|  |  |  |  |
| --- | --- | --- | --- |
| МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені Тараса Шевченка  ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ  **Кафедра програмних систем і технологій**    Дисципліна  **«Ймовірнісні основи програмної інженерії»**  **Лабораторна робота № 1**  **«Центральні тенденції та міра дисперсії»** | | | |
| **Виконав:** | Полюхович Данило Богданович | **Перевірила**: | Вечерковська А.С. |
| Група | ІПЗ-23 | Дата перевірки |  |
| Форма навчання | денна | Оцінка |  |
| Спеціальність | 121 |
| 2022 | | | |

**Мета –** навчитись використовувати на практиці набуті знання про центральні тенденції та міри.

**Постановку задачі:**

Вхідні дані: файл із М+1 рядків. Перший рядок містить число: M, де M —кількість переглядів і-того фільму.

Вихідні дані: запис у файл обчисленної інформації

Завдання

1. Побудувати таблицю частот та сукупних частот для переглянутих фільмів.

Визначити фільм, який був переглянутий частіше за інші.

2. Знайти Моду та Медіану заданої вибірки.

3. Порахувати Дисперсію та Середнє квадратичне відхилення розподілу.

4. Побудувати гістограму частот для даного розподілу.

5. Зробити висновок з вигляду гістограми, про закон розподілу.

**Математична модель:**

Для обчислення потрібних нам величин виведемо всі формули:

– середнє значення

– медіана для непарного

– медіана для парного

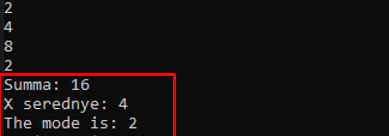
– Мода елемент із найвищою частотою. Якщо немає елемента в даних, що зустрічаються більше одного разу, тоді дані не мають моди. Якщо кілька елементів мають однакові значення частоти, яка перевищує всі інші частоти, то всі такі елементи є модами.

– об’єктивна дисперсія

– стандартне квадратичне відхилення

**Випробування алгоритму:**

Для перевірки алгоритму на суму, середнє значення та моду, використаймо наш алгоритм на тестових даних(2, 4, 8, 2) і подивимось чи правильно працює алгоритм:



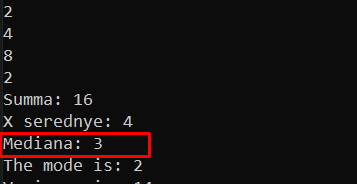
Алгоритм відпрацював без помилок.

Для перевірки медіани використаймо ті самі дані, але в нас буде два варінта: перший – парна вибірка, другий – непарна. Результат буде порівнювати з результатом, що порахуємо через вище зазначені формули:

Для парного варіанту (2,4,8,2) :

Оскільки масив чисел ми сортуємо перед обчисленням, вибірка стане 2,2,4,8

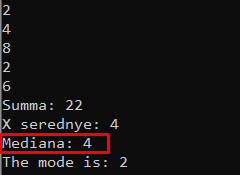
= (2 + 4) / 2 = 3



Для непарного варіанту (2,4,8,2,6) :

Оскільки масив чисел ми сортуємо перед обчисленням, вибірка стане 2,2,4,6,8

= 2.5 округлюємо до 3 – елемент під номером три – 4.

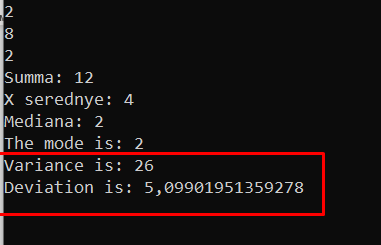


Як ми бачимо, результати збіглися.

Для дисперсії та віхлилення зробимо теж саме. Візьмемо вибірку: 2,8,2

= 26 – дисперсія

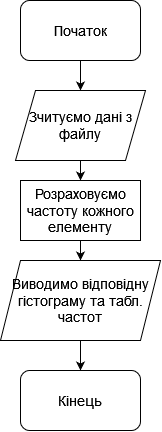
– стандартне квадратичне відхилення



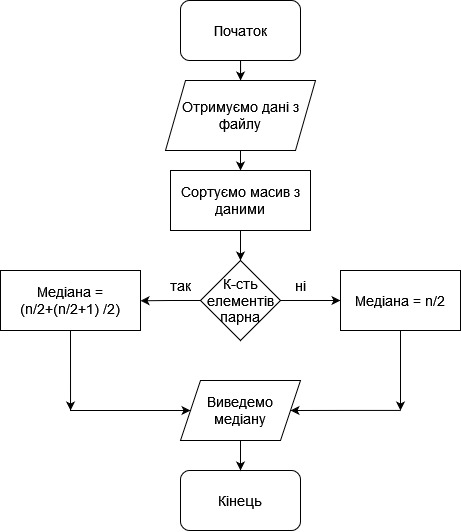
Алгоритм працює без помилок.

