

# Мера центральной тенденции

Средние величины и изучение  
вариации

# Мера центральных тенденций.

## Сущность и значение средних показателей

---

- Средняя величина – обобщенная количественная характеристика признака в статистической совокупности в конкретных условиях места и времени
- Сущность средней – в ней взаимопоглощаются отклонения значения признака отдельных единиц совокупности



- Средняя отражает типичный уровень признака

# Сущность и значение средних показателей

---

- Логическая формула средней:

$$\text{Средняя} = \frac{\text{Суммарное значение}}{\text{Число единиц}}$$

# Средняя арифметическая (простая)

- Используется для несгруппированных данных

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}, \quad (1)$$

где  $x_1, \dots, x_n$  — значения признака,  $n$  — число наблюдений.

Номер респондента	1	2	3	4	5	6
Стаж работы	10	3	5	12	11	7
Средний стаж	???					

- средняя арифметическая величина может быть дробным числом, если даже индивидуальные значения признака принимают только целые значения

# Средняя арифметическая (взвешенная)

- Используется для сгруппированных данных (дискретных или интервальных)

Сделка	Количество проданных акций	Курс продажи (руб.)
1	500	1080
2	300	1050
3	1100	1145

- Средняя 1112,9 руб.

# Средняя арифметическая (взвешенная)

Возраст	Число работников
До 25	7
26-30	13
31-40	38
41-50	42
51-60	16
Старше 60	5

- Ищем середину интервала
- Средний возраст 40,7 лет

# Задача:

## Результаты выполнения вопроса А1 ЕГЭ

А1	% верно ответивших
1 вариант	74,8
2 вариант	88,1
3 вариант	82,2
4 вариант	72,1
5 вариант	89,8
6 вариант	68,1
7 вариант	72,9
8 вариант	95,6

- **Вычисление средней в этом случае возможно только если количество отвечающих по каждому варианту совпадает**

# Задача

## Результаты выполнения вопроса А1 ЕГЭ-15

<b>Вопрос 1</b>	Количество человек	% верно ответивших
1 вариант	636	74,8
2 вариант	428	88,1
3 вариант	646	82,2
4 вариант	1282	72,1
5 вариант	859	89,8
6 вариант	865	68,1
7 вариант	48	72,9
8 вариант	45	95,6

# Медиана

---

- Медианой называется значение признака у той единицы совокупности, которая расположена в середине ряда частотного распределения.
- Медиана делит упорядоченный вариационный ряд на две равные по численности группы
- Медиана – это такое значение признака, которое делит всю совокупность пополам

# Медиана

## 1. Объем совокупности нечетный

Если объем совокупности нечетный и равен  $2n + 1$ , и варианты дискретного ряда размещены в порядке возрастания/убывания их значений, тогда медиана

$$Me = \frac{X_{n+1}}{2}$$

где  $X$  – значение признака,

$n$  – количество единиц совокупности

Респондент	1	2	3	4	5	6	7
Стаж работы	13	12	10	8	6	5	3

# Медиана

---

## 2. Объем совокупности четный

Если в ряду четное число членов ( $2k$ ), то медиана равна среднему арифметическому из двух срединных значений признака.

