

СРЕДНИЕ ВЕЛИЧИНЫ И ПОКАЗАТЕЛИ ВАРИАЦИИ

1. Средние величины. Общие принципы их применения.
 2. Расчет средних величин по результатам группировки.
 3. Структурные средние.
 4. Показатели вариации
-

СРЕДНЯЯ ВЕЛИЧИНА – ОБОБЩАЮЩИЙ ПОКАЗАТЕЛЬ, ХАРАКТЕРИЗУЮЩИЙ ТИПИЧЕСКИЙ УРОВЕНЬ ЯВЛЕНИЯ.

- Средняя величина выражает величину признака, отнесенную к единице совокупности.
- Средняя всегда обобщает количественную вариацию признака
- В средних величинах погашаются индивидуальные различия единиц совокупности, обусловленные случайными обстоятельствами.

ОБЩИЕ ПРИНЦИПЫ ПРИМЕНЕНИЯ СРЕДНИХ ВЕЛИЧИН:

- При определении средней величины нужно исходить из качественного содержания осредняемого признака, учитывать взаимосвязь изучаемых признаков и имеющиеся для расчета данные.
- Средняя величина должна рассчитываться по однородной совокупности.
- Общие средние должны подкрепляться групповыми средними.
- Необходим обоснованный выбор единицы совокупности, для которой рассчитывается средняя.

ВИДЫ СРЕДНИХ ВЕЛИЧИН

А. Степенные средние

• *Простая средняя :*

$$\bar{X} = \sqrt[m]{\frac{\sum X_i^m}{n}}$$

где X_i – варианты (значения) осредняемого признака;

m – показатель степени средней;

n – число вариантов.

$$\bar{X} = \sqrt[m]{\frac{\sum X_i^m f_i}{\sum f_i}}$$

• *Взвешенная средняя*

где X_i – варианты (значения) осредняемого признака или срединное значение интервала, в котором измеряется варианта;

m – показатель степени средней;

f_i – частота, показывающая, сколько раз встречается i -е значение осредняемого признака

В. Структурные средние (мода, медиана и др.)

В ЗАВИСИМОСТИ ОТ СТЕПЕНИ РАЗЛИЧАЮТ:

Вид степенной средней	Показатель степени (m)	Формула расчета	
		Простая	Взвешенная
Гармоническая	- 1	$\bar{X} = \frac{n}{\sum \frac{1}{x}}$	$\bar{X} = \frac{\sum W}{\sum \frac{W}{x}}$
Геометрическая	0	$\bar{X} = \sqrt[n]{\prod x} =$ $= \sqrt[n]{x_1 x_2 \dots x_n}$	$\bar{X} = \sqrt[\sum f]{\prod x^f} =$ $= \sqrt[\sum f]{x_1^{f_1} x_2^{f_2} \dots x_n^{f_n}}$
Арифметическая	1	$\bar{X} = \frac{\sum x}{n}$	$\bar{X} = \frac{\sum xf}{\sum f}$
Квадратическая	2	$\bar{X} = \sqrt{\frac{\sum x^2}{n}}$	$\bar{X} = \sqrt{\frac{\sum x^2 f}{\sum f}}$

ПРИМЕР 1, ПО ДАННЫМ ОБСЛЕДОВАНИЯ РАСХОДЫ НА УПЛАТУ АДМИНИСТРАТИВНЫХ ШТРАФОВ ОДНОЙ ИЗ ГРУПП НАСЕЛЕНИЯ СОСТАВИЛИ (РУБ. В МЕСЯЦ): 2020, 2250, 2310, 2320, 3020, 3280, 3650, 3980, 4210, 4800, 4920, 5430, 5670, 6120, 7320.

Данные не сгруппированы, поэтому среднюю рассчитываем по формуле средней арифметической простой:

$$\begin{aligned}\bar{X} &= \frac{\sum x}{n} = (2020 + 2250 + 2310 + \dots \\ &+ \dots + 7320) : 15 = \\ &61300 : 15 = 4088,7(\text{руб})\end{aligned}$$

ПРАВИЛА ОПРЕДЕЛЕНИЯ ВИДА (ГАРМОНИЧЕСКАЯ ИЛИ АРИФМЕТИЧЕСКАЯ) СРЕДНЕЙ:

- записать исходное отношение для определения средней
- если в исходном отношении не известен числитель – использовать среднюю арифметическую
- если в исходном отношении не известен знаменатель – использовать среднюю гармоническую

