

Свои конспекты прислать мне **ТОЛЬКО В**
ЛИЧНОЕ СООБЩЕНИЕ В КОНТАКТ

<https://vk.com/id588363475>

РАБОТЫ В КОММЕНТАРИЯХ НА САЙТЕ НЕ
ПРИНИМАЮ!!!

Переписать теоретический материал в тетрадь. Переписывается все без сокращений!!!

Тема «Основные понятия статистики. Этапы простой обработки статистических данных»

Каждое статистическое исследование состоит из сбора и обработки информации (статистических данных). Совокупность всех подлежащих обработке объектов называется **генеральной совокупностью**. **Выборочной совокупностью (выборкой)** называется совокупность объектов из генеральной совокупности, отобранных случайным образом. На основе полученных данных проводятся выработка различных прогнозов, оценка их достоверности. Важной задачей, без которой статистические данные теряют всякий смысл, является обработка полученных данных. Конкретные значения выборки, полученные в результате наблюдений, называют реализацией выборки или **вариантами**. Эти данные и подвергаются статистической обработке.

Этапы простейшей статистической обработки статистических данных:

В простейшем виде, порядок преобразований первоначально полученной информации примерно таков:

I) Сначала данные *упорядочивают* и *группируют*. Операция упорядочивания статистических данных (расположение значений статистических данных по возрастанию) в статистике называется *ранжированием*. Последовательность вариантов, расположенных в возрастающем порядке, называется *вариационным рядом* (вариация - изменение).

II) После группировки составляют *таблицы распределения данных*.

Для заполнения таблицы необходимо найти следующие характеристики:

Объем выборки N – число объектов выборки.

Частота варианты n_i – это число, показывающее сколько раз каждая варианта x_i из ряда данных, встречается в выборке.

Далее составляется таблица распределения данных, которая состоит из двух строк. В первой записываются варианты x_i , во второй строке – соответствующие им частоты n_i .

Полученная таким образом таблица является **статистическим распределением или статистическим рядом выборки**.

III) Таблицы распределения позволяют построить *графики распределения данных*

Гистограммой распределения называется ломанная, отрезки которой соединяют точки $(x_1;n_1), \dots, (x_2;n_2), \dots, (x_k;n_k)$, где x_i – варианты откладываются на оси Ox , n_i – соответствующие им частоты откладываются на оси Oy .

После простейшей статистической обработки данных составляют своего рода *паспорт данных* измерения, в котором собрано небольшое количество основных *числовых характеристик* полученной информации, а именно:

1. **Размах измерения** – разница между наибольшим и наименьшим значениями результатов измерений (наибольшей и наименьшей вариантами).
2. **Мода измерения** – чаще других встречающаяся варианта в выборке.
3. **Среднее значение** – частное от деления суммы всех результатов измерения на объем измерения.
4. **Медиана измерения** – средняя варианта в выборке.

Пример:

10 абитуриентов сдают тестирование по математике (тест оценивается от 0 до 5 баллов). В результате тестирования были набраны следующие баллы: 0,1,2,4,5,4,3,4,2,5. Проведите обработку данных. Постройте паспорт данных.

Решение:

Простейшая обработка данных проводится в 3 этапа:

- I) Сначала упорядочим данные. Для этого про ранжируем полученные данные:

0,1,2,2,3,4,4,4,5,5.

Полученная после ранжирования последовательность данных: 0,1,2,2,3,4,4,4,5,5 является вариационным рядом.

- II) Построим статистический ряд (статистическое распределение) выборки. Для этого найдем следующие характеристики:

N (объем выборки)=10

Варианты: $x_1=0$ $x_2=1$ $x_3=2$ $x_4=3$ $x_5=4$ $x_6=5$

Соответствующие им частоты (число, которое показывает сколько раз данная варианта встречается в выборке):

$n_1=1$ $n_2=1$ $n_3=2$ $n_4=1$ $n_5=3$ $n_6=2$

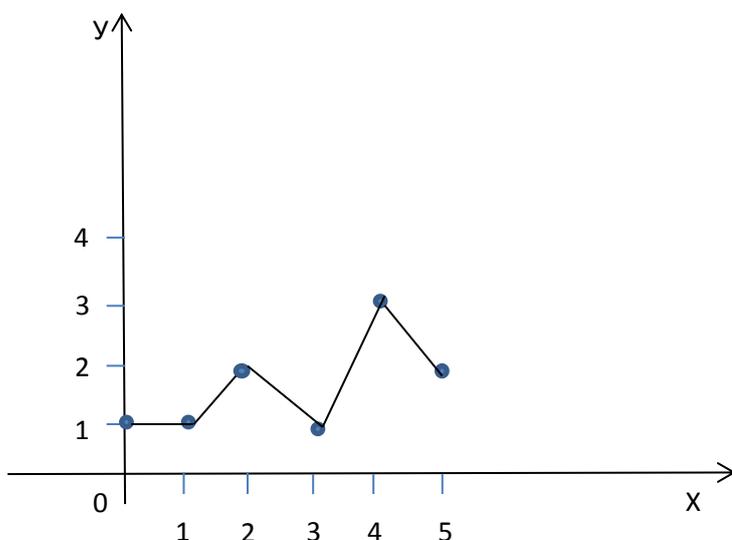
Построим статистический ряд. Занесем полученные данные в таблицу:

x_i	0	1	2	3	4	5	
n_i	1	1	2	1	3	2	Проверка: $1+1+2+1+3+2=10=N$

- III) Построим гистограмму распределения

Для этого в системе координат отложим точки с координатами $(x_i; n_i)$ (данные смотри в статистическом ряде (в таблице пункта II)) x_i – варианты откладываются на оси Ox , n_i – соответствующие им частоты

откладываются на оси Оу. Затем соединим построенные точки и получим ломанную, которая и является гистограммой распределения.



После простейшей обработки статистических данных построим паспорт данных. Для этого найдем следующие характеристики:

1. **Размах измерения** – разница между наибольшей и наименьшей вариантами: $5-0=5$
2. **Мода измерения** – чаще других встречающаяся варианта в выборке: $x_5=4$ (встречается в выборке чаще других $n_5=3$ раза)
3. **Среднее значение** – частное от деления суммы всех результатов измерения на объем измерения:

Смотри в таблице: складываются произведения всех вариантов на их частоты и все это делится на объем выборки

$$\frac{0 \cdot 1 + 1 \cdot 1 + 2 \cdot 2 + 3 \cdot 1 + 4 \cdot 3 + 5 \cdot 2}{10} = \frac{30}{10} = 3$$

4. **Медиана измерения** - средняя варианта в выборке:

Складываются все варианты и делятся на их количество

$$\frac{0+1+2+3+4+5}{6} = \frac{15}{6} = 2,5$$