ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені ІВАНА ФРАНКА

Факультет прикладної математики та інформатики

Кафедра дискретного аналізу

та інтелектуальних систем

Звіт з дисципліни

“Теорія ймовірності та математична статистика”

ІНДИВІДУАЛЬНЕ ЗАВДАННЯ № 2

Варіант 6

Виконала:

[redacted]

Оцінка

Перевірила:

Квасниця Г. А.

2021

Постановка задачі

1. Зчитати дані з текстового файлу, побудувати полігон 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 χ2.

Короткі теоретичні відомості

Для перевірки гіпотези за критерієм Пірсона виконуємо наступні кроки:

1. Статистичні дані записуємо у вигляді інтервального статистичного розподілу
2. Визначаємо теоретичні ймовірності попадання значень відповідно до закону розподілу
3. Обчислюємо емпіричне значення критерію узгодження Пірсона

$$\chi^2_{емп} = \sum_{i=1}^{m} \frac{(n_i - np_i)^2}{np_i}.$$

4. За даним рівнем значущості і кількістю ступенів вільності k = m - s - 1, де m - число часткових інтервалів, а s - число оцінюваних параметрів знаходимо критичну точку за таблицею критичний значень розподілу χ2.
5. Порівнюємо емпіричне та критичні значення. Якщо емпіричне більше за критичне, то гіпотезу відхиляють. Якщо емпіричне менше за критичне, гіпотезу приймають.

Гіпотетичні закони розподілу:

1. Біномний закон розподілу.
   Імовірності обчислюються
   $p_i = P(\xi = i) = C_N^i p^i (1-p)^{N-i}$, де $p = \frac{\bar{x}}{N}$, якщо невідоме
2. Закон розподілу Пуассона.
   $: P(\xi = i) = e^{-\lambda} \frac{\lambda^i}{i!}$, де $\lambda > 0$, якщо невідоме

3. Рівномірний закон розподілу.

   Функція має вигляд $$F(x)=\begin{cases} 0, & x<a \\ \dfrac{x-a}{b-a}, & a\le x<b \\ 1, & x\ge b \end{cases},$$
   де a та b можуть бути оцінені на основі даних вибірки

   $a=\overline{x}-\sqrt{3}s$, $b=\overline{x}+\sqrt{3}s$

4. Показниковий закон розподілу.

   Функція має вигляд $$F(x)=\begin{cases} 0, & x<0 \\ 1-e^{-\lambda x}, & x\ge 0 \end{cases}$$, де $\lambda=\dfrac{1}{\overline{x}}$

5. Нормальний закон розподілу.

   $$F(x)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{(t-a)^2}{2\sigma^2}}dt$$, де

