

# Элементы математической статистики




Тема:

*Предмет и задачи математической статистики. Представление данных.*



# Определение


Современную математическую статистику определяют как **науку о принятии решений в условиях неопределённости.**



**Задача математической статистики** состоит в создании методов сбора и обработки статистических данных для получения научных и практических выводов.


**Математическая статистика** возникла в XVII веке и развивалась параллельно с теорией вероятностей.

Дальнейшее развитие (вторая половина XIX века – начало XX века) обязано, в первую очередь, П. Л. Чебышеву, А. А. Маркову, А. М. Ляпунову, а так же К. Гауссу, А. Кетле, Ф. Гальтону, К. Пирсону и другие. XX век – советские учёные : В. И. Романовский, Е. Е. Слуцкий, А. Н. Колмогоров. Английские: Стьюдент, Фишер, Смирнов. Американские: С. Нейман, Вальд.



## **Генеральная и выборочная совокупность**

Пусть требуется изучить совокупность однородных объектов относительно некоторого качественного или количественного признака, характеризующего эти объекты.



# Определения выборочной и генеральной совокупности

**Def:** Выборочной совокупностью или выборкой называют совокупность случайно отобранных объектов.


**Def:** Генеральной совокупностью называют совокупность объектов из которых производится выборка.

**Def:** Объемом совокупности называют число объектов этой совокупности.



# Репрезентативность выборки.

Для того чтобы по данным выборки можно было достаточно уверенно судить об интересующем признаке генеральной совокупности, необходимо, чтобы объекты выборки правильно его представляли. Другими словами, выборка должна правильно представлять пропорции генеральной совокупности. Это требование коротко формулирует так: выборка должна быть **репрезентативной (представительной)**.




В силу закона больших чисел можно утверждать, больших что выборка будет репрезентативной, если её осуществлять случайно: каждый объект выборки отобран случайно из генеральной совокупности, если все объекты имеют одинаковую вероятность попасть в выборку.



# Признаки и их свойства.

В общем смысле под словом «признак» подразумевают свойство, проявлением которого один предмет отличается от другого ( в биологии: характерные особенности в строении и функциях живого).





# Признаки делятся на **качественные** и **количественные**.

**Качественные:** окраска цветов, особи разного пола и масти (серые, вороные, гнедые, пёстрые и другие), цвет глаз и волос. **Альтернативные признаки** ( женщина и мужчина, высокий и низкий).

**Количественные признаки** поддаются непосредственному измерению или счёту. Их делят на мерные или метрические, и счётные или меристические.

**Мерные признаки:** длина колосьев урожайность, мясная и молочная, продуктивность животных.

**Счётные признаки:** число зёрен или колосков в колосьях, яйценоскость и другие.

**Порядковые признаки** – объекту приписывают числа или баллы.



# Определение:

Характерным свойством биологических признаков является **варьирование** величины признаков в определённых пределах при переходе от одной единицы наблюдений к другой. Эти колебания величины одного и того же признака, наблюдаемые в массе однородных членов статистической совокупности, называют **вариациями** (от латинского *variatio* – изменение, колебания), а отдельные числовые значения варьирующего признака принято называть **вариантами** (от латинского *variants, variantis* - различимый, изменяющийся)



# Определение

**Def:** Вариационным рядом или рядом распределения называют двойной ряд чисел, показывающий, каким образом числовые значения признака связаны с их повторяемостью в данной статистической совокупности.



# Определение

Числа, показывающие, сколько раз отдельные варианты встречаются в данной совокупности называются **частотами** и обозначается строчной буквой латинского алфавита  $n_i$

Общая сумма частот вариационного ряда равна объёму данной совокупности, т.е.

$$\sum_{i=1}^k n_i = n$$

$n$ -общее число наблюдений.



# Определение

Частоты выражают не только абсолютными, но и относительными числами в долях единицы или в процентах от общей численности вариантов, составляющих данную совокупность. В таких случаях частоты называют **относительными**.

$$f_i = \frac{n_i}{n}$$



# ИНТЕРВАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Статистическое распределение можно задать так же в виде последовательности интервалов и соответствующих им частот ( в качестве частоты, соответствующей интервалу, принимают сумму частот, попавших в этот интервал).



# Пример:

Из урожая картофеля, собранного на одном из опытных делянок, случайным способом, т.е. наугад, отобрано 20 клубней, в которых подсчитывали число глазков. Результаты подсчёта оказались следующими:

**2 5 3 6 4 7 4 5 6 6 5 9 5 6 1 0 8 1 2 9 7 6**

# Проранжируем ряд

Под **ранжированием** (от французского ranger – выстраивать в ряд по ранжиру т.е. по росту) понимают расположение членов ряда в возрастающем (или убывающем) порядке.

**2 3 4 4 5 5 5 5 6 6 6 6 6 7 7 8 9 9 10 12**

**Распределение**

**абсолютных  
частот**

$X_i$	2	3	4	5	6	7	8	9	10	12
$n_i$	1	1	2	4	5	2	1	2	1	1

Растянутое, некомпактное представление: строится интервальное распределение .

Число интервалов (классов)  $K$  определяется по правилу.



# Правило Стерджеса

$$K=1+3,32 \lg n \text{ (} n\text{-число измерений } <100)$$

Классы  $K=5$

$$i=1,2,3,4,5.$$

Интервальное  
распределение  
абсолютных  
частот

$X_i - X_{i+1}$	[2-4)	[4-6)	[6-8)	[8-10)	[10-12]
$n_i$	2	6	7	3	2
$n_i^H$	2	8	15	18	20

Накопительная  
абсолютная частота