# Математическая статистика

Лекция 5

Математическая статистика изучает математические методы систематизации, обработки и использования статистических данных для научных и производственных целей

Основной метод обработки данных – *выборочный* 

Основа - *теория* вероятности, в которой изучаются математические модели реальных случайных явлений



Математическая статистика связывает реальные случайные явления и их математические вероятностные модели.

Математическая статистика возникла в 17 веке

ОППОВВОМОННО С ТООВИОЙ ВОВОЛЕНОСТИ

## Математическая статистика

Математическая статистика занимается изучением закономерностей, которым подчинены массовые случайные явления.



## Случайная величина

**Случайная величина** – это величина, значение которой зависит от случая.

Случайные величины могут быть *непрерывными*, т.е. принимать любые значения в некотором интервале (например, температура воздуха, измеряемая на данной метеостанции в случайный момент времени в течение года" - случайная величина).



## Случайная величина

Случайные величины могут быть *дискретными* т.е. принимать только конечное или счетное множество определенных значений (например, число очков при бросании игральной кости; число телефонных звонков, поступающих конкретному абоненту в течение суток). Мы будем рассматривать дискретные случайные величины.



# Числовые характеристики дискретных случайных величин



## Математическое ожидание (среднее значение)

Определение. *Математическим* ожиданием называется сумма произведений значения случайной величины на её вероятность:

$$Mx = \sum_{i} x_{i} p(x_{i})$$

# Дисперсия случайной величины

**Определение.** *Дисперсией* случайной величины *х* называется математическое ожидание квадрата отклонения значения величины от ее математического ожидания: Dx = M(x-Mx)<sup>2</sup>, для дискретной случайной величины:

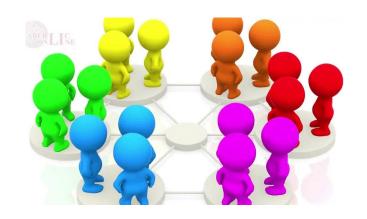
$$Dx = \sum_{i} (x_i - Mx)^2 p(x_i)$$

Пусть требуется оценить совокупность однородных объектов относительно некоторого качественного или количественного признака. Вся совокупность объектов, подлежащих оценке, называется **генеральной совокупностью** (ГС). При сплошном контроле обследуют каждый элемент ГС. Однако сплошной контроль недопустим при больших объемах ГС или в случае его разрушительности. Основным методом математической статистики

является выбо

**Выборочной совокупностью** (ВС), или просто выборкой, называют часть элементов ГС, отобранных для изучения требуемого признака. Объемом совокупности (ГС или ВС) называют число ее элементов. Выборка называется репрезентативной (представительной), если ее элементы относительно изучаемого признака правильно представляют всю ГС, то есть их качественные характеристики близки. Для обеспечения репрезентативности выборки нужно, чтобы у каждого элемента ГС была одинаковая вероятность попасть в нее.





**Генеральная совокупность** – совокупность всех исследуемых объектов

Выборочная совокупность (выборка) – совокупность случайно отобранных объектов

Случайный отбор – это такой отбор, при котором все объекты генеральной совокупности имеют одинаковую вероятность попасть в выборку

Числовые характеристики выборки

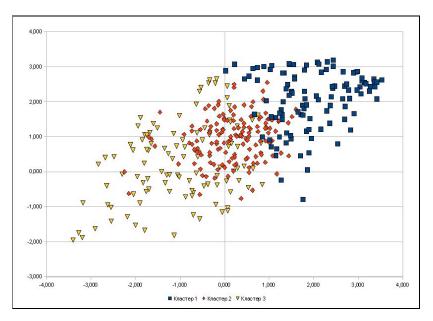
Объём выборки – это число равное количеству объектов генеральной или выборочной совокупности

## Пример

Из 10000 изделий для контроля отобрали 100 изделий.

Объем генеральной совокупности равен 10000, объем выборки – 100





Для статистической обработки результаты исследования объектов, составляющих выборку, представляют в виде **числовой выборки** (последовательность чисел).

Разность между наибольшим значением числовой выборки и наименьшим называется размахом выборки

Для статистической обработки результаты исследования объектов, составляющих выборку, представляют в виде числовой выборки (последовательность чисел).

Разность между наибольшим значением числовой выборки и наименьшим называется размахом выборки

Рассмотрим числовую выборку объема п, полученную при исследовании некоторой генеральной совокупности

Значение  $x_1$  встречается в выборке  $n_1$  раз  $x_2$  встречается  $n_2$  раза  $x_n$  встречается  $n_n$  раз

числа  $n_1, n_2, ..., n_n$  называются частотами значений.

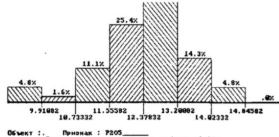
значений, при этом

$$\frac{n_1}{n}, \frac{n_2}{n}, ..., \frac{n_n}{n}$$
 Отношения частот к объему выборки 
$$\frac{n_1}{n}, \frac{n_2}{n}, ..., \frac{n_n}{n}$$
 называются **относительными частотами**

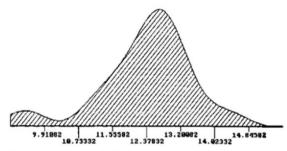
$$n_1 + n_2 + ... + n_n = n$$
,  $\frac{n_1}{n} + \frac{n_2}{n} + ... + \frac{n_n}{n} = 1$ 

Статистическим **распределением** выборки называют перечень вариант и соответствующих им частот или относительных частот. Статистическое распределение удобно представлять в табличной форме.

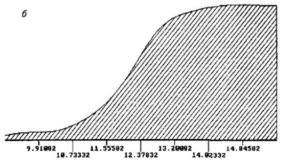
Закон распределения называют перечень вариант и соответствующих им вероятностей этих вариант.



Признак : Р205\_ Гистогранна распределения, наблюдений 63



Признак : Р205\_



63 жинадом фикция распределения, наблюдений

#### Статистическое распределение

X <sub>1</sub>	<b>x</b> <sub>2</sub>	X3	 Xn
$n_1$	n <sub>2</sub>	n <sub>3</sub>	 n <sub>n</sub>

#### Закон распределения

$\mathbf{x}_1$	X2	X3	 Xn
n <sub>1</sub> /n	n <sub>2</sub> /n	n <sub>3</sub> /n	 n <sub>n</sub> /n

### Пример1.

Для выборки определить объем, размах, найти статистический ряд и выборочное распределение: 3, 8, -1, 3, 0, 5, 3, -1, 3, 5

Объем: n = 10, paзмах = 8 - (-1) = 9

#### Статистическое распределение

Xi	-1	0	3	5	8
ni	2	1	4	2	1

#### Закон распределения

$x_i$	-1	0	3	5	8	0
n <sub>i</sub> /n	0,2	0,1	0,4	0,2	0,1	

## Пример 2.

В результате 10 опытов получена выборка 2,2,2,3,4,4,6,6,6,6, тогда для нее закон распределения будет?

#### Статистическое распределение

$\mathbf{X}_{i}$	2	3	4	6
$\mathbf{n}_{i}$	3	1	2	4

#### Закон распределения

$\mathbf{X}_{i}$	2	3	4	6	
$\mathbf{p}_{i}$	0,3	0,1	0,2	0,4	

#### Графические изображения выборки

Если выборка задана значениями и их частотами или статистическим рядом, то строится полигон – 'то ломаная с вершинами в точках  $(x_1;n_1),(x_2;n_2),...,(x_n;n_n)$ 

#### Полигон частот