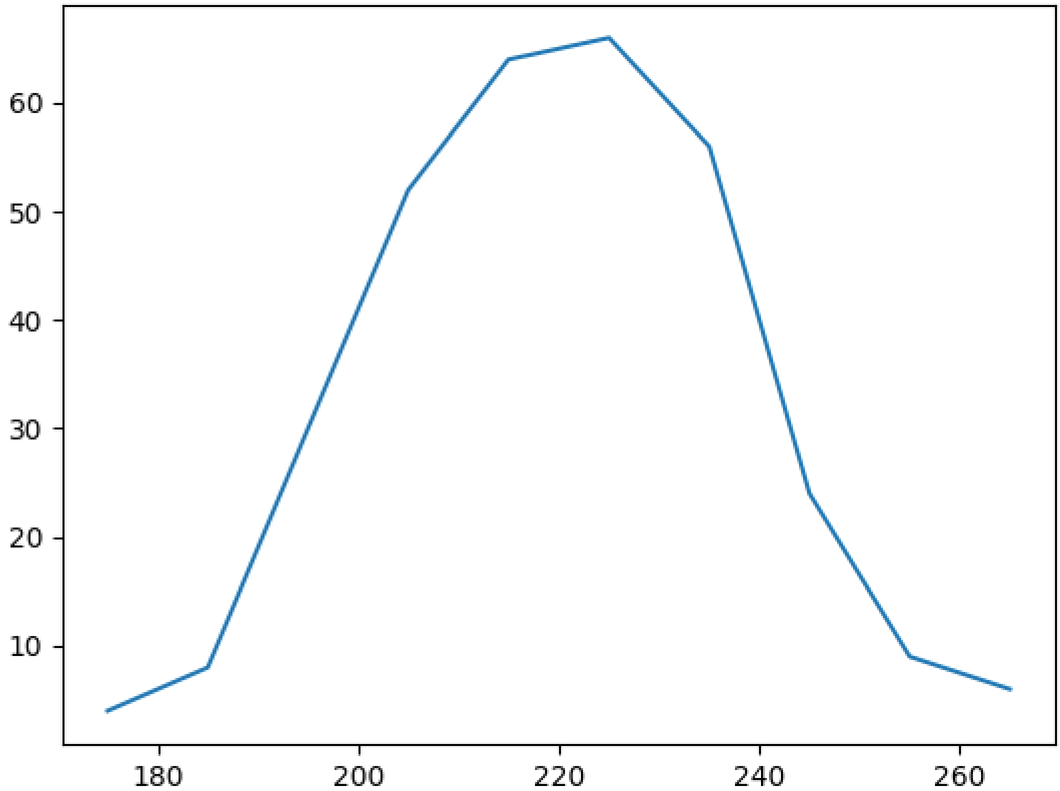
   параметри розподілу $a=\overline{x},\ \sigma=s$ оцінюються на основі вибірки.

Програмна реалізація

Реалізувала програму мовою Python в середовищі PyCharm. Взяла вибірку для задачі 1 та задачі 2 відповідно до свого варіанту 6. На основі вибірки побудувала полігон частот та обчислила потрібні числові характеристики. Перевірила гіпотезу за критерієм Пірсона. Для реалізації використовувала бібліотеки matplotlib, math та scipy.stats.

Задача 1. Інтервальний розподіл.

170 - 180	180 - 190	190 - 200	200 - 210	210 - 220	220 - 230	230 - 240	240 - 250	250 - 260	260 - 270
4	8	30	52	64	66	56	24	9	6



p[0] = 0.01251990884057308
p[1] = 0.03371028100236456
p[2] = 0.08420606409898296
p[3] = 0.15522563040207915
p[4] = 0.21119659979117744
p[5] = 0.21210308408824707
p[6] = 0.15723313783380544
p[7] = 0.08602921864660562
p[8] = 0.03473674855719744

p[9] = 0.013039326738967238

np[0] = 3.9938509201428123
np[1] = 10.753579639754296
np[2] = 26.861734447575564
np[3] = 49.51697609826325
np[4] = 67.37171533338561
np[5] = 67.66088382415082
np[6] = 50.15737096898393
np[7] = 27.443320748267194
np[8] = 11.081022789745983
np[9] = 4.159545229730549

np[9] < 10 => np[9] + np[8]
np[0] < 10 => np[0] + np[1]

