

5. Лабораторна робота. Застосування системи комп'ютерної алгебри Maxima до статистичних задач у теорії ймовірностей

Мета: ознайомлення з використанням можливостей роботи СКА Maxima для статистичного аналізу даних, вивчення основних характеристик вибірки, побудова розподілів і використання моделей у задачах теорії ймовірностей.

У Maxima є можливості для обробки статистичних даних, хоча вони не такі розвинені, як у спеціалізованих пакетах типу R або Python (NumPy/Pandas). Проте система дозволяє розв'язувати багато базових і прикладних задач статистики.

Теоретичні відомості та методичні рекомендації

СКА Maxima може наступне:

1. Описова статистика. СКА Maxima може обчислювати:

– середнє арифметичне, наприклад, (див. рис. 5.1):

```
(%i1) mean([1,2,3,4,5]);
```

```
(%o1) 3
```

Рисунок 5.1 – Результат виводу функції mean()

– медіану, наприклад, (див. рис. 5.2):

```
(%i2) median([1,2,3,4,5]);
```

```
(%o2) 3
```

Рисунок 5.2 – Результат виводу функції median()

– дисперсію і стандартне відхилення, наприклад, (див. рис. 5.3):

```
→ variance([1,2,3,4,5]);
```

```
std([1,2,3,4,5]);
```

```
(%o3) variance([1,2,3,4,5])
```

```
(%o4)  $\sqrt{2}$ 
```

Рисунок 5.3 – Результат виводу функцій variance() і std()

– мінімум і максимум, наприклад, (див. рис. 5.4):

```
(%i6) lmin([1,2,3,4,5]);
```

```
lmax([1,2,3,4,5]);
```

```
(%o5) 1
```

```
(%o6) 5
```

Рисунок 5.4 – Результат виводу функцій lmin() і lmax()

2. Робота з випадковими величинами. СКА Maxima підтримує генерацію випадкових чисел і обчислення ймовірностей:

– випадкове число від 0 до 1, наприклад, random(1.0);

- випадкове ціле число від 0 до 100, наприклад, `random(100);`
- множина випадкових чисел (набір вибірки), наприклад, 20 випадкових чисел від 0 до 99 (див. рис. 5.5):

```
(%i7) makelist(random(100), i, 1, 20);
```

```
(%o7) [67,29,85,34,24,27,23,70,88,61,53,7,78,44,3,74,31,28,12,30]
```

Рисунок 5.5 – Результат виводу функції `makelist(random())`

3. Ймовірнісні розподіли. СКА Maxima може працювати з деякими класичними розподілами:

- біноміальний розподіл, наприклад, кількість комбінацій з 5 по 2 (див. рис. 5.6):

```
(%i8) binomial(5,2);
```

```
(%o8) 10
```

Рисунок 5.6 – Результат виводу функції `binomial()`

- ймовірності для дискретних розподілів через формули (Бернуллі, біноміальний, геометричний);
- для неперервних розподілів можна використовувати інтеграли функцій густини.

4. Кореляція та регресія:

- коваріація та кореляція двох вибірок, наприклад, (див. рис. 5.7):

```
(%i10) covariance([1,2,3],[2,4,6]);
```

```
pearson_correlation([1,2,3],[2,4,6]);
```

```
(%o9) covariance([1,2,3],[2,4,6])
```

```
(%o10) pearson_correlation([1,2,3],[2,4,6])
```

Рисунок 5.7 – Результат виводу функцій `covariance()` і `pearson_correlation()`

- проста лінійна регресія через `lsquares_estimates(<дані>, [змінні], <модель>, <незалежна змінна>)` (знаходить коефіцієнти рівняння регресії $y = ax + b$). Ще треба завантажити пакет `lsquares` командою:

```
load("lsquares")
```

5. Візуалізація статистики. Через `draw` можна будувати:

- діаграми розподілу (гістограми);
- точки вибірки для аналізу регресії.

