

Средние величины



Сущность средних величин

Средние в статистике - это показатели, выражающие типичные размеры признака для данной совокупности. В них взаимопогашаются индивидуальные отклонения, присущие отдельным единицам и показываются значения признака, характерные для всей совокупности.



Виды средних

□ Средняя арифметическая простая $\bar{x} = \frac{\sum x}{n}$

□ Средняя арифметическая взвешенная $\bar{x} = \frac{\sum x \cdot f}{\sum f}$

□ Средняя гармоническая взвешенная $\bar{x} = \frac{\sum M}{\sum \frac{M}{x}}$



Для правильного определения средней величины признака, необходимо, прежде всего, понять сущность этого показателя, уяснить соотношение каких величин он выражает, что является исходной базой для расчета среднего значения данного признака



Допустим, необходимо определить средний размер заработной платы. Исходной базой для расчета средней заработной платы будет служить соотношение:

$$\frac{\text{фонд заработной платы}}{\text{число рабочих}}$$

Если фонд заработной платы обозначить M , число рабочих через f , а среднюю заработную плату через \bar{x} , то

$$\bar{x} = \frac{M}{f}$$

Выбор формулы для расчета средней величины зависит от имеющейся исходной информации



Имеются следующие данные о заработной плате рабочих бригады

Табельный номер рабочего	1	2	3	4	5	6	7	8	9	10
Зарботная плата, (д.е.)	300	250	400	350	275	325	425	375	225	425

Обозначив заработную плату через x , фонд заработной платы (M) можно определить как сумму заработной платы всех 10 рабочих, тогда

$$\bar{x} = \frac{\sum x}{n}$$

n – число значений x , равное числу рабочих бригады.

Следовательно, если каждое значение признака встречается один или одинаковое число раз, для расчета средней величины признака используется формула средней арифметической простой



Имеются следующие данные:

№ бригады	Число рабочих	Зарботная плата одного рабочего (д.ед.)
1	10	250
2	12	300

Необходимо определить среднюю заработную плату рабочего по 2 бригадам

Исходя из имеющихся данных фонд заработной платы по каждой бригаде определяется путем умножения заработной платы одного рабочего (x) на число рабочих (f), а по двум бригадам вместе как сумму $x \cdot f$, тогда

$$\bar{x} = \frac{\sum x \cdot f}{\sum f} = \frac{250 \cdot 10 + 300 \cdot 12}{10 + 12} = \frac{6100}{22} = 277,3 \text{ д.ед.}$$

Если каждое значение признака встречается несколько раз, для расчета средней величины используется формула средней арифметической взвешенной



Имеются следующие данные

№ бригады	Зарботная плата одного рабочего, д.е.	Фонд зарботной платы, д.е.
1	300	3000
2	350	2800

Исходной формулой для расчета средней зарботной платы является $\bar{x} = \frac{M}{f}$,
выражаем f как $\frac{M}{f}$

Тогда
$$\bar{x} = \frac{\sum M}{\sum \frac{M}{x}} = \frac{3000 + 2800}{\frac{3000}{300} + \frac{2800}{350}} = \frac{5800}{18} = 322,2 \text{ д.е.}$$

В данном случае используется формула средней гармонической взвешенной



Если варьирующий признак представлен в виде интервала «от-до», в качестве конкретных вариантов признака принимаются середины интервалов. Ширина открытого интервала принимается равной ширине примыкающего.

Среднее значение признака определяется по формуле средней арифметической взвешенной $\bar{x} = \frac{\sum x \cdot f}{\sum f}$.

Расчеты обычно располагают в виде таблицы.