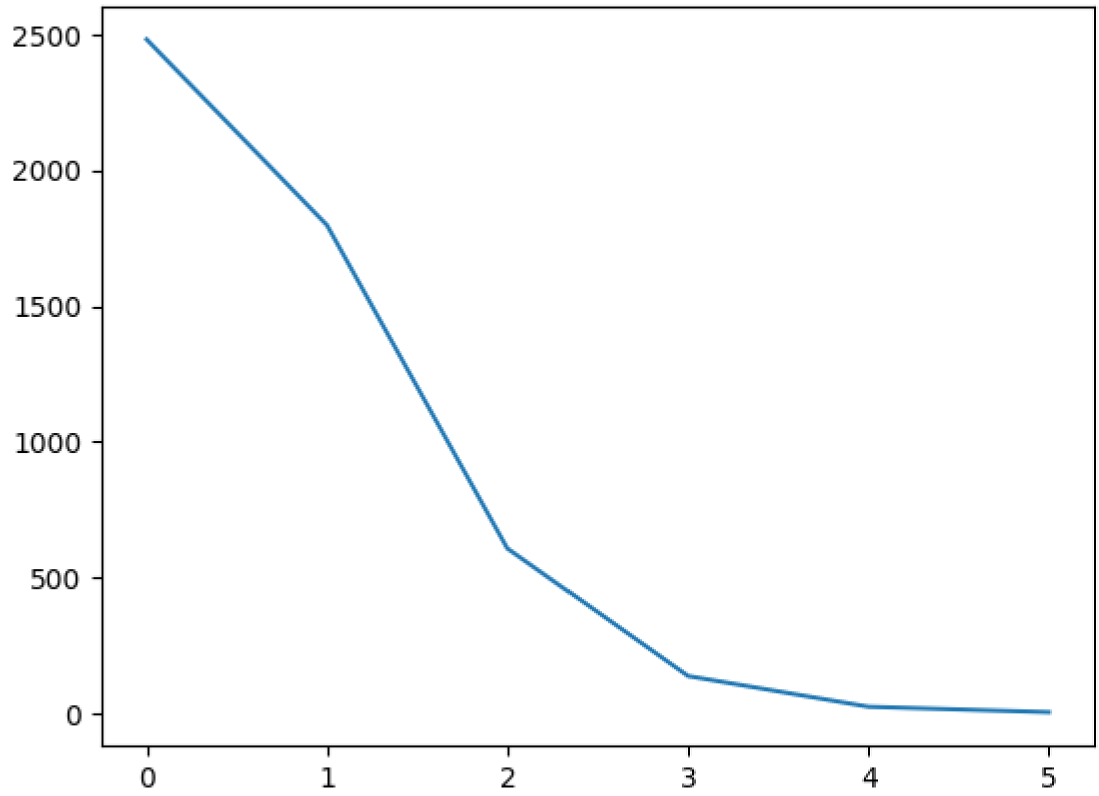
np[1] = 14.747430559897108
np[2] = 26.861734447575564
np[3] = 49.51697609826325
np[4] = 67.37171533338561
np[5] = 67.66088382415082
np[6] = 50.15737096898393
np[7] = 27.443320748267194
np[8] = 15.240568019476532

df = 8 - 2 - 1 = 5

Введіть рівень значущості: 0.01
Емпіричне: 2.328927263583102
Критичне: 15.086272469388991
Критичне значення більше емпіричного.
--- Гіпотезу приймаємо. ---

Задача 2. Дискретний розподіл.

0	1	2	3	4	5
2483	1799	608	138	25	5



З графічного зображення можна припустити, що статистичний матеріал розподілений за законом Пуассона. Перевіримо, використовуючи реалізовану програму.

p[0] = 0.49527145713040543
p[1] = 0.34800212705446043
p[2] = 0.12226172000311906
p[3] = 0.02863570270799296
p[4] = 0.005030213890085358
p[5] = 0.0007068952220388835

np[0] = 2505.0830301655906
np[1] = 1760.1947586414608
np[2] = 618.3997797757762

np[3] = 144.83938429702837
np[4] = 25.442821856051744
np[5] = 3.5754760330726727
np[5] < 10 => np[5] + np[4]

df = 5 - 1 - 1 = 3

Введіть рівень значущості: 0.05
Емпіричне: 2.1044552508708723
Критичне: 7.814727903251179
Критичне значення більше емпіричного.
--- Гіпотезу приймаємо. ---

Висновки

При виконанні індивідуального завдання 2 я навчилась на основі графічного представлення формулювати гіпотезу про закон розподілу досліджуваної ознаки генеральної сукупності, оцінювати на основі даних вибірки та для заданого користувачем рівня значущості перевіряти сформульовану гіпотезу за критерієм Пірсона ($\chi 2$).