МІНІСТЕРСТВО ОСВІТИ ТА НАУКИ УКРАЇНИ

ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені ІВАНА ФРАНКА

Кафедра дискретного аналізу

та інтелектуальних систем

**Індивідуальне завдання №3**

з курсу ''Теорія ймовірності та математична статистика''

Виконала:

студентка групи ПМі-21

Дудчак Валентина Юріївна

Оцінка

Перевірила:

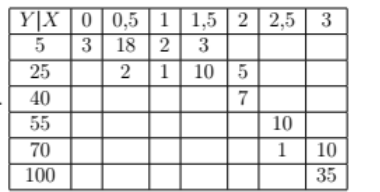
доц. Квасниця Г.А.

Львів 2024

**Постановка задачі:**

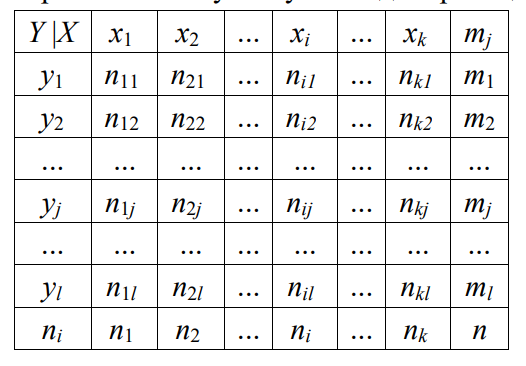
1. За даними кореляцiйної таблицi обчислити умовнi середнi yxi (i = 1,…,k ).
2. Побудувати поле кореляцiї, тобто нанести точки Mi(xi;yxi), i = 1,…,k , на координатну площину. На основi цього зробити припущення про вигляд функцiї регресiї (парабола, гiпербола i т.д.)
3. В залежностi вiд вигляду функцiї регресiї ((1), (3), (6) чи (9)) скласти вiдповiдну систему рiвнянь ((2), (5), (8) чи (11)). Розв’язати її i знайти невiдомi параметри вибраної функцiї регресiї.
4. Записати рiвняння кривої регресiї Y на X : yx = f(x) (з конкретною знайденою в пунктi 3 функцiєю регресiї f(x)) та побудувати її графік.
5. Обчислити дисперсiю (12) величини Y вiдносно кривої регресiї Y на X .
6. Визначити суму квадратiв вiдхилень δ2 умовних середнiх вiд значень функцiї регресiї за формулою (13).

**Варіант 4**

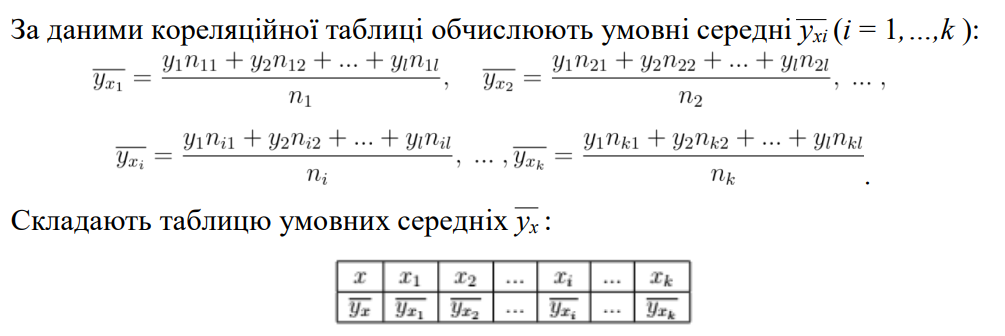


**Короткі теоретичні відомості:**

Нехай вибірка має вигляд:



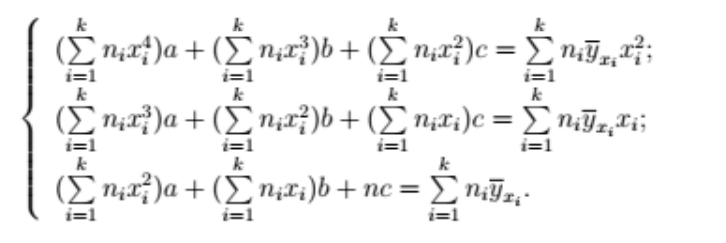
Де xi, yi — деякі значення, nij — частота появи події {X = xi, Y = yj}.



Для визначення вигляду функцiї регресiї будують точки (x;yx) (або (y;xy)) i за їх розмiщенням роблять висновок про приблизний вигляд функцiї регресiї.

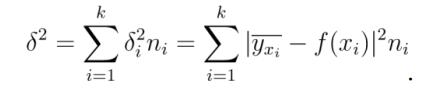
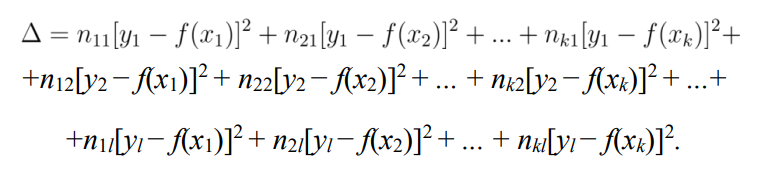
Наприклад, якщо функція має вигляд **параболічної** другого порядку:

Шукають коефіцієнти a, b, c за системою рівнянь:



В результаті одержують функцію кореляції f(x) = ax^2 + bx + c.

