Методы математической статистики

- **Математическая статистика** — раздел математики, разрабатывающий методы регистрации, описания и анализа данных наблюдений и экспериментов с целью построения вероятностных моделей массовых случайных явлений.
- Основана на использовании идеализированного предположения о существовании бесконечно большого числа измерений.

- **Генеральная совокупность** – множество всех измерений $n \to \infty$, $n \geq 20$
- **Выборочна совокупность** или **выборка** – набор данных конечного объема, извлекаемых из генеральной совокупности $2 < n \leq 20$.
- **n** – объем выборки.

- Математическая статистика позволяет на основе выборочных измерений делать заключения о поведении генеральной совокупности.

- **Случайная величина** — это величина, которая принимает в результате опыта одно и только одно значение, наперед неизвестное и зависящее от случайных причин, которые заранее не могут быть учтены.
- **Случайная величина** — это переменная величина, принимающая значения в зависимости от случая.

- Случайное событие - ***качественная характеристика*** испытаний, то случайная величина - его ***количественная характеристика***.
- Случайная величина – это не число, а функция случая.

Характеристики случайных величин

Случайные величины обозначают заглавными буквами X, Y, Z, а их значения прописными xi, yi, zi.

Случайная величина определяется

- областью ее измерения;
- вероятностью попадания в тот или иной интервал из области измерения.

Случайная величина бывает

- Непрерывной – принимает все значения из некоторого конечного или бесконечного промежутка; их пронумеровать нельзя;
- Дискретной – принимает отдельные изолированные возможные значения с определенными вероятностями; их можно пронумеровать.

Функции распределения случайной величины

- В математической статистике поведение случайных величин принято описывать специальными **функциями**, связывающие значение, которое принимает случайная величина с вероятностью его реализации.

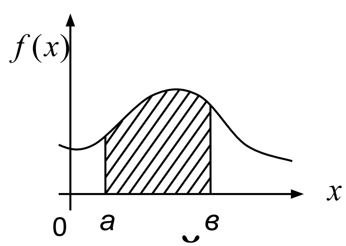
- Они могут быть представлены в виде:

- таблицы (ряд распределения),

X	x_1	x_2	x_3	…	x_n
P (x)	P_1	P_2	P_3	…	P_n

$$\sum_{i=1}^{n} p_i = 1$$

- графика,



- явной функциональной зависимости. $f(x)=ax^3+bx^2+cx$

- Они могут быть

- интегральными,
- дифференциальными

Интегральная функция распределения случайной величины

Или функция накопления F(X)

Значение интегральной функции распределения в точке x есть вероятность того, что случайная величина X примет значения меньше или равное x.

$$F(x)=P(X\le x)$$

Свойства

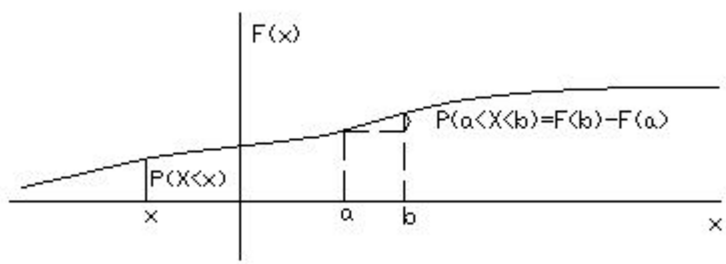
1. Значение функции распределения принадлежит отрезку [0,1]: $0 \le F(x) \le 1$.
2. Функции распределения есть неубывающая функция: $x_1<x_2$ $F(x_1)<F(x_2)$
3. Вероятность того, что случайная величина X примет значение, заключенное в интервале (a, b), равна приращению функции распределения на этом интервале:

$$P(a < X < b) = F(b) – F(a).$$

4. Если все возможные значения случайной величины X принадлежат интервалу (a, b), то

$$F(x) = 0 \text{ при } x \le a; \quad F(x) = 1 \text{ при } x \ge b$$

Или $F(-\infty) = 0$ $\quad F(\infty) = 1$



Пример. Закон распределения дискретной случайной величины задан таб

$X = x_i$	-2	-1	2	5	7
$P(X = x_i) = p_i$	0,1	0,2	0,4	0,1	0,2

Построить функцию распределен $F(x)$ и ее график.

