

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



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

Виконав  
студент 4 курсу  
факультету інформатики  
Іванюк-Скульський Б. В.

## 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

Робота була виконана на мові Python в форматі Jupyter Notebook та за допомогою середовища Jupyter.

- 1) Для формування вибірки було використано модуль random відповідним діапазоном значень.
- 2) Для обчислення загальновибірових характеристик яскравості пікселів зеленого кольору було сформовано масив, отриманий шляхом «сплюснення» початкового зображення зеленого каналу.

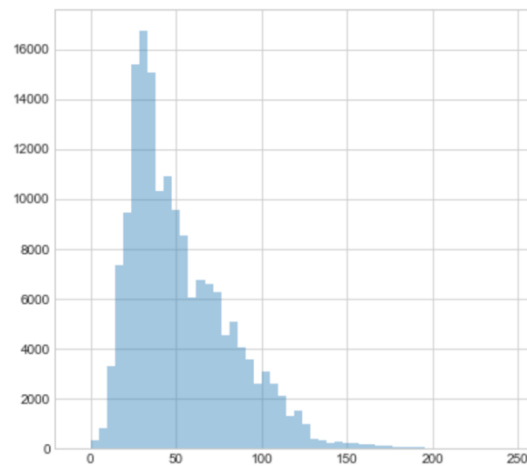


Рис.1. Гістограма яскравості пікселів зеленого каналу

Числові характеристики даного розподілу мають наступний вигляд:

Математичне очікування	52.68
Дисперсія	895.4
Медіана	45.0
Інтерквартильний розмах	40.0
Коефіцієнт ексцесу	1.393
Коефіцієнт асиметрії	1.12

Бачимо, що мат. Очікування більше за медіану, правий хвіст розподілу більший за лівий, що підтверджується додатнім коефіцієнтом асиметрії.

Розподіл є більш гострим ніж нормальний розподіл, про що свідчить коефіцієнту ексцесу. Маємо відносно велику дисперсію.

### 3) Обробка та апроксимація отриманих гістограм

Для вибірки з 250 зображень розподіл отриманих апроксимацій за видом апроксимуючого розподілу (за критерієм мінімізації середньоквадратичного відхилення):

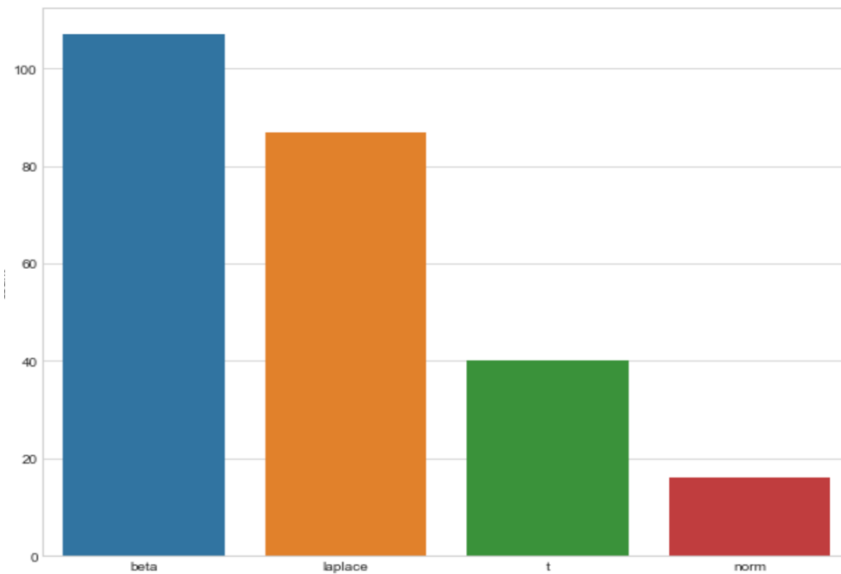


Рис. 2. Розподіл отриманих апроксимацій за видом оптимального апроксимуючого розподілу

Бачимо, що отриманий розподіл має виражений головний елемент – бета розподіл. Трохи менш оптимальним є розподіл Лапласа і далі за графіком: Стюдента та нормальний.

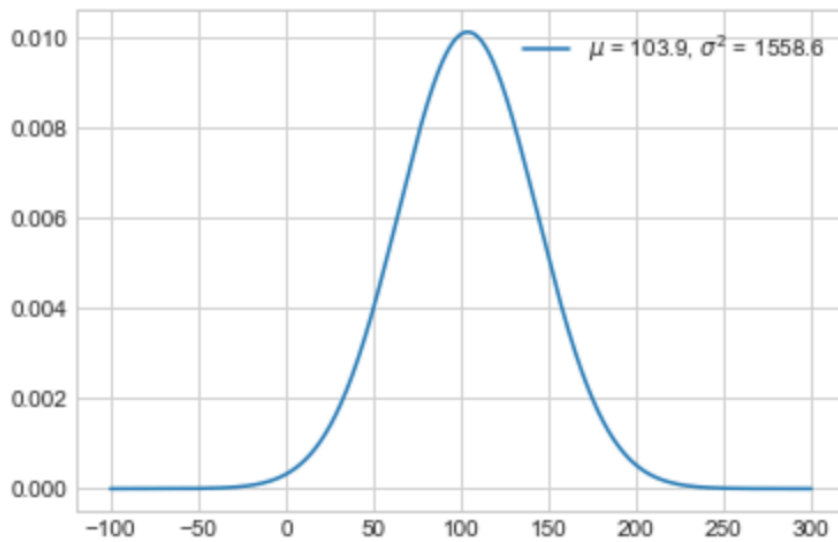
#### 4) Багатовимірні гаусові моделі

Сформувавши вектори параметрів розподілу значень яскравості пікселів для тестових зображень, ми сформували наступні гаусові моделі:

##### 1. Математичне очікування

Вектор середнього	103.92
Матриця коваріації	1558.55

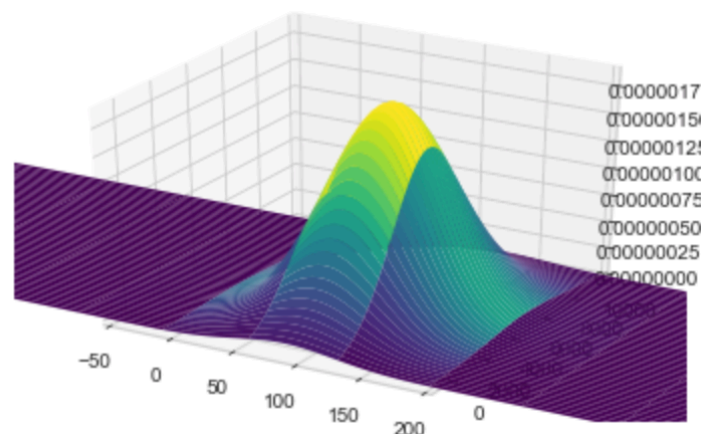
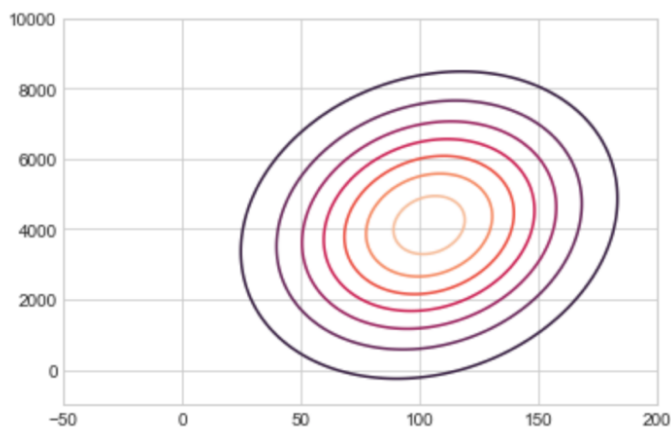
Табл.1.Параметри багатовимірної Гаусової моделі побудованій на мат. очікуванні



##### 2. Математичне очікування та дисперсія

Вектор середнього	103.92 ; 4116.82
Матриця коваріації	$\begin{bmatrix} 1558.55 & 14806.59 \\ 14806.59 & 4732305.20 \end{bmatrix}$

Табл.2.Параметри багатовимірної Гаусової моделі побудованій на мат. очікуванні та дисперсії



### 3. Математичне очікування, дисперсія та коефіцієнт асиметрії

Вектор середнього	103.92 ; 4116.82; 0.3527
Матриця коваріації	[[ 1558.55    14806.59    -1.0807    ] [ 14806.59    4732305.20    -3187.02    ] [ -1.08        -3187.02.        13.9031    ]]

### 4. Математичне очікування, дисперсія, коефіцієнти асиметрії та ексцесу

Вектор середнього	103.92 ; 4116.82; 0.3527; 0.3236
Матриця коваріації	[[ 1558.55    14806.59    -1.0807    -33.00    ] [ 14806.59    4732305.20    -3187.02    -148.73    ] [ -1.08        -3187.02.        13.9031    0.34.    ] [ -33.00       -148.73        0.34        0.9171    ]]

### **3. Висновки**

З отриманих даних MIRFlickr відібрав випадковим чином 250 картинок. Відібрав зелений канал та провів розрахунок наступних характеристик: математичного сподівання, дисперсії, медіани, інтерквартильного розмаху, коефіцієнтів асиметрії та ексцесу. Потім візуалізував гістограми значень яскравості пікселів та апроксимував за допомогою одного із зазначених розподілів: Нормального, Лапласа, Стюдента, Бета за критерієм мінімізації середньо квадратичного відхилення. Останнім кроком побудував багатовимірні гаусові моделі з різними вхідними векторами: (Математичне сподівання), (Математичне сподівання; Дисперсія), (Математичне сподівання; Дисперсія; Коефіцієнт асиметрії), (Математичне сподівання; Дисперсія; Коефіцієнт асиметрії; Коефіцієнт Ексцесу). Та вивів отримані параметри моделі для кожного з представлених векторів