Имеются данные о распределении рабочих по размеру заработной платы

Зарботная плата, (д.е.).	Число рабочих (f)	Середина интервала (x)	$x \cdot f$
До 250	10	237,5	2375,0
250 – 275	15	262,5	3937,5
275 – 300	18	287,5	5175,0
300 – 325	12	312,5	3750,0
325 и более	5	337,5	1687,5
Итого	60		16925



Серединное значение первого интервала равно 237,5

$\left(\frac{225 + 250}{2}\right)$ второго - **262,5** $\left(\frac{220 + 275}{2}\right)$ и т. д.

$$\bar{x} = \frac{\sum x \cdot f}{\sum f} = \frac{16925}{60} = 282,08 \text{ д.ед}$$



Мода и медиана

□ Мода в статистике - это значение признака, наиболее часто встречающееся в изучаемой совокупности.

□ Медиана - это значение признака у единицы совокупности, делящей ранжированный ряд пополам (или стоящей в середине ранжированного ряда).

В дискретном ряду распределения модой является вариант признака, имеющий наибольшую частоту.



Распределение рабочих по тарифному разряду

Разряд	1	2	3	4	5	6	Итого
Число рабочих	5	6	18	16	11	9	67

Наибольшее число рабочих (18) имеют третий разряд. Следовательно, мода для данной совокупности – 3 разряд.

Для нахождения медианы строится ряд накопленных частот.



Разряд	Число рабочих	Накопленная частота
1	5	5
2	8	5+8=13
3	18	13+18=31
4	16	31+16=47
5	11	47+11=58
6	9	58+9=67
Итого	67	



В данной совокупности состоящей из 67 единиц, в середине ранжированного ряда будет находиться 34-й рабочий . Рабочих с 1, 2, 3 разрядом насчитывается 31. Эта величина меньше порядкового номера медианы. Накопленная частота для 4 разряда - 47, т. е. превышает порядковый номер медианы. Отсюда следует, что рабочий, имеющий порядковый номер 34 принадлежит к 4-й тарифной группе. Следовательно, медиана в примере - четвертый разряд.



Графическое изображение рядов распределения

Наглядно ряды распределения представляются при помощи графических изображений.

Ряды распределения изображаются в виде:

- Полигона*
- Гистограммы*
- Кумуляты*
- Огивы*



Полигон

При построении полигона на горизонтальной оси (ось абсцисс) откладывают значения варьирующего признака, а на вертикальной оси (ось ординат) — частоты или частости

Домохозяйства, состоящие из:	одного человека	дзух человек	трех человек	5 или более	всего
Число домохозяйств в %	19,2	25,2	22,6	20,5	100,0



Полигон построен по данным микропереписи населения России



- **Условие:** Приводятся данные о распределении 25 работников одного из предприятий по тарифным разрядам: 4; 2; 4; 6; 5; 6; 4; 1; 3; 1; 2; 5; 2; 6; 3; 1; 2; 3; 4; 5; 4; 6; 2; 3; 4
- **Задание:** Построить дискретный вариационный ряд и изобразить его графически в виде полигона распределения.
- **Решение:** В данном примере вариантами является тарифный разряд работника. Для определения частот необходимо рассчитать число работников, имеющих соответствующий тарифный разряд.
- Для построения полигона распределения по оси абсцисс (X) откладываем количественные значения варьирующего признака — варианты, а по оси ординат — частоты или частости.



Тарифный разряд X_j	Число работников f_j
1	3
2	5
3	4
4	6
5	3
6	4
Итого:	25



Гистограмма

- Если значения признака выражены в виде интервалов, то такой ряд называется интервальным.
- Интервальные ряды распределения изображают графически в виде гистограммы.
- Гистограмма – график, на котором ряд изображен в виде смежных друг с другом столбиков. Для построения гистограммы по оси абсцисс указывают значения границ интервалов и на их основании строят прямоугольники, высота которых пропорциональна частотам (или частостям).



