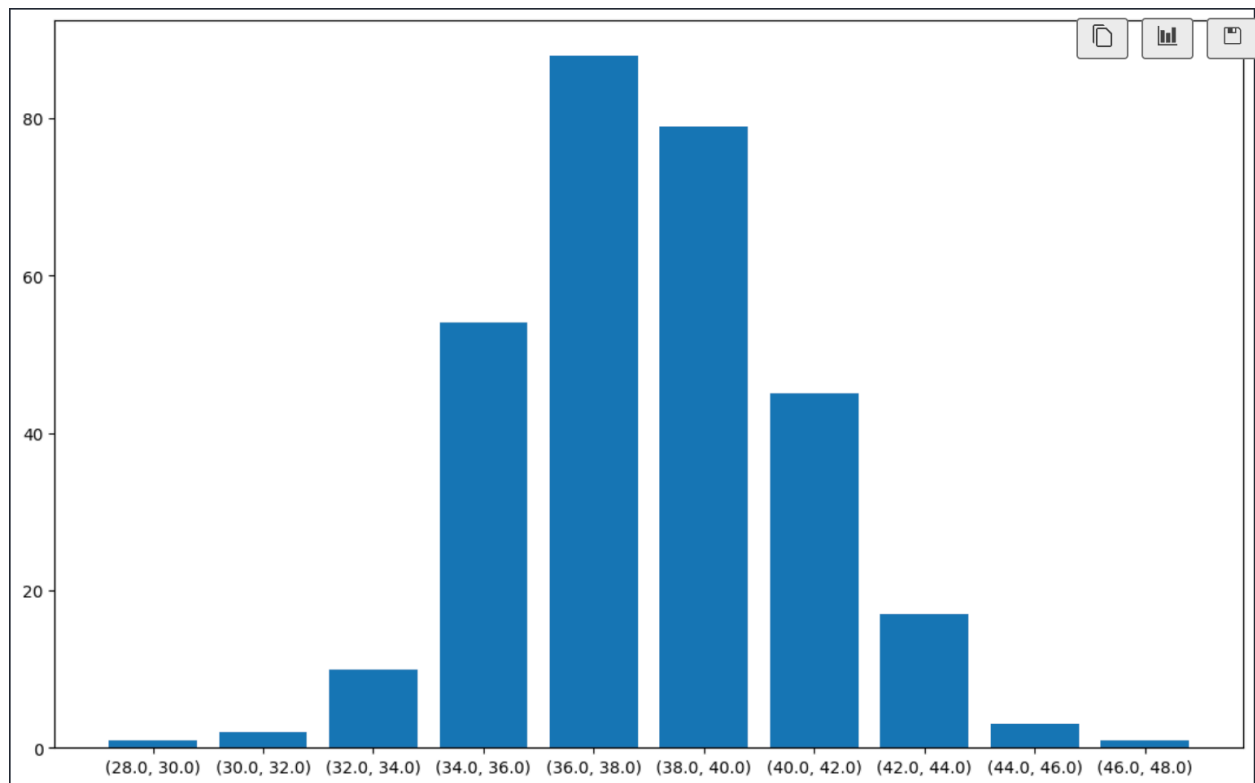
Після цього я перевіряю, чи для всіх рядків виконується критерій Пірсона, і якщо ні, то об'єдную ті з них, де це потрібно. Роблю це в окремій функції, використовуючи методи pandas — `iloc` та `concat`.

В кінці обчислюю значення  $\chi^2$  квадрат емпіричного та критичного. Емпіричне рахую за формулою, а критичне — за допомогою функції `chi2.ppf()`. Якщо емпіричне значення менше за критичне, то гіпотеза приймається, якщо ні — відхиляється.

## Аналіз отриманих результатів:

### Перша задача:

0	(28.0, 30.0)	1
1	(30.0, 32.0)	2
2	(32.0, 34.0)	10
3	(34.0, 36.0)	54
4	(36.0, 38.0)	88
5	(38.0, 40.0)	79
6	(40.0, 42.0)	45
7	(42.0, 44.0)	17
8	(44.0, 46.0)	3
9	(46.0, 48.0)	1



Гіпотеза  $H_0$  – нормальний закон розподілу

	xi	ni	avg_interval	pi
0	(28.0, 30.0)	1	29.0	0.001359
1	(30.0, 32.0)	2	31.0	0.010856
2	(32.0, 34.0)	10	33.0	0.054245
3	(34.0, 36.0)	54	35.0	0.158640
4	(36.0, 38.0)	88	37.0	0.271917
5	(38.0, 40.0)	79	39.0	0.273371
6	(40.0, 42.0)	45	41.0	0.161200
7	(42.0, 44.0)	17	43.0	0.055713
8	(44.0, 46.0)	3	45.0	0.011270
9	(46.0, 48.0)	1	47.0	0.001427

	xi	ni	pi
0	(28.0, 36.0)	67	0.225101
1	(36.0, 38.0)	88	0.271917
2	(38.0, 40.0)	79	0.273371
3	(40.0, 48.0)	66	0.229612

