Лекція 20. Вибіркові числові характеристики

1) *вибіркова середня величина* $\bar{x}_{\text{В}}$. Величину, яка визначається формулою

$$\bar{x}_{\text{В}} = \frac{\sum x_i n_i}{n} \tag{20.1}$$

називають *вибірковою середньою величиною дискретного статистичного розподілу вибірки*.

Тут x_i — варіанта варіаційного ряду вибірки;

n_i — частота цієї варіанти;

n — обсяг вибірки ($n = \sum n_i$).

Якщо всі варіанти з'являються у вибірці лише по одному разу, тобто $n_i = 1$, то

$$\bar{x}_{\text{В}} = \frac{\sum x_i}{n} \tag{20.2}$$

2) *відхилення варіант*. Різницю $(x_i - \bar{x}_{\text{В}})n_i$ називають відхиленням варіант.

При цьому

$$\sum (x_i - \bar{x}_{\text{В}})n_i = \sum x_i n_i - \sum \bar{x}_{\text{В}} n_i = n \cdot \bar{x}_{\text{В}} - n \cdot \bar{x}_{\text{В}} = 0.$$

Отже, сума відхилень усіх варіант варіаційного ряду вибірки завжди дорівнює нулеві;

3) *мода* (Mo^*). *Модою дискретного статистичного розподілу вибірки* називають варіанту, що має найбільшу частоту появи.

Мода може бути кілька. Коли дискретний статистичний розподіл має одну моду, то він називається *одномодальним*, коли має дві моди — *двомодальним* і т. д.;

4) *медіана* (Me^*). *Медіаною дискретного статистичного розподілу вибірки* називають варіанту, яка поділяє варіаційний ряд на дві частини, рівні за кількістю варіант;

5) *дисперсія.* Для вимірювання розсіювання варіант вибірки відносно $\bar{x}_{\text{В}}$ вибирається дисперсія.

Дисперсія вибірки — це середнє арифметичне квадратів відхилень варіант відносно $\bar{x}_{\text{В}}$, яке обчислюється за формулою

$$D_{\text{В}} = \frac{\sum (x_i - \bar{x}_{\text{В}})^2 n_i}{n} \tag{20.3}$$

або

$$D_{\text{В}} = \frac{\sum x_i^2 n_i}{n} - (\bar{x}_{\text{В}})^2 \tag{20.4}$$

6) *середнє квадратичне відхилення вибірки* $\sigma_{\text{В}}$. При обчисленні $D_{\text{В}}$ відхилення підноситься до квадрата, а отже, змінюється одиниця виміру ознаки X, тому на основі дисперсії вводиться середнє квадратичне відхилення

$$\sigma_{\text{В}} = \sqrt{D_{\text{В}}} \tag{20.5}$$

яке вимірює розсіювання варіант вибірки відносно $\bar{x}_{\text{В}}$, але в тих самих одиницях, в яких вимірюється ознака X;

7) *розмах* (R). Для грубого оцінювання розсіювання варіант відносно $\bar{x}_{\text{В}}$ застосовується величина, яка дорівнює різниці між найбільшою $x_{\max}$ і найменшою $x_{\min}$ варіантами варіаційного ряду. Ця величина називається *розмахом*

$$R = x_{\max} - x_{\min} \tag{20.6}$$

8) *коефіцієнт варіації V*. Для порівняння оцінок варіацій статистичних рядів із різними значеннями $\bar{x}_{\text{В}}$, які не дорівнюють нулеві, вводиться коефіцієнт варіації, який обчислюється за формулою

$$V = \frac{\sigma_{\text{В}}}{\bar{x}_{\text{В}}} 100\% \tag{20.7}$$

Приклад 1. За заданим статистичним розподілом вибірки

$X = x_i$	2,5	4,5	6,5	8,5	10,5
n_i	10	20	30	30	10

потрібно:

1) обчислити $\bar{x}_{\text{В}}$, $D_{\text{В}}$, $\sigma_{\text{В}}$;

2) знайти Mo*, Me*;

3) обчислити R, V.

Розв'язання. Оскільки $n = \sum n_i = 100$, то дістанемо:

$$\bar{x}_{\text{В}} = \frac{\sum x_i n_i}{n} = \frac{2,5 \cdot 10 + 4,5 \cdot 20 + 6,5 \cdot 30 + 8,5 \cdot 30 + 10,5 \cdot 10}{100} = 6,7;$$

$$\bar{x}_{\text{В}} = 6,7 .$$

Для обчислення $D_{\text{В}}$ визначається

$$\frac{\sum x_i^2 n_i}{n} = \frac{(2,5)^2 \cdot 10 + (4,5)^2 \cdot 20 + (6,5)^2 \cdot 30 + (8,5)^2 \cdot 30 + (10,5)^2 \cdot 10}{100} = 50,05.$$

Тоді

$$D_{\text{В}} = \frac{\sum x_i^2 n_i}{n} - (\bar{x}_{\text{В}})^2 = 50,05 - (6,7)^2 = 50,05 - 44,89 = 5,16 .$$

$$D_{\text{В}} = 5,16.$$

$$\sigma_{\text{В}} = \sqrt{D_{\text{В}}} = \sqrt{5,16} \approx 2,27.$$

$$\sigma_{\text{В}} = 2,27.$$

$$\text{Mo}^* = 6,5;\ 8,5.$$

Отже, наведений статистичний розподіл вибірки буде двомодальним. Me* = 6,5, оскільки варіанта $x = 6,5$ поділяє варіаційний ряд 2,5; 4,5; **6,5**; 8,5; 10,5 на дві частини: 2,5; 4,5 і 8,5; 10,5, які мають однакову кількість варіант.

$R = x_{\max} - x_{\min} = 10,5 - 2,5 = 8.$

$$V = \frac{\sigma_{\text{В}}}{\bar{x}_{\text{В}}} 100\% = \frac{2,27}{6,7} 100\% = 33,88\%.$$