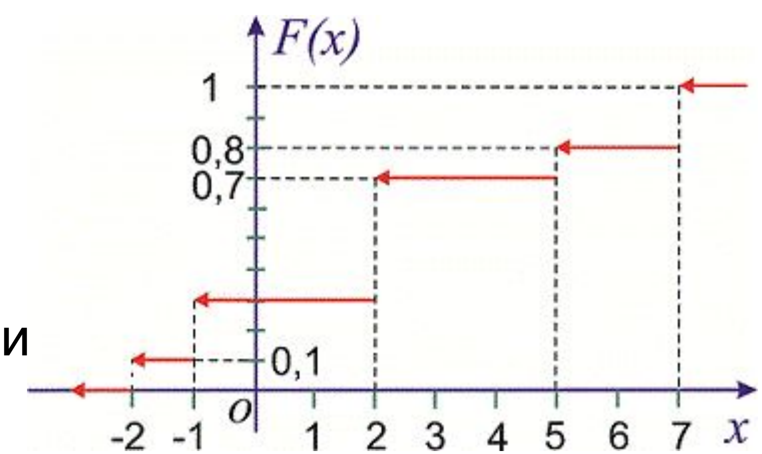
Решение. Согласно свойствами функ $F(x)$ получим приведенные дальше значение.

$$F(x) = P(X < -2) = 0;$$
$$F(x) = P(X < -1) = P(X = -2) = 0,1;$$
$$F(x) = P(X < 2) = P(X = -2) + P(X = -1) = 0,3;$$
$$F(x) = P(X < 5) = P(X = -2) + P(X = -1) + P(X = 2) = 0,7;$$
$$F(x) = P(X < 7) = 0,1 + 0,2 + 0,4 + 0,1 = 0,8;$$
$$F(x) = P(X > 7) = P(X < 7) + P(X = 7) = 0,8 + 0,2 = 1.$$

Компактно функция распределения иметь запись

$$F(x) = P(X < x) = \begin{cases} 0, & x \le -2; \\ 0,1, & -2 < x \le -1; \\ 0,3, & -1 < x \le 2; \\ 0,7, & 2 < x \le 5; \\ 0,8, & 5 < x \le 7; \\ 1, & x > 7. \end{cases}$$

График функции распределения имеет ви



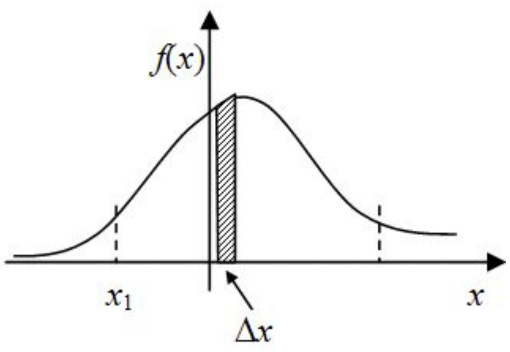
Дифференциальная функция распределения случайной величины

- Или функция плотности вероятности
- *называется* первая производная от интегральной функции распределения, т. е. $F'(x)=f(x)$.

$$f(x)dx=P(x \le X \le x+dx)$$

- *Смысл:* Произведение $f(x)dx$ есть вероятность того, что случайная величина получит значение из интервала от x до dx.

- *Графически*: вероятность попадания значения случайной величины в бесконечно малый интебрвал dx есть площадь криволинейной трапеции под графиком функции $f(x)$.



- функция плотности вероятности существует только для непрерывных случайных величин.

Свойство дифференциальной функции распределения случайной величины

- **Свойство 1**. дифференциальная *функция* распределения *f(x)* не отрицательна для любого *x* из ее области определения.