**Діаграми алгоритмів:**

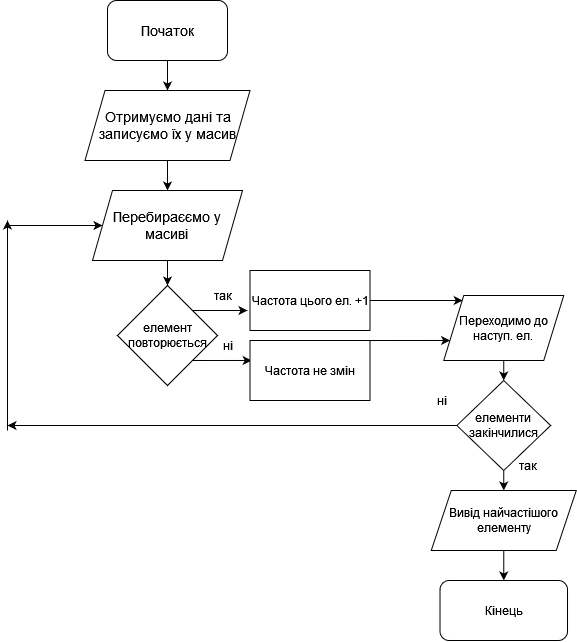
1. Для гістограми та табл. частот:

****

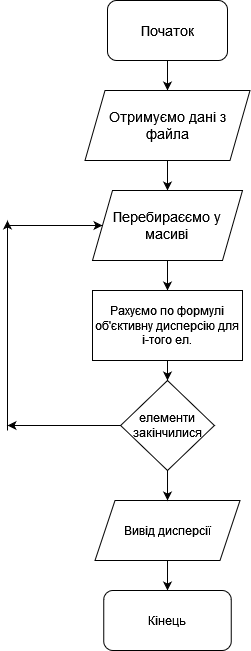
1. Для медіани:

****

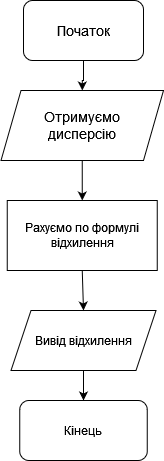
1. Для моди:

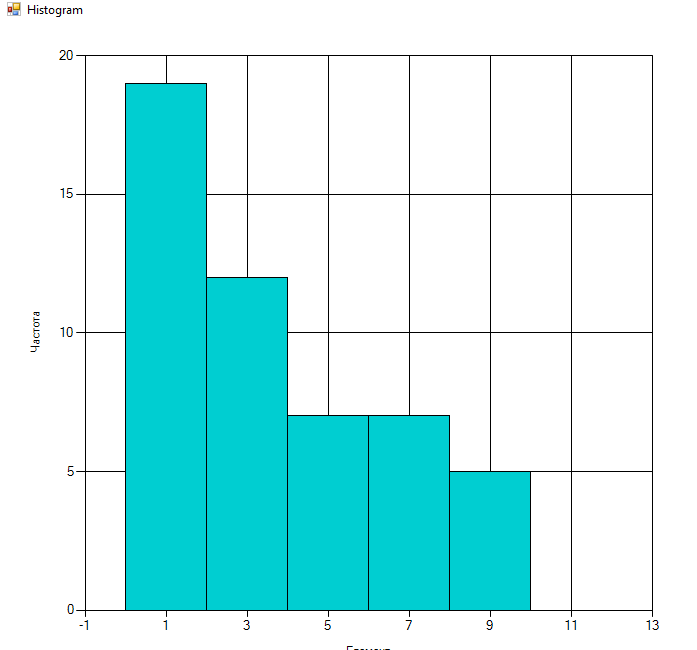
****

1. Для дисперсії:



1. Для відхилення:





**Висновок:** Проаналізувавши гістограму з нашими даними, можемо точно сказати, що графік не є семетричним відносно медіани і не є рівномірним. Оскільки більшість даних з правої сторони, можемо стверджувати, що наш графік з правим перекосом. Медіана більша за моду, але менша за середнє.