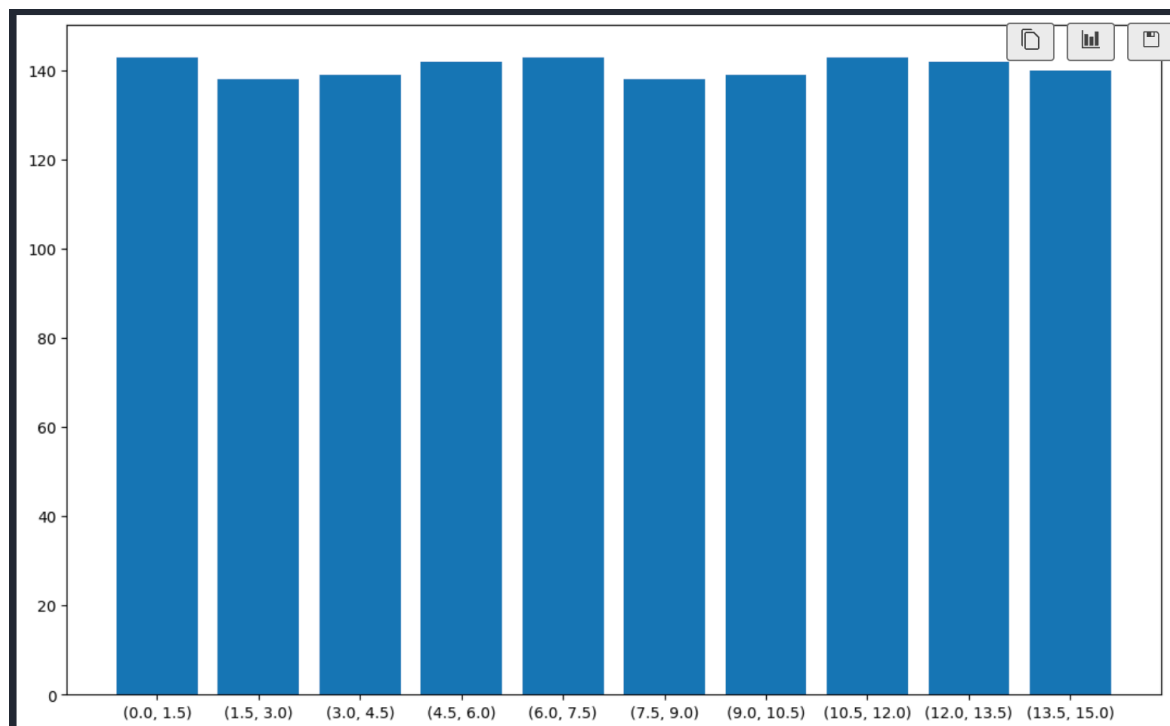
$\chi^2$  емпіричне: 0.7414786515411623

$\chi^2$  критичне: 3.841458820694124

Емпіричне < критичне – гіпотезу  $H_0$  приймаємо

**Друга задача:**

	$x_i$	$n_i$
0	(0.0, 1.5)	143
1	(1.5, 3.0)	138
2	(3.0, 4.5)	139
3	(4.5, 6.0)	142
4	(6.0, 7.5)	143
5	(7.5, 9.0)	138
6	(9.0, 10.5)	139
7	(10.5, 12.0)	143
8	(12.0, 13.5)	142
9	(13.5, 15.0)	140



Гіпотеза  $H_0$  – рівномірний закон розподілу

	xi	ni	avg_interval	pi
0	(0.0, 1.5)	143	0.75	0.1
1	(1.5, 3.0)	138	2.25	0.1
2	(3.0, 4.5)	139	3.75	0.1
3	(4.5, 6.0)	142	5.25	0.1
4	(6.0, 7.5)	143	6.75	0.1
5	(7.5, 9.0)	138	8.25	0.1
6	(9.0, 10.5)	139	9.75	0.1
7	(10.5, 12.0)	143	11.25	0.1
8	(12.0, 13.5)	142	12.75	0.1
9	(13.5, 15.0)	140	14.25	0.1

	xi	ni	pi
0	(0.0, 1.5)	143	0.1
1	(1.5, 3.0)	138	0.1
2	(3.0, 4.5)	139	0.1
3	(4.5, 6.0)	142	0.1
4	(6.0, 7.5)	143	0.1
5	(7.5, 9.0)	138	0.1
6	(9.0, 10.5)	139	0.1
7	(10.5, 12.0)	143	0.1
8	(12.0, 13.5)	142	0.1
9	(13.5, 15.0)	140	0.1

$\chi^2$  емпіричне: 0.2850035536602701

$\chi^2$  критичне: 8.31117091082631

Емпіричне < критичне – гіпотезу  $H_0$  приймаємо



**Висновок:**

Виконуючи дане індивідуальне завдання, я детальніше ознайомилась з темою перевірка гіпотез про закон розподілу випадкової величини методом Пірсона. Навчилась обчислювати  $\chi^2$  квадрат емпіричне та критичне та пригадала, як обчислювати ймовірності для різних законів розподілу.

Виконуючи це індивідуальне завдання, я отримав більше розуміння теми перевірки гіпотез щодо розподілу випадкової величини за допомогою методу Пірсона. Отримав навички у обчисленні емпіричного та критичного значень статистики  $\chi^2$ -квадрат, а також пригадав, як обчислювати ймовірності для різних розподілів.