

Міністерство освіти і науки України
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «КИЄВО-МОГИЛЯНСЬКА АКАДЕМІЯ»
Факультет інформатики



Протокол до лабораторної роботи №1
З дисципліни „Математичні методи машинного навчання ”

Виконала студентка 3 р.н.
спеціальності “Прикладна математика”
Антошина Катерина

1. Завдання лабораторної роботи

- 1) Сформувати тестову вибірку зображень з вихідного пакету;
- 2) Для зеленого каналу кольору тестових зображень обчислити наступні характеристики розподілу значень яскравості пікселів:
 - a) Математичне очікування та дисперсію;
 - b) Медіану та інтерквартильний розмах;
 - c) Коефіцієнти асиметрії та ексцесу (нормалізований);
 - d) Гістограму значень яскравості пікселів (нормалізовану);
- 3) Провести обробку отриманих гістограм:
 - a) Провести апроксимацію гістограм з використанням імовірнісних розподілів:
 - i) Нормального (гаусового) розподілу;
 - ii) Розподілу Лапласа;
 - iii) Розподілу Стюдента;
 - iv) Бета розподілу;
 - b) Для кожної гістограми визначити найкращий тип апроксимуючого розподілу за критерієм мінімізації середньо квадратичного відхилення;
 - c) Побудувати розподіл отриманих апроксимацій за видом апроксимуючого розподілу;
- 4) Побудувати багатовимірні гаусові моделі:
 - a) Сформувати вектори параметрів розподілу значень яскравості пікселів тестових зображень;
 - i) Математичне очікування;
 - ii) Математичне очікування та дисперсія;
 - iii) Математичне очікування, дисперсія та коефіцієнт асиметрії;
 - iv) Математичне очікування, дисперсія, коефіцієнти асиметрії та ексцесу;
 - b) Визначити параметри багатовимірних гаусових моделей для кожної групи векторів, відобразити їх графічно та у вигляді таблиць.
- 5) Підготувати звіт за отриманими результатами лабораторної роботи.

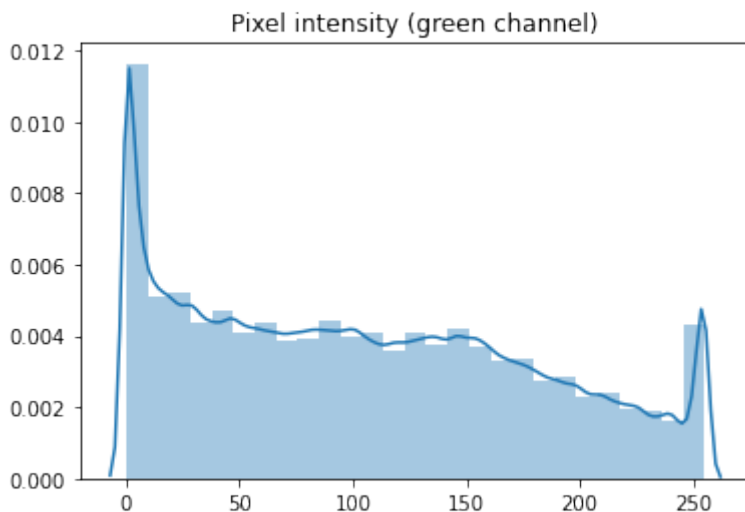
2. Порядок виконання роботи та отримані результати

1) Вибірка була сформована з використанням функції `random()`, в основі якої лежить генератор Мерсена, зі стартовим значенням 1. Згенеровані 250 чисел є номерами картинок, які витягуються з архіва `mirflickr25k.zip` та зберігаються в нову теку на Google Drive, із назвою `mirflickr`. Після того, зображення конвертуються в масиви `numpy` і зберігаються в масиві `imgs`.

