

Теория вероятностей и математическая статистика

Элементы математической статистики



Основные понятия и задачи математической статистики

- **Математическая статистика** – раздел математики, изучающий математические методы сбора, систематизации, обработки и интерпретации результатов с целью выявления статистических закономерностей
- **Задача:** по имеющимся данным эксперимента установить неизвестное распределение вероятностей или объективно оценить параметры распределения

Основные понятия и задачи математической статистики

- **Объект изучения:** генеральные совокупности, исследуемые на основе выборки
- **Генеральная совокупность** – множество результатов наблюдений за значениями одного или нескольких признаков, которые могут быть сделаны при данном комплексе условий (*рассматривают как случайную величину X*)
- **Выборочная совокупность (выборка)** – множество результатов случайно отобранных из генеральной совокупности

Выборочное наблюдение

- Метод отбора определяет возможность продолжения участия отобранной единицы в процедуре отбора:
 1. Повторная выборка
 2. Бесповторной выборка
- Выборка должна быть **репрезентативной** – правильно отражать пропорции генеральной совокупности. Это достигается случайностью отбора (все объекты генеральной совокупности имеют одинаковую вероятность быть отобранными)

Вариационные ряды и их графическое изображение

- Различные значения изучаемого признака (СВ X) называют **вариантами** (обозначаются через x)
- **Вариационный ряд** – ранжированный в порядке возрастания/убывания ряд вариантов с соответствующими им весами
- При большом объеме выборки (большом числе наблюдений n) варианты **группируются** (разбиваются на интервалы)

Вариационные ряды и их графическое изображение

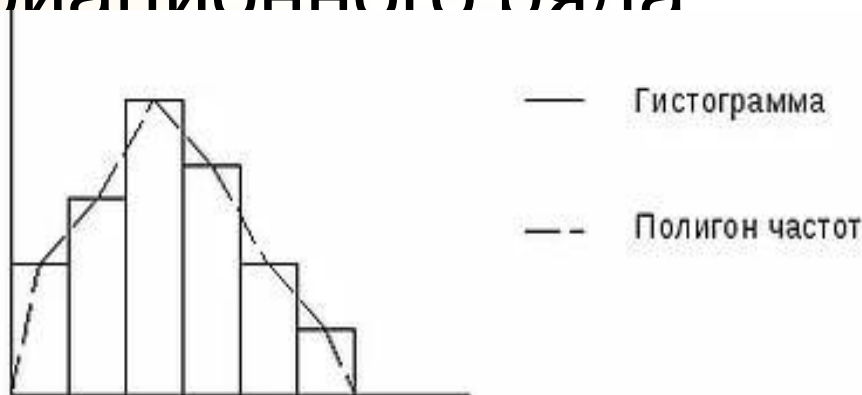
- Рекомендуемое число интервалов:
формула Стерджеса: $m = 1 + 3,221 * \lg n$
- Величина интервала: $h = \frac{x_{max} - x_{min}}{m}$
- **Частоты** – числа, которые показывают, как часто встречаются те или иные варианты в ряду распределения
- Частоты, выраженные в долях единицы или в процентах к итогу, называются **частностями** (ω), соответственно их сумма равна 1 или 100%

Виды рядов распределения:

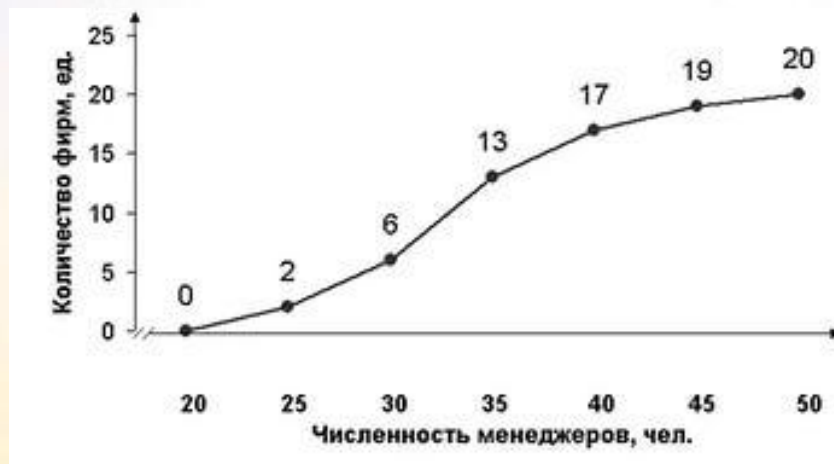
1. **Атрибутивные** (группы строятся по качественному признаку).
2. **Вариационные** (группы строятся по количественному признаку):
 - **Дискретные** (варианты отличаются на постоянную величину).
 - **Интервальные** (группы строятся по признаку, принимающему в определенном интервале любые значения).

Вариационные ряды и их графическое изображение

- **Полигон** – это ломаная кривая, строящаяся на основе прямоугольной системы координат.
- **Гистограмма** применяется для изображения интервального вариационного ряда



- **Кумулятивная кривая (кумулята)** – кривая накопленных частот. Накопленная частота показывает, сколько наблюдалось вариантов со значением признака, меньших x .



Показатели вариации

- **Размах вариации** характеризует границы вариации изучаемого признака $R = X_{max} - X_{min}$
- **Средняя арифметическая** вариационного ряда:

$$\bar{x} = \frac{\sum x_i}{n} \text{ или } \bar{x} = \frac{\sum x_i * f_i}{\sum f_i}$$

- **Мода (M_o)** – варианта, которая чаще всего встречается в данной совокупности.
- **Медиана (M_e)** – варианта, которая находится в середине вариационного ряда. $N_{me} = \frac{n+1}{2}$

Показатели вариации

Дисперсия

*Среднее
квадратическое
отклонение*

*Коэффициент
вариации*

*Среднее
линейное
отклонение*