Лекція №10

Середні значення. Дисперсія.

1. Середня арифметична величина.

2.Середнє гармонійне двох додатних чисел.

3. Середнє квадратичне відхилення. Дисперсія

1. Середня арифметична величина

Середня арифметична є здавна відомою величиною. Вона широко застосовується у науці і техніці. Немає буквально жодної біологічної роботи, де б не зустрічалася в тій чи іншій формі середня арифметична.

Озн. 1: Середня арифметична – це узагальнююча величина, яка носить і абстрактний, і конкретний характер. Абстрактність середньої арифметичної полягає в тому, що вона може не дорівнювати за своїм числовим значенням жодному елементу вибірки. А конкретність виявляється в тому, що вона виражається в тих же одиницях виміру, що і варіанти вибірки. Її значення є середнім між найменшим та найбільшим значеннями вибірки.

Розглянемо вибірку з n вимірів, кожен з яких позначимо $x_1, x_2, x_3, \dots , x_n$. Тоді середнє арифметичне $\overline{X}$ значення для n вимірів:

$$\overline{X} = \frac{x_1 + x_2 + \cdots + x_n}{n}, \tag{1}$$

де n – це кількість вимірів певної ознаки.

Середня величина характеризує групу в цілому і вона більш об’єктивно відбиває аналізований показник, ніж будь-який один вимір (маса однієї тварини). Середнє значення величини X позначають буквою $\overline{X}$ (для будь якої ознаки).

Приклад 1: З трьох дослідних ділянок було зібрано відповідно 324 кг, 360 кг і 270 кг помідорів. Визначити, скільки в середньому було зібрано помідорів з однієї ділянки?

Розв’язання: Складемо статистичну таблицю:

Ділянка	перша	друга	третя
Кількість (x), кг	324	360	270

Середня кількість помідорів, зібраних з ділянки становитиме:

$$\overline{X} = \frac{324 + 360 + 270}{3} = 318(кг).$$

Задачі такого типу є найпростішими задачами на обчислення середнього арифметичного. Для його визначення потрібно суму окремих значень даної ознаки поділити на число одиниць, що мають цю ознаку.

П р и к л а д 2: У таблиці подані відомості про ціну та кількість реалізованого у трьох крамницях товару.

Крамниця	*A*	*B*	*C*
Ціна товару (варіанта x), грн.	3	4	3,6
Кількість товару (чатота n), кг.	350	240	360

Визначити середню ціну реалізованого товару.

<u>Розв’язання:</u> Розрахунок середньої ціни товару здійснюємо таким чином:

$$середня \;\; ціна = \frac{вартість \;\; реалізованого \;\; товару}{кількість \;\; товару}.$$

Отже, середня ціна реалізованого товару буде такою:

$$\overline{X} = \frac{3 \cdot 350 + 4 \cdot 240 + 3{,}6 \cdot 360}{350 + 240 + 360} = \frac{3306}{950}(грн.)$$

Проведені обчислення можна узагальнити формулою:

$$\overline{X} = \frac{x_1 \cdot n_1 + x_2 \cdot n_2 + \cdots + x_n \cdot n_k}{n_1 + n_2 + \ldots + n_k}, \qquad (2)$$

де x_i – варіанта, n_i – частота.

У цій задачі, на відміну від попередньої, варіанти мають різні частоти. Тому формула ускладнюється.

2.Середнє гармонійне двох додатних чисел.

Поряд з поняттям середнього арифметичного двох додатних чисел $\frac{a+b}{2}$ у математичній статистиці існує поняття середнього гармонійного цих чисел $\frac{2}{\frac{1}{a}+\frac{1}{b}}$.

Озн. 2: Середнє гармонійне – це узагальнююча величина, яка обчислюється за формулою: $h=\frac{n}{\frac{1}{a_1}+\frac{1}{a_2}+\ldots+\frac{1}{a_n}}$ (3)

За переказами, поняття середнього гармонійного ввів Піфагор (VI ст. до н.е.), який установив, що разом із струною довжиною 12*l*, співзвучно зливаючись з нею звучать струни такого самого натягу, що мають довжини 6 *l*, 8 *l*, 9 *l*. Неважко помітити, 9 є середнім арифметичним чисел 6 і 12: $\frac{6+12}{2}=9$, 8 є середнім гармонійним цих чисел: $\frac{2}{\frac{1}{6}+\frac{1}{12}}=8$.