2) Характеристики розподілу значень яскравості пікселів для зеленого каналу кольору тестових зображень:

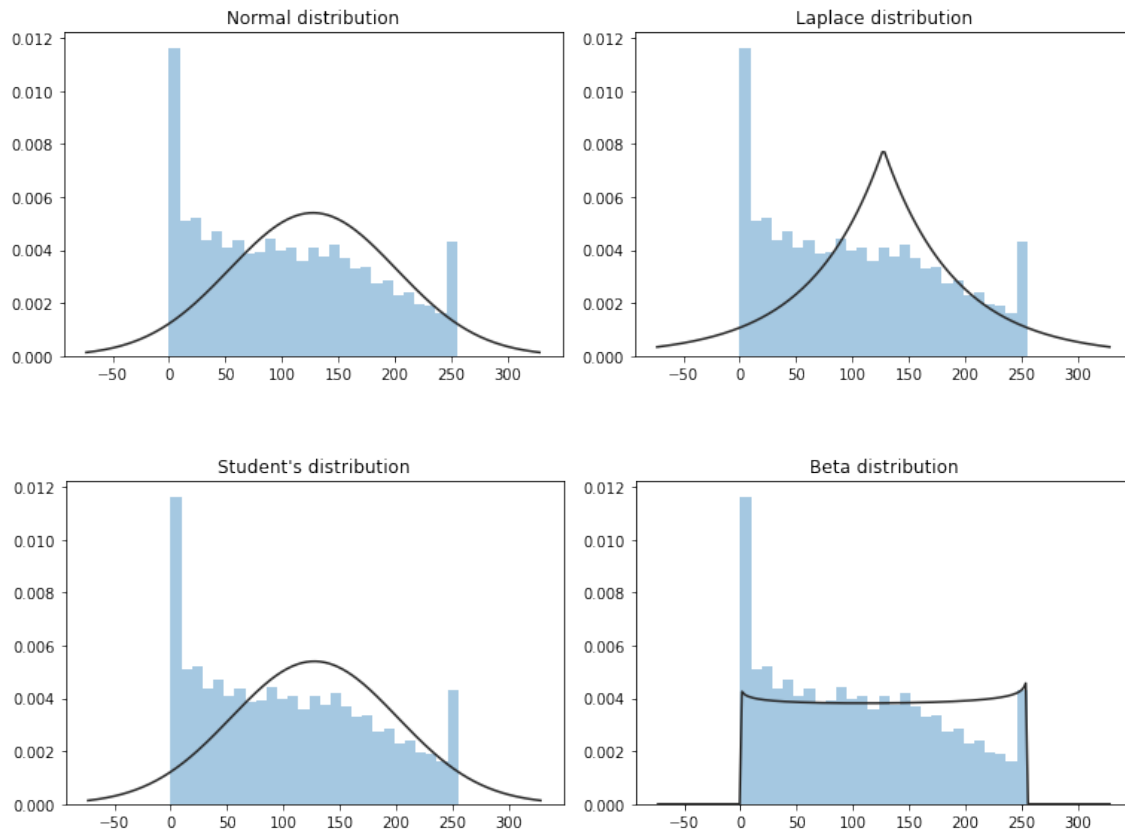
Математичне очікування	104.32698693938991
Дисперсія	5572.828828999129
Медіана	97.0
Інтерквартильний розмах	123.0
Коефіцієнти асиметрії	0.32621085086243945
Коефіцієнт ексцесу	-3.9947769281332453

Гістограма значень яскравості пікселів (нормалізована):