$$Me = \frac{x_n + x_{n+1}}{2}.$$

# Медиана

Вычисление медианы для интервального ряда

$$Me = x_0 + \delta \frac{\frac{1}{2}n - n_H}{n_{Me}},$$

# Медиана

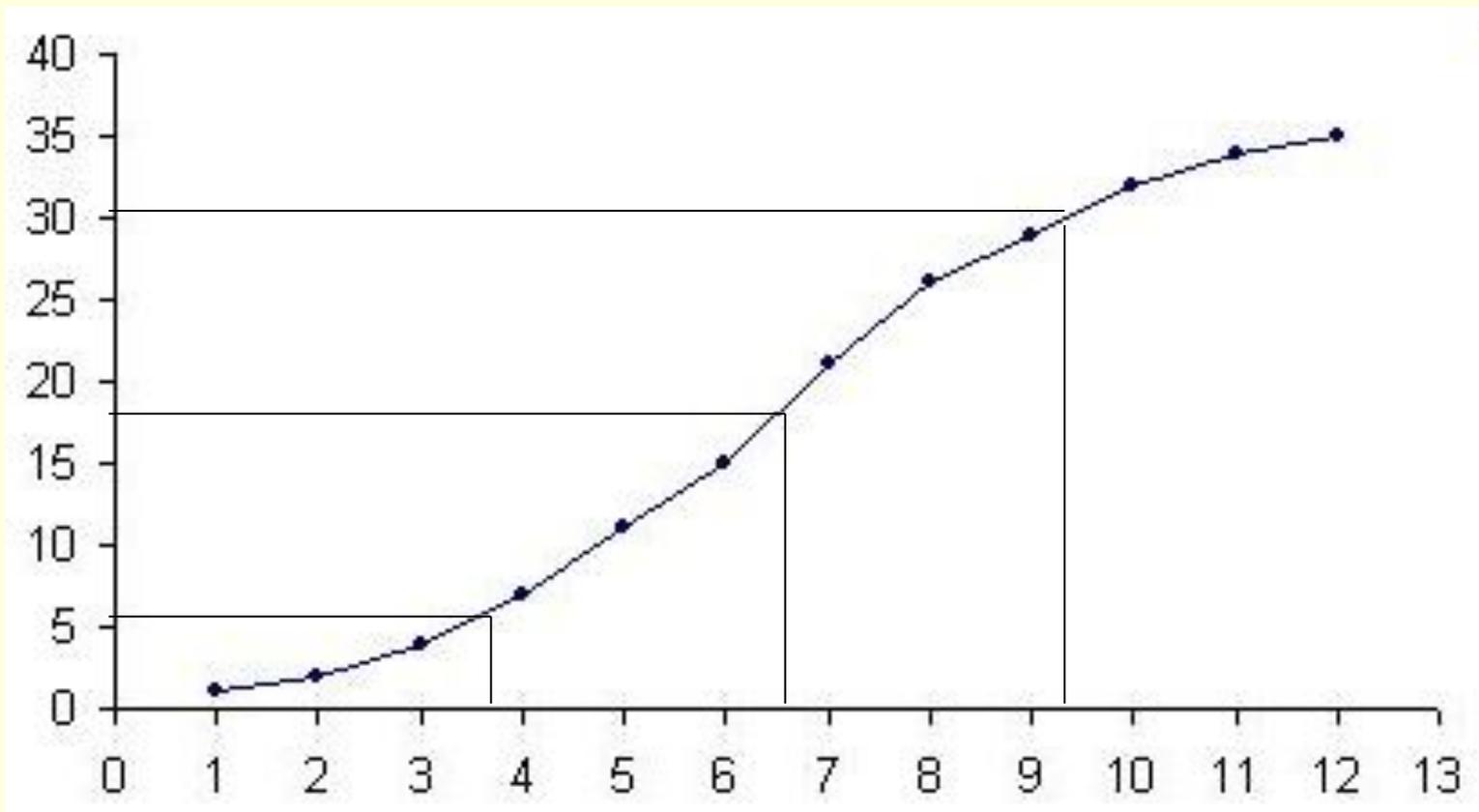
- Медиану можно определить и для порядковых качественных данных.
- Пример:
- Совокупность пяти военнослужащих, имеющих воинские звания: рядовой, ефрейтор, младший сержант, сержант, старший сержант. Эти данные упорядочены по возрастанию званий рядового и сержантского состава. В этой совокупности 5 элементов. Медианой является среднее, третье, т. е. "младший сержант".
- Если же в подобной совокупности четное число данных, причем средние данные различны, то считают, что медианой является пара средних данных: ведь найти их среднее арифметическое нельзя. Если к перечисленным военнослужащим добавить одного с воинским званием старшина, то медианой совокупности, состоящей из 6 элементов, является пара "младший сержант и сержант".

# Квантили

---

- делят ряд распределения на 4, 10 и т.д. равных частей
- Квантили, которые делят ряд на 4 равные по объему совокупности, называются квартелями.
- **Проценти** делят множество наблюдений на 100 частей с равным числом наблюдений в каждой.
- Децилы делят множество наблюдений на десять равных частей.

# Квантили легко вычисляются по распределению накопленных частот (по кумуляте).



# Мода

- Модой в статистике называется наиболее часто встречающееся значение признака, т. е. значение, с которым наиболее вероятно можно встретиться в серии зарегистрированных наблюдений
- В дискретном ряду мода ( $M_o$ ) — это значение с наибольшей частотой.
- Если все значения в вариационном ряду встречаются одинаково часто, то считают, что этот ряд **не имеет моды**
- В интервальном ряду (с равными интервалами) модальным является класс с наибольшим числом наблюдений. Значение моды находится в его пределах и вычисляется по формуле

# Мода

$$M_o = x_0 + h \frac{n_2 - n_1}{(n_2 - n_1) + (n_2 - n_3)},$$

- где  $x_0$  - начальное значение модального интервала, т. е. интервала, который содержит моду;  $h$  - длина модального интервала;  $n_2$  - частота модального интервала;  $n_1$  - частота интервала, предшествующего модальному;  $n_3$  - частота интервала, следующего за модальным.
- В совокупностях, в которых может быть произведена лишь операция классификации объектов по какому-нибудь качественному признаку, **вычисление моды является единственным способом указать некий центр тяжести совокупности.**