<u>Приклад 2:</u> У таблиці подано інформацію про реалізацію трьох сортів картоплі протягом певного часу. Визначити середню ціну реалізованої картоплі.

Сорт	*A*	*B*	*C*
Ціна 1 кг картоплі (x), грн.	3	4	3,6
Вартість реалізованої картоплі ($m_i=x_i\cdot n_i n$), кг.	864	864	864

<u>Розв'язання:</u> Визначимо кількість реалізованого товару: $\left(\frac{864}{3}+\frac{864}{4}+\frac{864}{3{,}6}\right)(кг)$. Тоді середнє значення дорівнює:

$$h=\frac{864+864+864}{\frac{864}{3}+\frac{864}{4}+\frac{864}{3{,}6}}=\frac{1}{\frac{1}{3}+\frac{1}{4}+\frac{1}{3{,}6}}\approx 3{,}48.$$

Отже, середня ціна реалізованої картоплі дорівнює 3,48 грн.

3. Середнє квадратичне відхилення. Дисперсія.

Знаючи середнє арифметичне значення даних експерименту, виникає питання: як обчислити середню величину, на яку відрізняються дані від середнього арифметичного?

Озн. 3: Різницю між будь-яким виміром з вибірки і середнім арифметичним цієї ж вибірки називають відхиленням варіанти x_i від $\overline{X}$: $x_i - \overline{X}$.

Якщо обчислити відхилення для усіх варіант, то серед отриманих значень будуть від'ємні і додатні, які у сумі даватимуть 0, тобто взаємно компенсуються. Це означає, що беззмістовно обчислювати середнє відхилення як середнє арифметичне відхилень. Для того, щоб уникнути компенсації додатних і від'ємних значень, існує декілька способів. Найпоширеніший – піднесення кожної різниці ($x_i - \overline{X}$) до квадрату. (Квадрати як від'ємних, так і додатних величин є величинами додатними.) Додаючи квадрати усіх різниць і ділячи на кількість цих різниць, отримаємо величину, яка називається дисперсією. Фактично вона показує середнє арифметичне квадратів відхилень. Для того, щоб позбутися квадрату величини, обчислюємо корінь квадратний з дисперсії. Отримане значення називається середнім квадратичним відхиленням. Розрізняють формули середнього квадратичного відхилення для генеральної і вибіркової сукупностей.

При існуючих даних генеральної сукупності використовують таку формулу: $\sigma = \sqrt{\dfrac{\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2}{n}}$, (4)

де X_i − значення *i*-тої варіанти, $i=1,...,n$, $\overline{X}$ – середнє арифметичне, n – об'єм генеральної сукупності.

Оскільки $D(X) = \sigma^2$, то $D(X) = \dfrac{\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2}{n}$ − зміщена дисперсія. (5)

Якщо ж є тільки дані вибірки, то для визначення середнього квадратичного відхилення застосовується така формула:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2}{n-1}}, \tag{6}$$

А дисперсія, що називають незміщеною, відповідно:

$$D\left(X\right) = \frac{\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2}{n-1}, \tag{7}$$

де X_i – значення i-тої варіанти, $i=1,...,n$, $\overline{X}$ – середнє арифметичне, n – об'єм вибіркової сукупності.

<u>Приклад 3.</u> Було зроблено 5 вимірів вмісту кальцію в крові (в умовних одиницях): 11,27; 11,36; 11,09; 11,16; 11,47. Обчислити середнє квадратичне відхилення та дисперсію вмісту кальцію у крові (5 даних вимірів розглядаємо як вибіркову сукупність):

<u>Розв'язання:</u> $\overline{X} = \frac{11{,}27 + 11{,}36 + 11{,}09 + 11{,}16 + 11{,}47}{5} = \frac{56{,}35}{5} = 11{,}27$.

Отже, середнє арифметичне значення вмісту кальцію у крові дорівнює 11,27.

Тоді середнє квадратичне відхилення:

$$\sigma = \sqrt{\frac{\left(11{,}27 - 11{,}27\right)^2 + \left(11{,}36 - 11{,}27\right)^2 + \left(11{,}09 - 11{,}27\right)^2 + \left(11{,}16 - 11{,}27\right)^2 + \left(11{,}47 - 11{,}27\right)^2}{5-1}} \approx$$

$\approx 0{,}152$.

А дисперсія $D\left(X\right) = \sigma^2 = 0{,}152^2 \approx 0{,}023$.

Д.з.: «Т. й.» Шевченко Р.Л., Ревицька У.С. Розділ VII. § 4.