

ОПИСАТЕЛЬНАЯ СТАТИСТИКА

Статистика (лат. *status* – состояние) - наука
изучающая количественные данные о
массовых явлениях жизни

Экономическая статистика

(изучает изменение цен спроса и предложения на товары
прогнозирует рост и падение производства)

Медицинская статистика

(изучает эффективность методов лечения ...)

Демографическая статистика

(изучает численность населения, его состав ...)

СРЕДНИЕ ВЕЛИЧИНЫ

Средней величиной называют показатель который характеризует обобщенное значение признака или группы признаков исследуемой совокупности

- Выборочное среднее
- Мода
- Медиана

СРЕДНИЕ ВЕЛИЧИНЫ

Средним арифметическим нескольких чисел называется число, равное отношению суммы этих чисел к их количеству.

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

В статистике эту величину называют еще **средним значением** или **выборочным средним**

Средний балл

Средняя зарплата

Средний рост

СРЕДНЕЕ АРИФМЕТИЧЕСКОЕ

Пример 1

Ученик в четверти получил по алгебре оценки:

$$5 \ 2 \ 4 \ 5 \ 5 \ 4 \ 4 \ 5 \ 5 \ 5 \quad (n=10)$$

Найдем его средний балл — то есть среднее арифметическое всех членов ряда:

$$\bar{x} = \frac{5 + 2 + 4 + 5 + 5 + 4 + 4 + 5 + 5 + 5}{10} = 4,4$$

СРЕДНИЕ ВЕЛИЧИНЫ

Модой M_o числового ряда называют число которое встречается в этом ряду наиболее часто.

Пример 1 Для ряда:

5 2 4 **5** 5 4 **5** **5** 5 5 **5** **5** **5**

мода $M_o=5$

ошибка в формуле

Пример 2 Для ряда:

4 2 3 5

МОДЫ НЕТ.

ошибка в формуле

СРЕДНИЕ ВЕЛИЧИНЫ

- Медианой Me числового ряда называют число этого ряда (или полусумму двух его чисел) слева и справа от которого на числовой прямой лежит одинаковое количество членов ряда.

Медиана - число ряда которое делит его ровно пополам.

Пример 1 Дан ряд:

5 2 4 5 5 4 4 5 5 5

Ранжируем ряд:

2 4 4 4 5 5 5 5 5 5 (n=10)


$$Me = \frac{5 + 5}{2} = 5$$

Пример 2 Дан ряд:

1 2 4 9 10 4 4 6 5

Ранжируем ряд:

1 2 4 4 4 5 6 4 10 (n=9)



Ряд содержит нечетное число членов \Rightarrow медиана - число
которое находится ровно в середине ряда

$$Me = 4$$

Особенности средних величин

Далеко не всегда имеет смысл вычислять все три характеристики

Во многих ситуациях какая-то из характеристик может не иметь никакого содержательного смысла

Пример 3. Гвозди в магазине продают на вес. Чтобы оценить сколько гвоздей содержится в одном килограмме дядя Вася решил найти массу одного гвоздя. Он взвесил на несколько гвоздей и получил следующий ряд чисел (масса гвоздей в граммах):

4,47 4,44 4,64 4,32 4,45 4,32 4,54 4,58

$$\bar{x} = 4,47 \quad Mo = 4,32 \quad Me = 4,46.$$

Самой подходящей характеристикой по смыслу задачи является среднее арифметическое. Несильно отличается от него и медиана.

Мода здесь вряд ли подойдет

Пример 4. Дан ранжированный ряд представляющий данные о времени дорожно-транспортных происшествий на улицах города в течение одних суток (в виде ч : мин):

0:15	0:55	1:20	3:20	4:10	6:10	6:30
7:15	7:45	8:40	9:05	9:20	9:40	10:15
10:15	11:30	12:10	12:15	13:10	13:50	14:10
14:20	14:25	15:20	15:20	15:45	16:20	16:25
17:05	17:30	17:30	17:45	17:55	18:05	18:15
18:45	18:50	19:45	19:55	20:30	20:40	21:30
21:45	22:10	22:35				

$$\bar{x} = 13,33 \quad M_0 = \quad M_e = \quad .$$

Ответ дает интервальная мода

Пример 5. На школьной спартакиаде проводится несколько квалификационных забегов на 100 метров, по результатам которых в финал выходит ровно половина от числа всех участников. Перед вами результаты всех спортсменов. Какой результат позволяет пройти в финал?

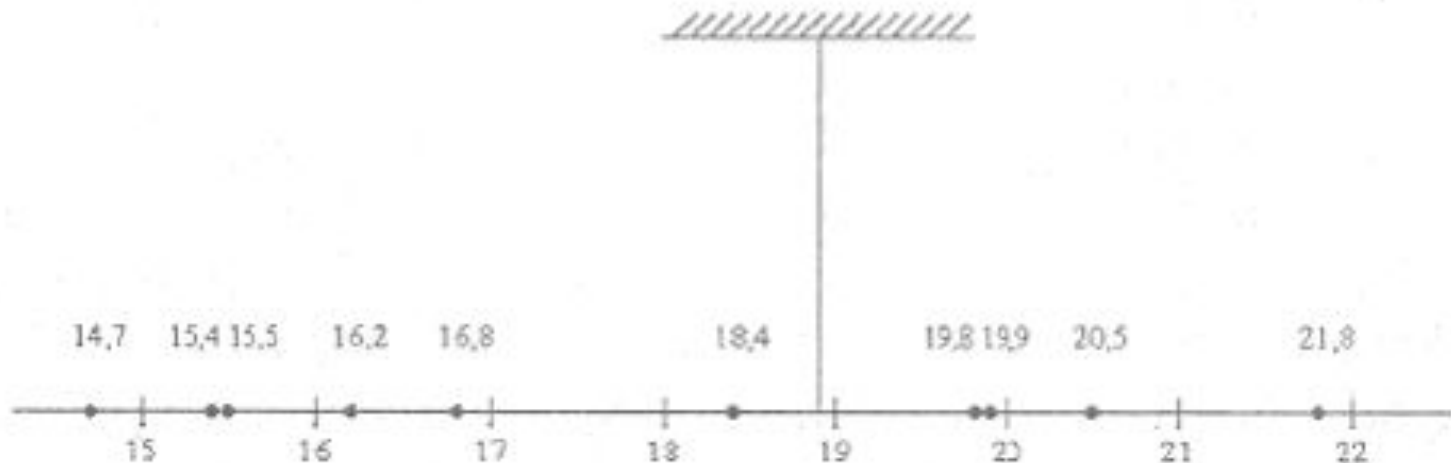
15,5 16,8 21,8 18,4 16,2 32,3 19,9 15,5 14,7 19,8 20,5 5,4