Наприклад, побудова випадкової вибірки (див. рис. 5.8):

```
load("draw") $
data : makelist(random(100), i, 1, 20) $
draw2d(
  points(data)
);
```

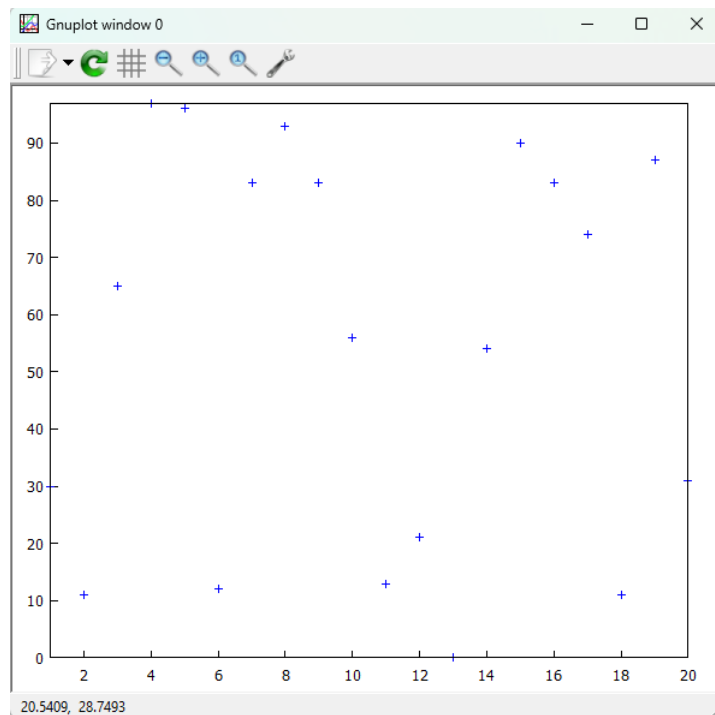


Рисунок 5.8 – Побудова випадкової вибірки

Отже, у Махіта можна розв'язувати такі статистичні задачі:

- 1) обчислення середнього, медіани, дисперсії, стандартного відхилення;
- 2) генерація випадкових вибірок і моделювання експериментів;
- 3) робота з класичними ймовірнісними розподілами;
- 4) коваріація, кореляція, проста лінійна регресія;
- 5) візуалізація даних через точки та гістограми.

Завдання до лабораторної роботи

1. Згенерувати випадкову вибірку випадкових величин, об'ємом 100 елементів (цілі числа). Знайдіть:

- середнє значення;
- медіану;
- моду;
- дисперсію;
- стандартне відхилення.

2. Згенеруйте 1000 випадкових значень з біноміального розподілу з параметрами $n = 10, p = 0,3$. Обчисліть відносну частоту появи кожного значення. Знайдіть ймовірність того, що випадкова змінна менше або дорівнює 3. Використовуйте команди `binomial_pdf(k,n,p)` або `binomial_cdf(k,n,p)` після `load("distrib")`.

3. Задані експериментальні дані:

`data: [[1,2], [2,3], [3,5], [4,4], [5,6]];`

Знайдіть рівняння прямої регресії: $y = ax + b$. Побудуйте графік даних і регресійної прямої.

4. У коробці лежить 5 червоних, 3 синіх і 2 зелених кульки. Витягують 2 кульки без повернення. Знайдіть ймовірність того, що обидві будуть одного

кольору. Моделюйте цей процес 1000 разів і порівняйте експериментальний та теоретичний результат.

Питання для самоконтролю

1. Як обчислити середнє арифметичне засобами СКА Maxima?
2. Як обчислити медіану вибірки засобами СКА Maxima?
3. Як обчислити дисперсію і стандартне відхилення засобами СКА Maxima?
4. Яка функція допоможе знайти мінімум і максимум у СКА Maxima?
5. Яка функція може згенерувати множину випадкових чисел у СКА Maxima?
6. Чи можна у СКА Maxima обчислювати ймовірнісні розподіли? Охарактеризуйте можливості СКА.
7. Як обчислити кореляцію та регресію засобами СКА Maxima?
8. Охарактеризуйте можливості візуалізації даних через точки та гістограми у СКА Maxima.