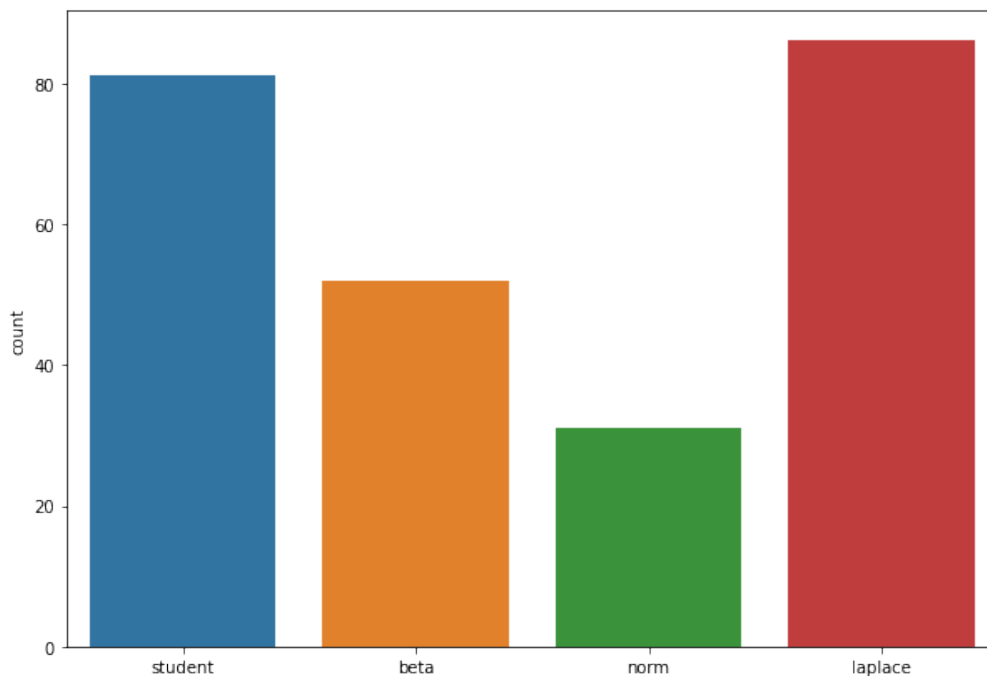


3) Обробка отриманих гістограм.

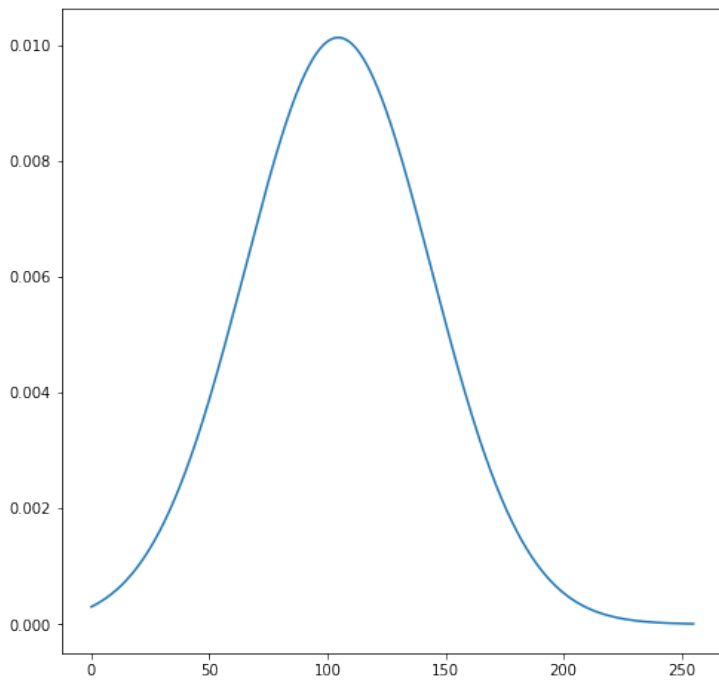
Апроксимація нормалізованої гістограми значень яскравості пікселів кривими нормального, Лапласового, Стюдента та бета-розподілу:



По всій вибірці було здійснено визначення найкращого типу апроксимуючого розподілу за критерієм мінімізації середньо квадратичного відхилення. Відповідний графік, що ілюструє розподіл отриманих апроксимацій, наведено нижче. Бачимо, що “найпопулярнішими” є розподіли Стюдента та Лапласа.

