**Протокол до лабораторної роботи №1**

**З дисципліни «Математичні методи машинного навчання»**

Виконав

студент 4 курсу

факультету інформатики

Таран Д. Г.

1. Завдання лабораторної роботи

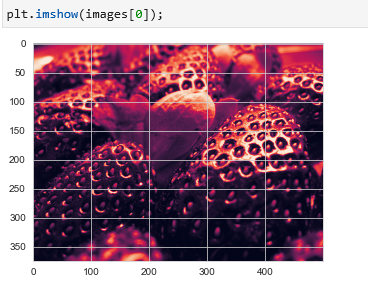
1. Сформувати тестову вибірку зображень з вихідного пакету;
2. Для зеленого каналу кольору тестових зображень обчислити наступні характеристики розподілу значень яскравості пікселів:
   1. Математичне очікування та дисперсію;
   2. Медіану та інтерквартильний розмах;
   3. Коефіцієнти асиметрії та ексцесу (нормалізований);
   4. Гістограму значень яскравості пікселів (нормалізовану);
3. Провести обробку отриманих гістограм:
   1. Провести апроксимацію гістограм з використанням імовірнісних розподілів:
      1. Нормального (гаусового) розподілу;
      2. Розподілу Лапласа;
      3. Розподілу Стьюдента;
      4. Бета розподілу;
   2. Для кожної гістограми визначити найкращий тип апроксимуючого розподілу за критерієм мінімізації середньо квадратичного відхилення;
   3. Побудувати розподіл отриманих апроксимацій за видом апроксимуючого розподілу;
4. Побудувати багатовимірні гаусові моделі:
   1. Сформувати вектори параметрів розподілу значень яскравості пікселів тестових зображень;
      1. Математичне очікування;
      2. Математичне очікування та дисперсія;
      3. Математичне очікування, дисперсія та коефіцієнт асиметрії;
      4. Математичне очікування, дисперсія, коефіцієнти асиметрії та ексцесу;
   2. 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. Виконання роботи та отримані результати

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

1. **Формування тестової вибірки зображень з вихідного пакету.**

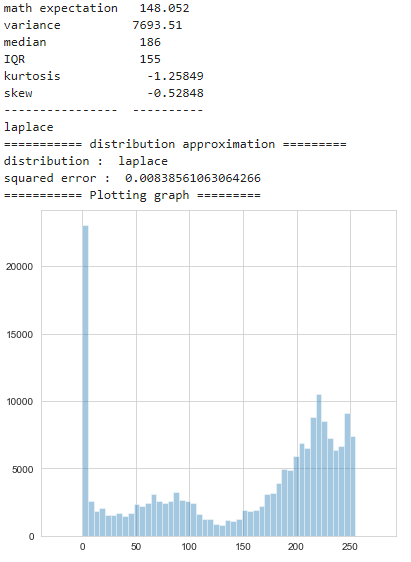
Для формування вибірки було використано модуль random з заданим відповідним діапазоном значень. У якості прикладу було виведено зображення.

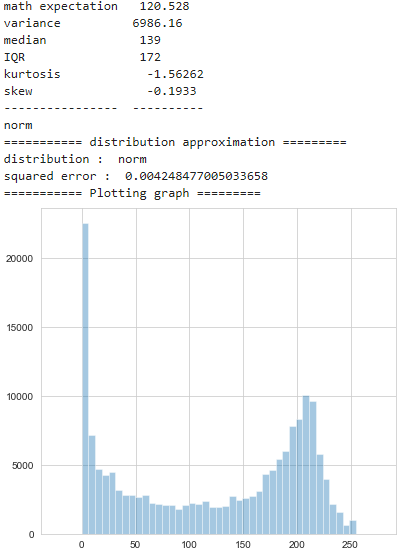


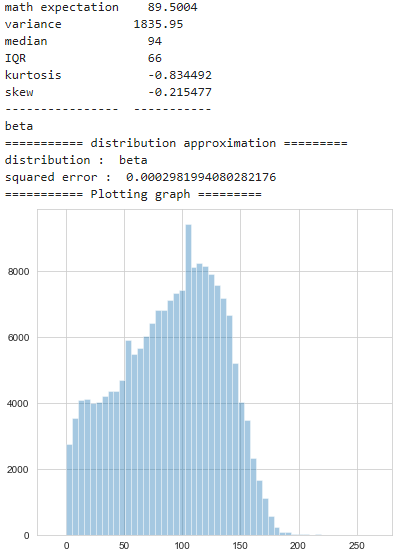
1. **Обчислення характеристик розподілу значень яскравості пікселів зеленого каналу кольору.**

Для обчислення характеристик розподілу зеленого каналу конкретного зображення було сформовано масив значень яскравості усіх пікселів даного зображення.

Для прикладу взяли 3 картинки. На основі цього отримали такі діаграми розподілів:



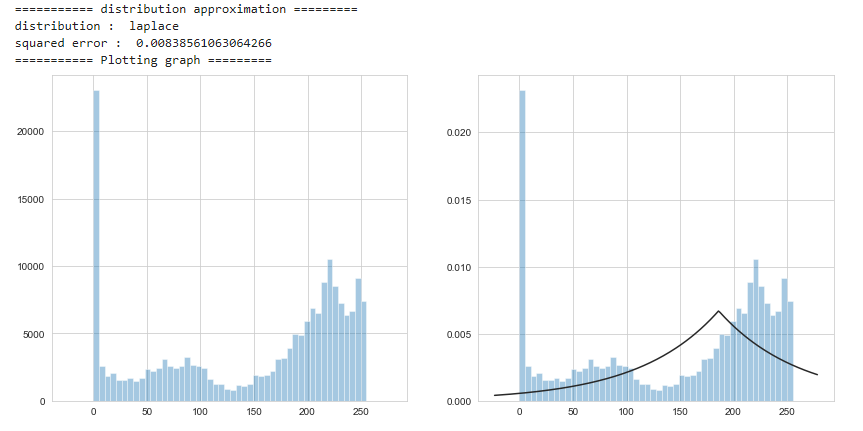




1. **Обробка та апроксимація отриманих гістограм.**

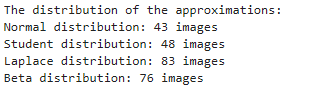
Були проведені апроксимації отриманих гістограм зображень з використанням імовірнісних розподілів: нормального(гаусового), Лапласа, Стьюдента та бета.

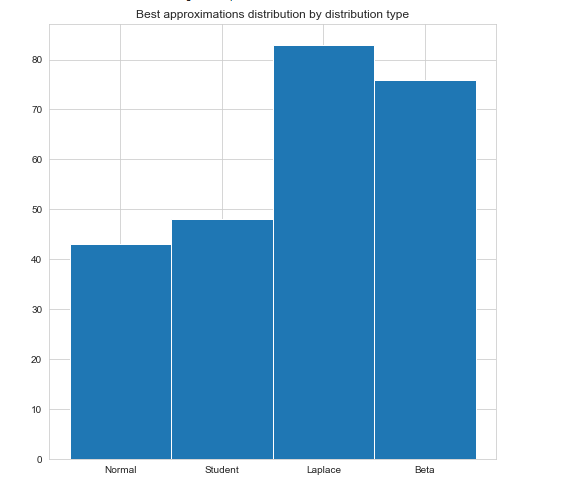
Для зображення з прикладу отримані апроксимації мали такий вигляд:



Далі для кожного зображення вибірки були знайдені дані апроксимації і з них обрані розподіли, за якими апроксимація була найточніша за критерієм мінімізації середньоквадратичного відхилення.

Після опрацювання всіх 250 зображень вибірки маємо наступний розподіл отриманих найкращих апроксимацій за видом апроксимуючого розподілу:





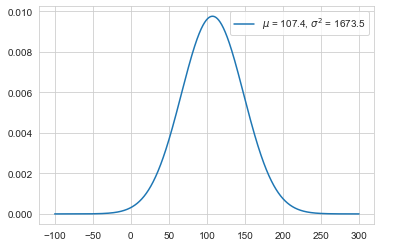
Бачимо, що найкраще з апроксимацією зображень цієї вибірки впорались бета-розподіл та розподіл Лапласа, набагато гірше – Стьюдента, і найгірше – нормальний розподіл.

1. **Багатовимірні гаусові моделі**

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

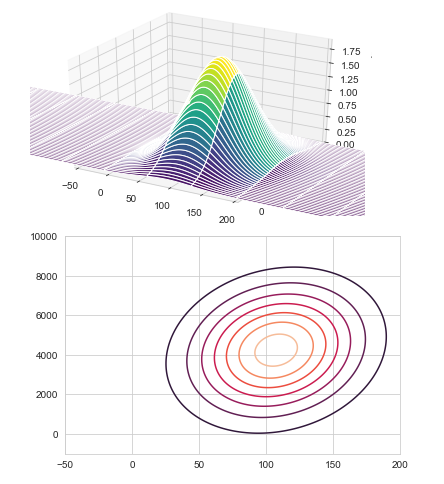
1. Математичне очікування

|  |  |
| --- | --- |
| Вектор середнього | 107.4 |
| Матриця коваріації | 1673.5 |



1. Математичне очікування та дисперсія

|  |  |
| --- | --- |
| Вектор середнього | (107.4, 4243.4) |
| Матриця коваріації | [[ 1673.53594569 13880.98473311] [ 13880.98473311 4366445.12780987]] |



1. Математичне очікування, дисперсія та коефіцієнт асиметрії

|  |  |
| --- | --- |
| Вектор середнього | (107.41545217506075 4243.408919072519 0.4199595228895673) |
| Матриця коваріації | [[ 1673.53594569 13880.98473311 -29.28626586] [ 13880.98473311 4366445.12780987 -3078.0709774 ] [ -29.28626586 -3078.0709774 9.5595654 ]] |

1. Математичне очікування, дисперсія, коефіцієнти асиметрії та ексцесу

|  |  |
| --- | --- |
| Вектор середнього | (107.41545217506075 4243.408919072519 0.4199595228895673 0.27798516994334505) |
| Матриця коваріації | [[ 1673.53594569 13880.98473311 -29.28626586 -38.8191856 ] [ 13880.98473311 4366445.12780987 -3078.0709774 -91.67834187] [ -29.28626586 -3078.0709774 9.5595654 0.94267026] [ -38.8191856 -91.67834187 0.94267026 1.14124183]] |

3. Висновки

В ході лабораторної роботи ми проаналізували розподіли яскравості зеленого каналу пікселів випадково згенерованої вибірки зображень з пакету.

Апроксимація зображень вибірки заданими розподілами і мінімізація їх похибок за критерієм середньоквадратичного відхилення показала, які розподіли краще апроксимують отримані гістограми, а саме бета-розподіл та розподіл Лапласа. Найгірше ж зображення отриманої вибірки апроксимував нормальний розподіл.

Також по векторах параметрів розподілів яскравості зеленого каналу ми побудували багатовимірні гаусові моделі та визначили їхні вектори середнього та матриці коваріації.