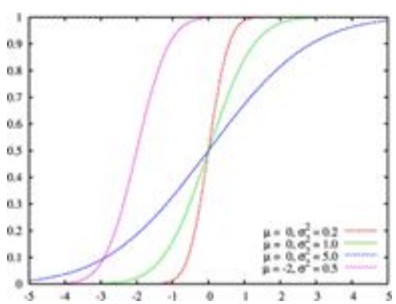
$$f(x) \geqslant 0$$

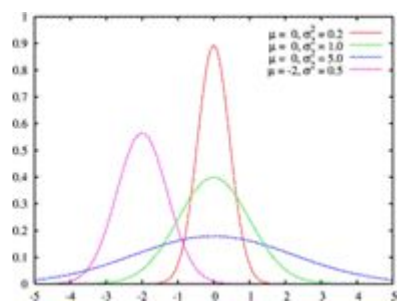
- **Свойство 2**. *Вероятность* попадания непрерывной случайной величины *x* в *интервал* [а,в] равен определенному интегралу от функции плотности распределения *f(x)* на этом интервале

$$P(\alpha \leqslant x \leqslant \beta) = \int\limits_{\alpha}^{\beta} f(x)dx$$

- **Свойство 3**. Интегральная *функция* распределения случайной величины может быть выражена через функцию плотности вероятностей *по* формуле

$$F(x) = \int\limits_{-\infty}^{x} f(x)dx$$





- **Свойство 4**. *Площадь* под кривой плотности распределения на всей ее области определения равен единице

$$\int\limits_{-\infty}^{+\infty} f(x)dx = 1$$

Параметры распределения случайных величин

- Математическое ожидание непрерывной случайной величины характеризует ее наиболее вероятное значение (истинное значение при отсутствии систематической погрешности) и определяется через функцию плотности вероятности

$$M(X) = \int_{-\infty}^{\infty} x f(x) dx$$

- Для дискретной случайной величины

$$M(X) = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n$$

$$M(X) = \sum_{i=1}^{\infty} x_i p_i,$$

- Для равновероятных событий

$$P_i = const = P_1 = P_2 = P_3 = .. = P_n = 1/n$$

$$M(X) = \overline{X} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Свойства математического ожидания.

- Постоянный множитель можно выносить за знак математического ожидания:

$$M(C \cdot X) = C \cdot M(X)$$

- Математическое ожидание постоянной величины *C* равно самой этой величине:

$$M(C) = C$$

- Математическое ожидание суммы двух случайных величин равно сумме их математических ожиданий:

$$M(X + Y) = M(X) + M(Y)$$

- Математическое ожидание произведения независимых случайных величин равно произведению математических ожиданий этих величин:

$$M(X \cdot Y) = M(X) \cdot M(Y)$$

Дисперсия случайной величины

Дисперсией дискретной случайной величины – мера рассеивания (разброса значений) случайной величины относительно математического ожидания – это математическое ожидание квадрата ее отклонения:

$$D(X) = M(X - M(X))^2$$

Свойства дисперсии.

- Дисперсия постоянной величины *C* равна 0:

$$D(C) = 0$$

- Если *X*- случайная величина, а *C* – постоянная, то

$$D(C \cdot X) = C^2 D(X)$$

$$D(X + C) = D(X)$$

- Если *X* и *Y*- независимые случайные величины, то

$$D(X + Y) = D(X) + D(Y)$$

- Для дискретных случайных величин

$$D(X) = (x_1 - M(X))^2 p_1 + (x_2 - M(X))^2 p_2 + \ldots + (x_n - M(X))^2 p_n =$$
$$= \sum P_i (x_i - M(X))^2 = \sum P_i (x_i - \bar{x})^2$$

- Для вычисления дисперсий более удобной является формула

$$D(X) = M(X^2) - (M(X))^2$$

- Для непрерывных случайных величин

$$D(X) = \int_{-\infty}^{\infty} (x - M(X))^2 f(x) dx$$

Расчет дисперсии генеральной и выборочной совокупности

Генеральная совокупность

$n \to \infty$, $n \geq 20$

$D \equiv \sigma^2$ – стандартное отклонение характеризует рассеяние, разброс результатов измерений относительно наиболее вероятного значения x, и называется ***генеральной дисперсией***;

σ – называют генеральным средним квадратичным отклонением.

$$\sigma = \sqrt{D}$$

$$\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}}$$

Выборочная совокупность

$2 \leq n \leq 20$

S^2 – ***выборочная дисперсия***

S – ***выборочное стандартное отклонение***

$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}$$