Все население	В том числе в возрасте								
		до 10	10- 20	20- 30	30- 40	40- 50	50- 60	60- 70	70 и старше
Численность населения	12,1	15,7	13,6	16,1	15,3	10,1	9,8	7,3	100,0

Необходимо распределить население страны по возрастным группам



Гистограмма

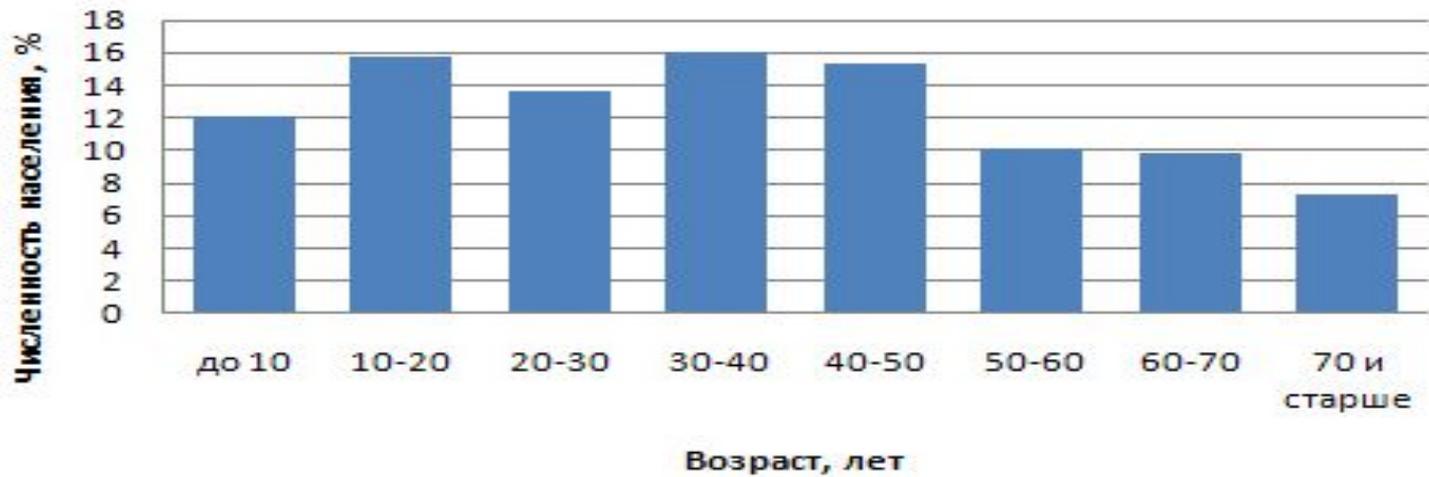


Рис. 6.2. Распределение населения России по возрастным группам



Имеется распределение 30 работников фирмы по размеру месячной заработной платы

Размер заработной платы руб. в месяц	Численность работников чел.
до 5000	4
5000 — 7000	12
7000 — 10000	8
10000 — 15000	6
Итого:	30

Необходимо изобразить интервальный вариационный ряд графически в виде кумуляты



Решение:

- 1. Неизвестная граница открытого (первого) интервала определяется по величине второго интервала: $7000 - 5000 = 2000$ руб. С той же величиной находим нижнюю границу первого интервала: $5000 - 2000 = 3000$ руб.**
- 2. Для построения гистограммы в прямоугольной системе координат по оси абсцисс откладываем отрезки, величины которых соответствуют интервалам вариационного ряда.**

Эти отрезки служат нижним основанием, а соответствующая частота (частость) — высотой образуемых прямоугольников.

- 3. Построим гистограмму**



При построении кумуляты накопленная частота (частость) соответствующего интервала присваивается его верхней границе



Огива

- Огива строится аналогично кумуляте с той лишь разницей, что накопленные частоты помещают на оси абсцисс, а значения признака — на оси ординат.
- Разновидностью кумуляты является кривая концентрации или график Лоренца. Для построения кривой концентрации на обе оси прямоугольной системы координат наносится масштабная шкала в процентах от 0 до 100. При этом на оси абсцисс указывают накопленные частоты, а на оси ординат — накопленные значения доли (в процентах) по объему признака.
- Равномерному распределению признака соответствует на графике диагональ квадрата (рис. 6.4). При неравномерном распределении график представляет собой вогнутую кривую в зависимости от уровня концентрации признака.



Огива



6.4. Кривая концентрации