ИСХОДНЫЕ ОТНОШЕНИЯ:

- Среднедушевые доходы на душу населения = Сумма доходов по всем источникам поступления / Численность группы
- Средний срок расследования = Время, затраченное на расследование всех дел (по отделению, у того или иного следователя и т.п.) / Число дел
- Средняя нагрузка на одного следователя = Число дел расследованных за изучаемый период / Число следователей и т.д.
- Среднее число расследованных дел = Общее число расследованных дел / Число дел, находящихся в производстве

ПРИМЕР2, ИМЕЮТСЯ СЛЕДУЮЩИЕ ДАННЫЕ О ВЫПЛАТАХ АДМИНИСТРАТИВНЫХ ШТРАФОВ ЗА

ДВА ПЕРИОДА :

Основание выплаты штрафа	Базисный период		Отчетный период	
	Средний размер штрафа руб.	Число плательщиков тыс.чел.	Средний размер штрафа руб.	Общая сумма выплат, млн.руб.
1	5120	1,8	5140	13782,5
2	5150	2,1	5180	11856,3
3	5135	1,8	5170	9074,9

ТОГДА СРЕДНИЕ ВЕЛИЧИНЫ ОПРЕДЕЛЯЕМ:

- В базисном периоде по формуле средней арифметической:

$$\bar{X} = \frac{\sum X_i f_i}{\sum f_i} = \frac{5120 * 1,8 + 5150 * 2,1 + 5135 * 1,8}{1,8 + 2,1 + 1,8} = \frac{29274}{5,7} = 5135,79(\text{руб.})$$

- В отчетном периоде по формуле средней гармонической:

$$\bar{X} = \frac{\sum W_i}{\sum \frac{W_i}{X_i}} = \frac{13782,5 + 11856,3 + 9074,9}{\frac{13782,5}{5140} + \frac{11856,3}{5180} + \frac{9074,9}{5170}} = \frac{34713,7}{6,74} = 5161,57(\text{руб.})$$

ЕСЛИ СРЕДНЯЯ ОПРЕДЕЛЯЕТСЯ НЕ ПО ДИСКРЕТНОМУ, А ПО ИНТЕРВАЛЬНОМУ РЯДУ РАСПРЕДЕЛЕНИЯ:

- И интервалы открыты, то сначала закрывают интервалы, полагая, что величина открытого интервала равна величине предыдущего или последующего интервала.
- В качестве варианты берут середину интервала, определяя её как полусумму нижней и верхней границы интервала:

Например,

$$\bar{X}_i = \frac{X_{max} + X_{min}}{2}$$

РАСПРЕДЕЛЕНИИ НАСЕЛЕНИЯ ОДНОГО ИЗ РЕГИОНОВ ПО РАЗМЕРУ СРЕДНЕМЕСЯЧНОГО ДУШЕВОГО РАСХОДА НАСЕЛЕНИЯ НА ПРОДУКТЫ ПИТАНИЯ ПРЕДСТАВЛЕНО В ГРАФЕ 1 И 2. РАССЧИТАЕМ СРЕДНИЕ ВЕЛИЧИНЫ:

	f	X_i	$X_i \cdot f_i$	S_{me}	плотность распределения
до 600	4,3	300	1290	4,3	0,007167
600-1200	64,9	900	58410	69,2	0,108167
1200-1800	163,3	1200	195960	232,5	0,272167
1800-2400	228,6	2100	480060	461,1	0,381
2400-3000	252,2	2700	680940	713,3	0,420333
3000-4000	403,4	3500	1411900	1116,7	0,4034
4000-5000	340,7	4500	1533150	1457,4	0,3407
5000-6000	270,5	5500	1487750	1727,9	0,2705
6000-7000	209,1	6500	1359150	1937	0,2091
7000-8000	160,1	7500	1200750	2097,1	0,1601
8000-9000	122,4	8500	1040400	2219,5	0,1224
9000-10000	93,7	9500	890150	2313,2	0,0937
10000 и более	349,3	10500	3667650	2662,5	0,3493
итого	2662,5		14007560		

СРЕДНЯЯ АРИФМЕТИЧЕСКАЯ:

$$\bar{X} = \frac{\sum X_i f_i}{\sum f_i} =$$

$$14007560 : 2662,5 = 5261 \text{ (руб.)}$$

МОДА- НАИБОЛЕЕ ЧАСТО ВСТРЕЧАЮЩЕЕСЯ ЗНАЧЕНИЕ ПРИЗНАКА

$$M_o = X_{M_o} + i \frac{f_{M_o} - f_{M_o-1}}{(f_{M_o} - f_{M_o-1}) + (f_{M_o} - f_{M_o+1})} = 2400 + 600 \frac{0,42 - 0,39}{(0,42 - 0,39) + (0,42 - 0,4)} = 2760(\text{руб})$$