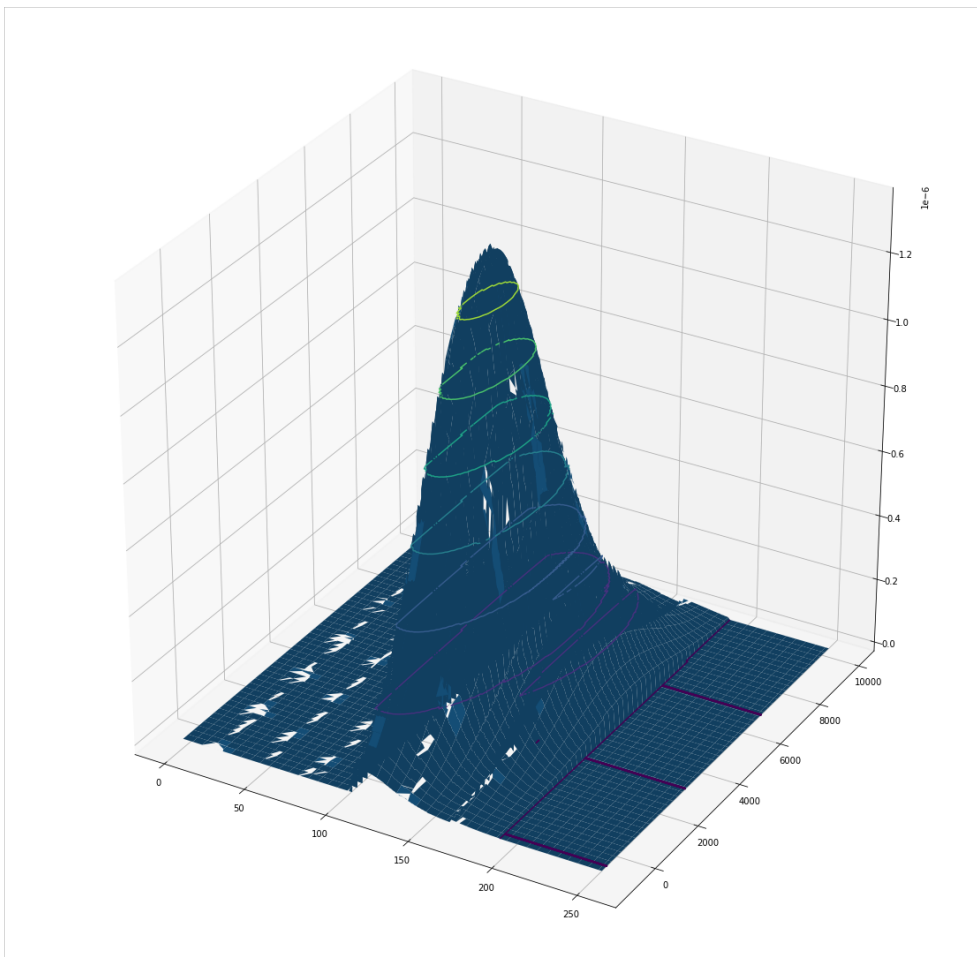


4) Побудова багатовимірних Гаусових моделей.
Графічне зображення розподілів наведено нижче.
i) математичне очікування:



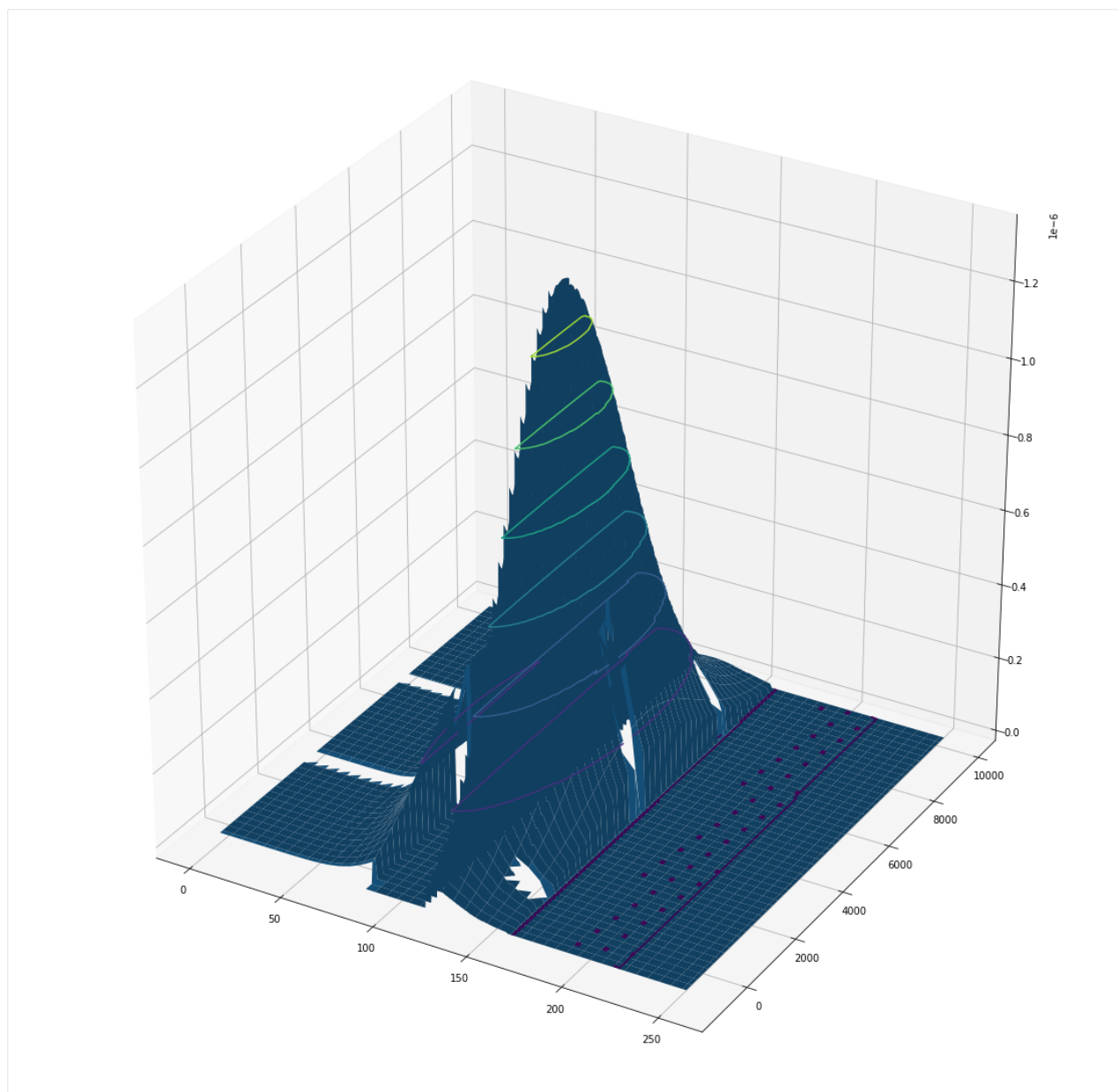
mean = 104.55681374933789
cov = 1552.6609550619742

ii) математичне очікування та дисперсія:



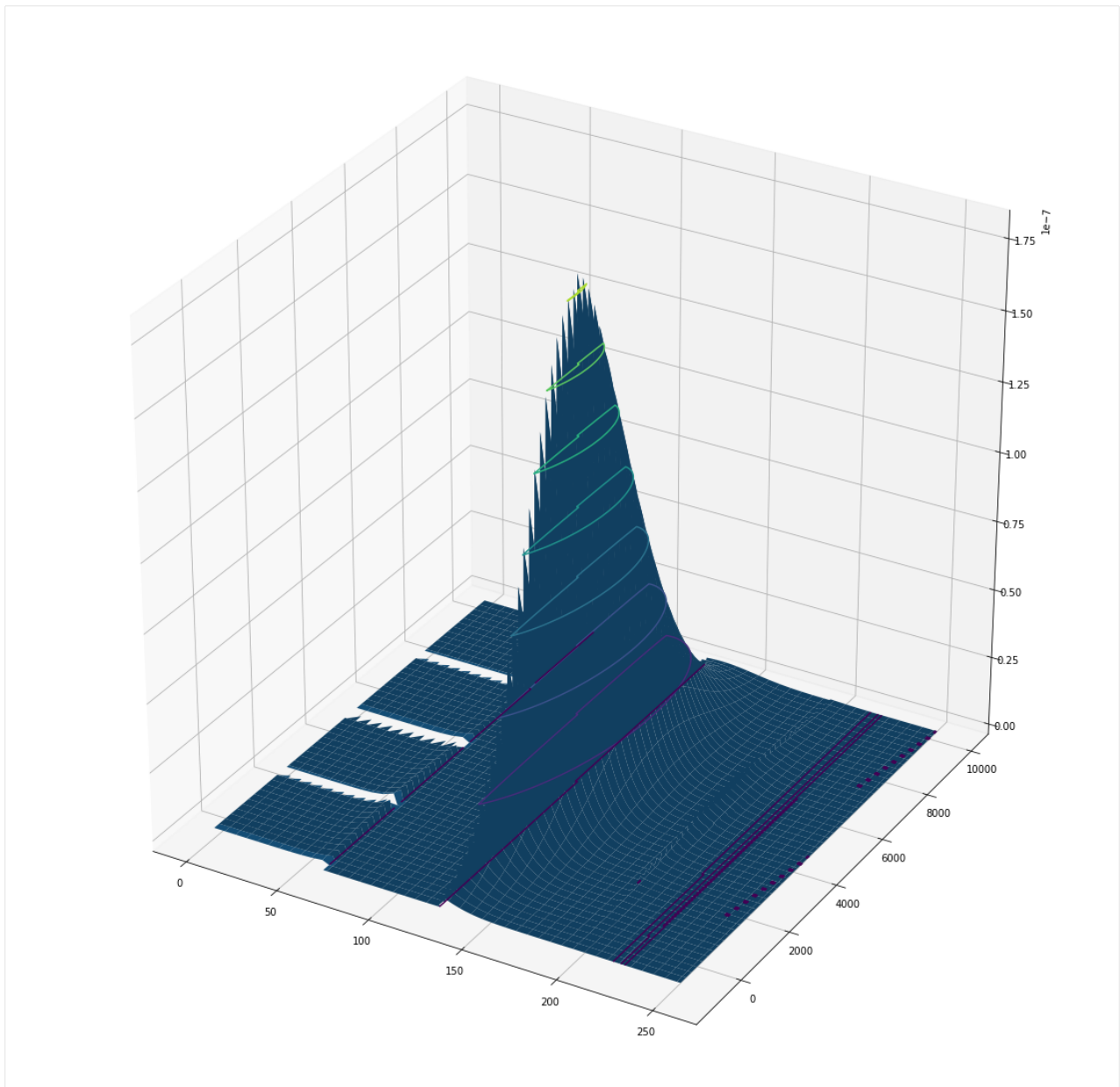
mean = [104.55681375
4041.70367757]
cov = [[1.55266096e+03
1.97717851e+04]
[1.97717851e+04
4.87940131e+06]]

iii) математичне очікування, дисперсія та коефіцієнт асиметрії:



```
mean = [1.04556814e+02 4.04170368e+03 3.72161389e-01]
cov = [[ 1.55266096e+03 1.97717851e+04 -3.33604262e+01]
 [ 1.97717851e+04 4.87940131e+06 -3.61346551e+02]
 [-3.33604262e+01 -3.61346551e+02 9.82485690e-01]]
```

iv) математичне очікування, дисперсія, коефіцієнт асиметрії та ексцесу:



```
mean = [1.04556814e+02 4.04170368e+03 3.72161389e-01 3.93080073e-01]
cov = [[ 1.55266096e+03 1.97717851e+04 -3.33604262e+01 -6.48548763e+01]
 [ 1.97717851e+04 4.87940131e+06 -3.61346551e+02 -3.17102390e+03]
 [-3.33604262e+01 -3.61346551e+02 9.82485690e-01 2.44847517e+00]
 [-6.48548763e+01 -3.17102390e+03 2.44847517e+00 1.53996440e+01]]
```