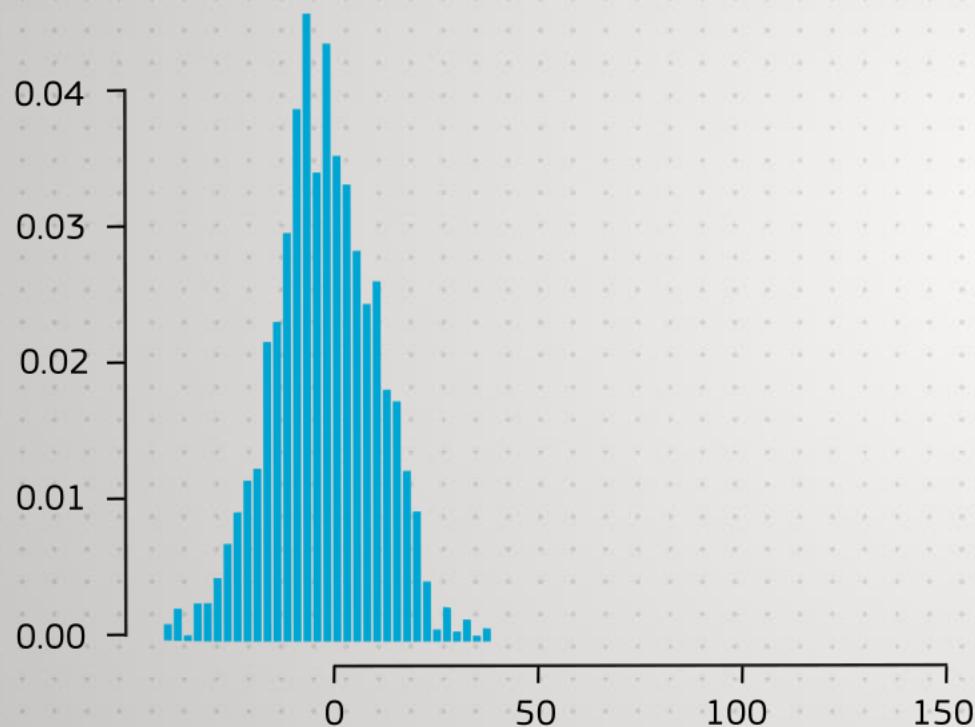


Гистограмма и методы её построения

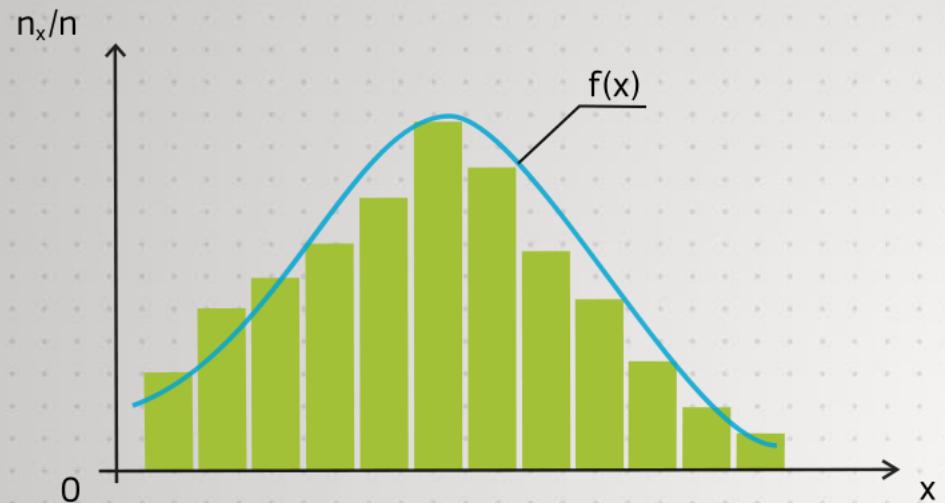
Графический анализ данных

Гистограмма и методы её построения

Гистограмма — это геометрическое изображение эмпирической функции плотности вероятности некоторой случайной величины, построенное по выборке.



Как правильно построить гистограмму?



$F(x)$ — функция плотности

1. Поделить область значений на k равных интервалов.

$$-\infty = x_0 < \dots < x_k = \infty$$

2. Посчитать количество наблюдений n_i .

$$h_i = \frac{n_i}{n(x_{i+1} - x_i)}$$

h — высота столбцов

Свойство функции плотности

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Как определить количество интервалов?

1. Метод Стёрджесса.

$$k = 1 + [\log_2 n]$$

2. $k = \lceil \sqrt{n} \rceil$

Как определить ширину интервалов?

1. Метод Скотта

$$w_i = x_{i+1} - x_i = 3.5 \frac{\sigma^2}{\sqrt[3]{n}}$$

2. Метод Фридмана — Диакониса

$$w_i = x_{i+1} - x_i = \frac{2\tau}{\sqrt[3]{n}}$$