где X_{M_o} - нижняя граница модального интервала
(2400);

i - величина модального интервала
(600=3000-2400)

f_{M_o} , f_{M_o-1} и f_{M_o+1} - плотность модального, до
модального и после модального интервалов
соответственно (0,42; 0,387 и 0,40)

МЕДИАНА- ЗНАЧЕНИЕ ПРИЗНАКА, НАХОДЯЩЕГОСЯ В СЕРЕДИНЕ РЯДА РАСПРЕДЕЛЕНИЯ

$$Me = X_{Me} + i_{Me} \cdot \frac{\frac{\sum f}{2} - S_{Me-1}}{f_{Me}} = 4000 + 1000 \frac{2662,5 / 2 - 1116,7}{340,7} = 4629,7(\text{руб})$$

X_{Me} - нижняя граница медианного интервала
(4000);

i - величина медианного интервала
(1000=4000-3000)

f_{me} - частота медианного интервала (340,7);

S_{me-1} - сумма накопленных частот,
предшествующих медианному интервалу (1116.7)

ПОКАЗАТЕЛИ ВАРИАЦИИ:

- Размах вариации: $H = X_{\max} - X_{\min}$.

- Среднее линейное отклонение:

$$L = \frac{\sum |X_i - \bar{X}| f_i}{\sum f_i}$$

$$\sigma^2 = \frac{\sum (X_i - \bar{X})^2 f_i}{\sum f_i}$$

- Дисперсия

или
$$\sigma^2 = \bar{X}^2 - (\bar{X})^2 = \frac{\sum X_i^2 f_i}{\sum f_i} - \left(\frac{\sum X_i f_i}{\sum f_i} \right)^2$$

- Среднее квадратическое отклонение:

$$\sigma = \sqrt{\sigma^2}$$

- Коэффициент вариации

$$V = \frac{\sigma}{\bar{X}} 100$$

ПО РАССМОТРЕННОМУ ПРИМЕРУ ДАДИМ ОЦЕНКУ ВАРИАЦИИ РАСХОДОВ РЕГИОНА:

X_i	f	$X - \bar{X}$	$(X - \bar{X})^2$	$ x - \bar{x} \cdot f_i$	$(X - \bar{X})^2 \cdot f_i$
	1	2	3	4	5
300	4,3	-4961	24611521	21332,3	1,06E+08
900	64,9	-4361	19018321	283028,9	1,23E+09
1200	163,3	-4061	16491721	663161,3	2,69E+09
2100	228,6	-3161	9991921	722604,6	2,28E+09
2700	252,2	-2561	6558721	645884,2	1,65E+09
3500	403,4	-1761	3101121	710387,4	1,25E+09
4500	340,7	-761	579121	259272,7	1,97E+08
5500	270,5	239	57121	64649,5	15451231
6500	209,1	1239	1535121	259074,9	3,21E+08
7500	160,1	2239	5013121	358463,9	8,03E+08
8500	122,4	3239	10491121	396453,6	1,28E+09
9500	93,7	4239	17969121	397194,3	1,68E+09
10500	349,3	5239	27447121	1829982,7	9,59E+09
	2662,5		1,43E+08	6611490,3	2,31E+10

ПОКАЗАТЕЛИ ВАРИАЦИИ:

- Среднее линейное отклонение:

$$L = 6611490,3 / 2662,5 = 2483,29 \text{ (руб.)}$$

- Дисперсия: $2311392 : 2662,5 = 8681296$

- Среднее квадратическое отклонение:

$$\sigma = \sqrt{\sigma^2} = 2946,4 \text{ (руб.)}$$

- Коэффициент вариации:

$$V = \frac{\sigma}{X} 100 = 2946,4 / 5261 * 100 = 56(\%)$$

СВОЙСТВА ДИСПЕРСИИ:

- Дисперсия постоянной величины равна 0.
- Уменьшение всех значений признака на одну и ту же величину (A) не изменит величину дисперсии:
- Уменьшение всех значений признака в K раз уменьшает дисперсию в K^2 раз:

$$\sigma^2(X_i - A) = \sigma^2(X_i)$$

$$\sigma^2\left(\frac{X_i}{K}\right) = \frac{\sigma^2(X_i)}{K^2}$$

РАССМОТРЕННЫЕ СВОЙСТВА ПОЗВОЛЯЮТ УПРОСТИТЬ РАСЧЕТ ДИСПЕРСИИ, ИСПОЛЬЗОВАТЬ СПОСОБ МОМЕНТОВ:

$$\sigma^2 = K^2 \cdot (m_2 - m_1^2)$$

- Дисперсия:

где $m_1 = \frac{\sum X_i \cdot f_i}{\sum f_i}$ - момент 1-го порядка,

$m_2 = \frac{\sum X_i^2 \cdot f_i}{\sum f_i}$ - момент 2-го порядка,

A – постоянная величина (константа)

K – величина интервала

Например,
 $\frac{X_i - A}{K}$

ДАНО РАСПРЕДЕЛЕНИЕ РАБОТАЮЩИХ ОДНОГО ИЗ ПРЕДПРИЯТИЙ ПО РАЗМЕРУ СРЕДНЕЙ ЗАРАБОТНОЙ ПЛАТЫ:

	Число сотрудников	X_i	$X'_i = \frac{X_i - 8000}{2000}$	$X'_i \cdot f_i$	$X_i'^2 \cdot f_i$
До 3000	5	2000	-3	-15	45
3000-5000	7	4000	-2	-14	28
5000-7000	12	6000	-1	-12	12
7000-9000	17	8000	0	0	0
9000-11000	14	10000	1	14	14
11000-13000	11	12000	2	22	44
свыше 13000	9	14000	3	27	81
Итого				22	224

НАЙДЕМ МОМЕНТЫ ПЕРВОГО И ВТОРОГО ПОРЯДКА:

$$m_1 = \frac{\sum X_i' \cdot f_i}{\sum f_i} = 22 / 75 = 0,293 \quad m_2 = \frac{\sum X_i'^2 \cdot f_i}{\sum f_i} = 224 / 75 = 2,987$$

- **средняя заработная плата по предприятию составит:**

- **Дисперсия:** $X = K * X' + A = 2000 * 0,086 + 8000 = 8586,7(\text{руб.})$

- **Среднее квадратическое отклонение:**
 $\sigma^2 = K^2 \cdot (m_2 - m_1) = 4000000 * (2,987 - 0,086) = 11602489$

- **Коэффициент вариации:**

$$\sigma = \sqrt{\sigma^2} = 3406,2(\text{руб.})$$

$$V = \frac{\sigma}{X} 100 = 3406,2 / 8586,7 * 100 = 39,7(\%)$$

НАРЯДУ С ОБЩЕЙ ДИСПЕРСИЕЙ, ИЗМЕРЯЮЩЕЙ ВАРИАЦИЮ

ПРИЗНАКА ПО ВСЕЙ

СОВОКУПНОСТИ РАССЧИТЫВАЮТ:

- Внутригрупповые дисперсии: $\sigma_i^2 = \frac{\sum (X_i - \bar{X}_i)^2 f_i}{\sum f_i}$
- Среднюю из внутригрупповых: $\bar{\sigma}_i^2 = \frac{\sum \sigma_i^2 f_i}{\sum f_i}$
- Межгрупповую дисперсию: $\sigma_{мг}^2 = \frac{\sum (\bar{X}_i - \bar{X})^2 f_i}{\sum f_i}$
- И согласно правилу сложения дисперсий: общая дисперсия равна сумме межгрупповой и средней из групповых дисперсий, т.е.

$$\sigma^2 = \bar{\sigma}_i^2 + \sigma_{мг}^2$$

КОЭФФИЦИЕНТ ДЕТЕРМИНАЦИИ - ДОЛЯ МЕЖГРУППОВОЙ ДИСПЕРСИИ В ОБЩЕЙ ДИСПЕРСИИ ПРИЗНАКА РЕЗУЛЬТАТА

$$\eta^2 = \frac{\sigma_{мг}^2}{\sigma^2}$$

- он показывает влияние фактора X положенного в основание группировки на часть общей вариации признака результата Y.

Эмпирическое корреляционное отношение- корень квадратный из коэффициента детерминации оценивает тесноту связи:

$$\eta = \sqrt{\frac{\sigma_{мг}^2}{\sigma^2}}$$

Соотношения Чаддока: до 0,3 –слабая; 0,3-0,5-умеренная; 0,5-0,7 – заметная; 0,7-0,9 – тесная ; св.0,9 - весьма тесная