$$\bar{x} = 18,9 \quad Mo = 15,5 \quad Me = 17,6 \quad .$$

Ответ дает медиана $Me = 17,6$

Свойства средних величин

Среднее арифметическое числового ряда является его наиболее естественным «центром». Если нарисовать все члены ряда на числовой прямой, то среднее арифметическое будет центром их масс. Пусть в каждой из точек x_1, x_2, \dots, x_n на числовой оси находятся грузы одинаковой массы. Если теперь «подвесить» числовую ось в точке \bar{x} , то вся система будет находиться в равновесии.



Особенностью **МОДЫ** является еще и то, что ее можно использовать не только в числовых рядах.

Если, например, опросить большую группу учеников, какой школьный предмет им нравится больше всего, то модой этого ряда ответов окажется тот предмет, который будут называть чаще остальных.

Это одна из причин, по которой мода широко используется при изучении спроса и проведении других социологических исследований.

Достоинством **медианы** является ее большая по сравнению со средним арифметическим «устойчивость к ошибкам».

Представим себе, что в таблицу результатов из примера 5 вкралась досадная оплошность: при записи одного из чисел мы пропустили десятичную запятую и вместо 21,8 написали 218. Тогда среднее арифметическое результатов возрастет с 18,9 секунд до 35,25 секунд, а медиана будет по-прежнему 17,6 секунд!

Вычисление средних по таблице частот

результаты наблюдений удобно представлять не в виде последовательности чисел, а в виде таблицы частот

Отметка (a_j)	2	4	5
Абсолютная частота (n_j)	1	3	6
Относительная частота (f_j)	0,1	0,3	0,6

$$\bar{x} = \frac{2 \cdot 1 + 4 \cdot 3 + 5 \cdot 6}{10} = 4,4$$

$$2 \cdot 0,1 + 4 \cdot 0,3 + 5 \cdot 0,6 = 4,4.$$

Среднее арифметическое выборки может быть найдено по любой из двух формул:

$$\bar{x} = \frac{a_1 \cdot n_1 + a_2 \cdot n_2 + \dots + a_k \cdot n_k}{n_1 + n_2 + \dots + n_k}$$

$$\bar{x} = a_1 \cdot f_1 + a_2 \cdot f_2 + \dots + a_k \cdot f_k$$

Для вычисления моды нужно найти максимальное значение в столбце абсолютных или относительных частот и выбрать соответствующее ему значение числового ряда.

Если максимальных частот в таблице несколько, то выборка не имеет моды (будет полимодальной).

Отметка (a_i)	2	4	5
Абсолютная частота (n_i)	1	3	6
Относительная частота (f_i)	0,1	0,3	0,6

$$M_0=5$$

Для вычисления **медианы** нужно найти первое значение накопленной частоты, превосходящее 0,5, и выбрать соответствующее ему значение числового ряда. В нашем случае накопленная частота впервые превосходит 0,5 только в последней строке таблицы, значит, медианой выборки будет 5.

Отметка (a_i)	2	4	5
Абсолютная частота (n_i)	1	3	6
Относительная частота (f_i)	0,1	0,3	0,6

0,4

$Me=5$

Размах числового ряда

Размах R - это разность наибольшего и наименьшего значений ряда данных

Пример 6. Дан ряд:

1 4 7 4 12 5 3 9 3

$$R=12-1=11$$

Пример 7. Температура на Меркурии колеблется
от -150° до $+350^{\circ}$

$$R=350^{\circ} - (-150^{\circ}) = 500^{\circ}$$

$$\bar{x} = +15^{\circ}$$

ЗАДАЧИ

1) На стадионе «Локомотив» была зафиксирована следующая посещаемость первых четырех футбольных матчей: 24000, 18000, 22000, 24000.

Какова была средняя посещаемость этих матчей?

Сколько зрителей должно посетить следующий матч, чтобы средняя посещаемость выросла?

ЗАДАЧИ

2) Президент компании получает зарплату 100 000 руб., четверо его заместителей получают по 20 000 руб., а 20 служащих компании - по 10 000 руб.

Найдите все средние характеристики (среднее арифметическое, моду, медиану) зарплат в компании.

Какую из этих характеристик выгоднее использовать президенту в рекламных целях?

ЗАДАЧИ

- 3) Каждое число исходного числового ряда увеличили на 10. Что произойдет с его средним арифметическим? модой? медианой?
- 4) Все числа исходного числового ряда увеличили в два раза. Что произойдет с его средним арифметическим? модой? медианой?
- 5) Найдите для числового ряда 1, 2, 3, 4, x все возможные значения x , при которых:
- а) среднее арифметическое ряда равняется 3;
 - б) мода равняется 3;
 - в) медиана равняется 3.