Також можна обчислити оцінку щільності кореляційного зв’язку: дисперсію та суму квадратів відхилень:



**Програмна реалізація:**

Для написання коду я використала мову програмування Python, середовище

Jupyter Notebook і наступні бібліотеки: pandas, numpy, matplotlib

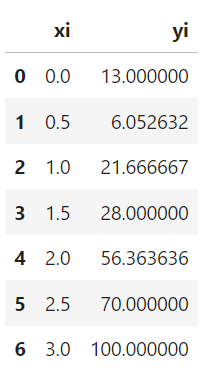
Спочатку я зчитую дані вибірки з файлу за допомогою методів pandas та показую їх у вигляді таблиці. Рахую умовні y-ки за допомогою окремої функції. Потім записую xi та умовні yi в два масиви і через них виводжу графік функції. Роблю припущення про те, яка це кореляція.

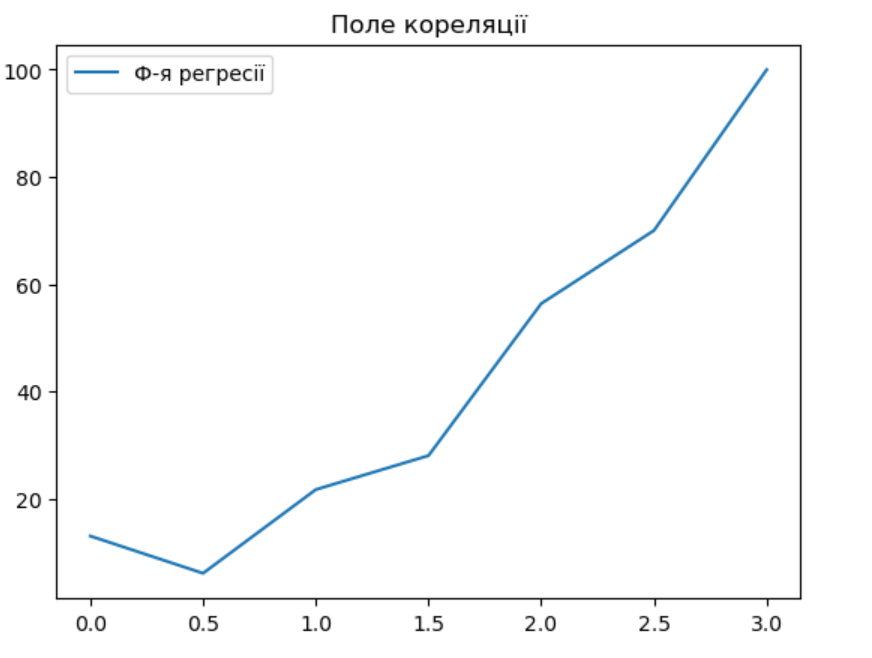
За допомогою методу Гауса розв’язую систему рівнянь з формули (2). Для цього в окремій функції спочатку обчислюю потрібні значення з формули, а потім знаходжу коефіцієнти a, b, c за допомогою методу трикутника. Маючи коефіцієнти, виводжу рівняння кореляції. Записую це рівняння як функцію та виводжу два графіки: початкову функцію та отриману.

В кінці оцінюю щільність кореляційного зв’язку, обчислюючи дисперсію величини Y вiдносно кривої регресiї Y на X та суму квадратів відхилень умовних середніх від функції регресії.

**Аналіз отриманих результатів:**

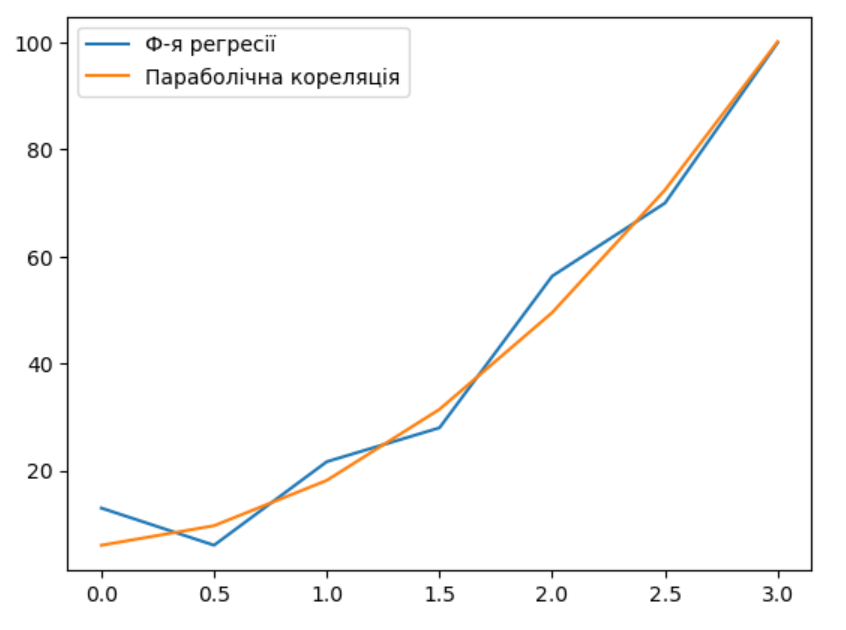
Умовні середні yi:





Припускаємо, що кореляція – параболічна







Значення дисперсії та відхилення досить великі, оскільки значення уі великі.

**Висновок:**

Виконуючи дане індивідуальне завдання, я детальніше ознайомилась з темою нелінійна регресія. Навчилась знаходити функцію кореляції, залежно від значень умовних yi, що буде наближена до початкової функції.