

Теория вероятностей и математическая статистика
Характеристики непрерывных случайных величин. Распределения.

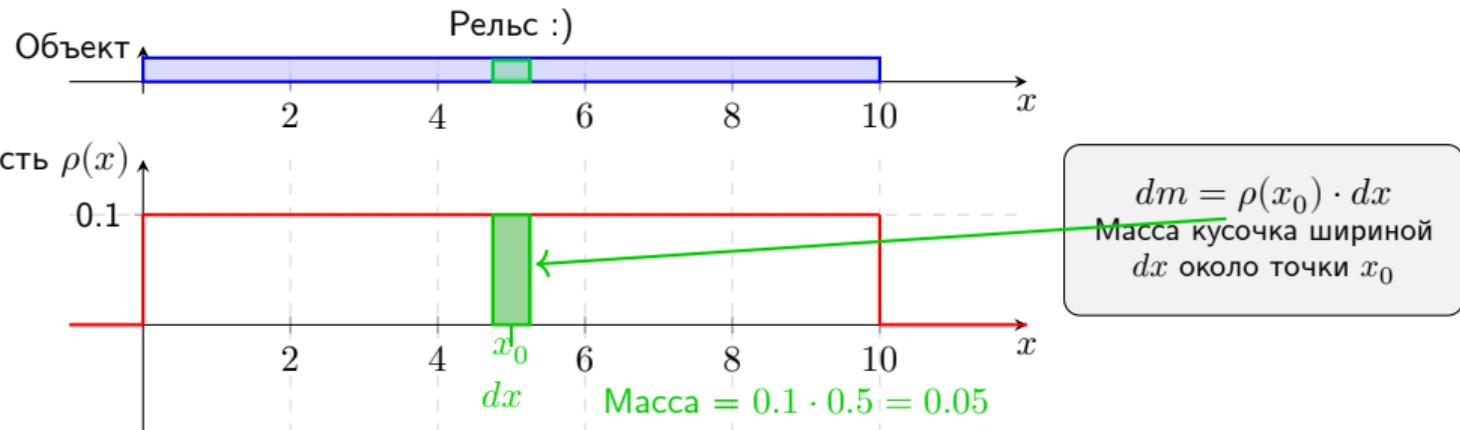
Глеб Карпов

ВШБ Бизнес-информатика

В прошлых сериях...

Физическая аналогия: значение плотности вещества в точке

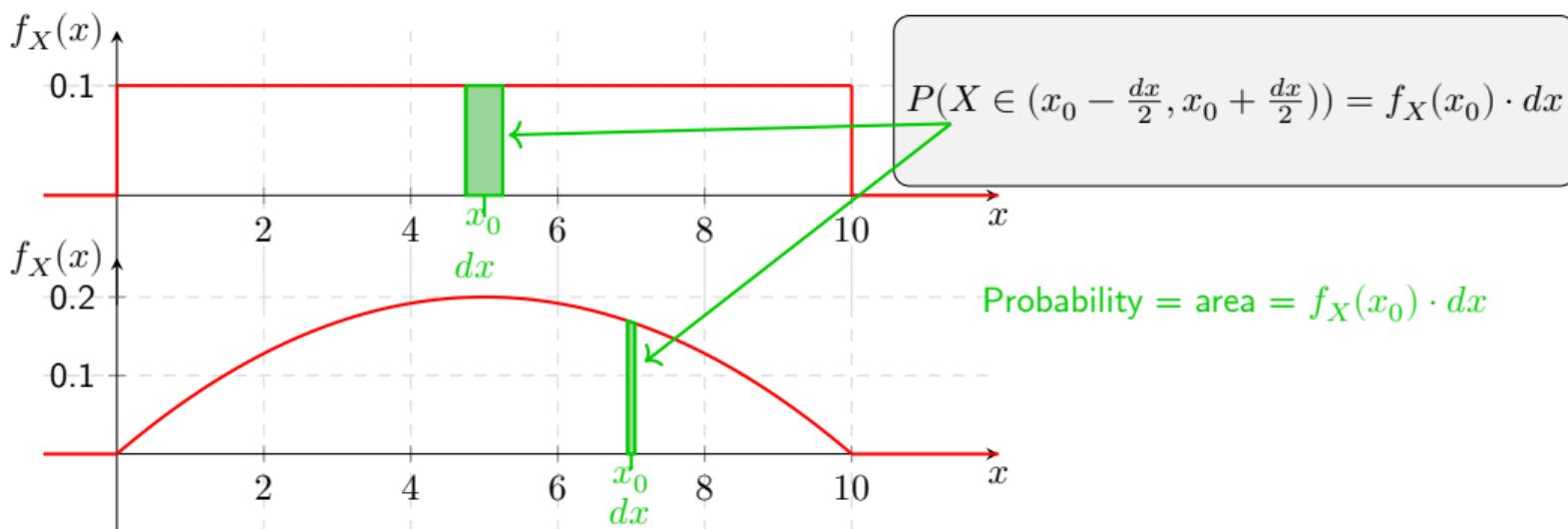
- Само по себе значение функции плотности $\rho(x_0)$ не обозначает массу в точке x_0 . Смысл несет именно произведение плотности на длину, вспомним, чтобы сократились размерности $[kg] = \left[\frac{kg}{m}\right] [m]$



В прошлых сериях...

Значение плотности вероятности в точке

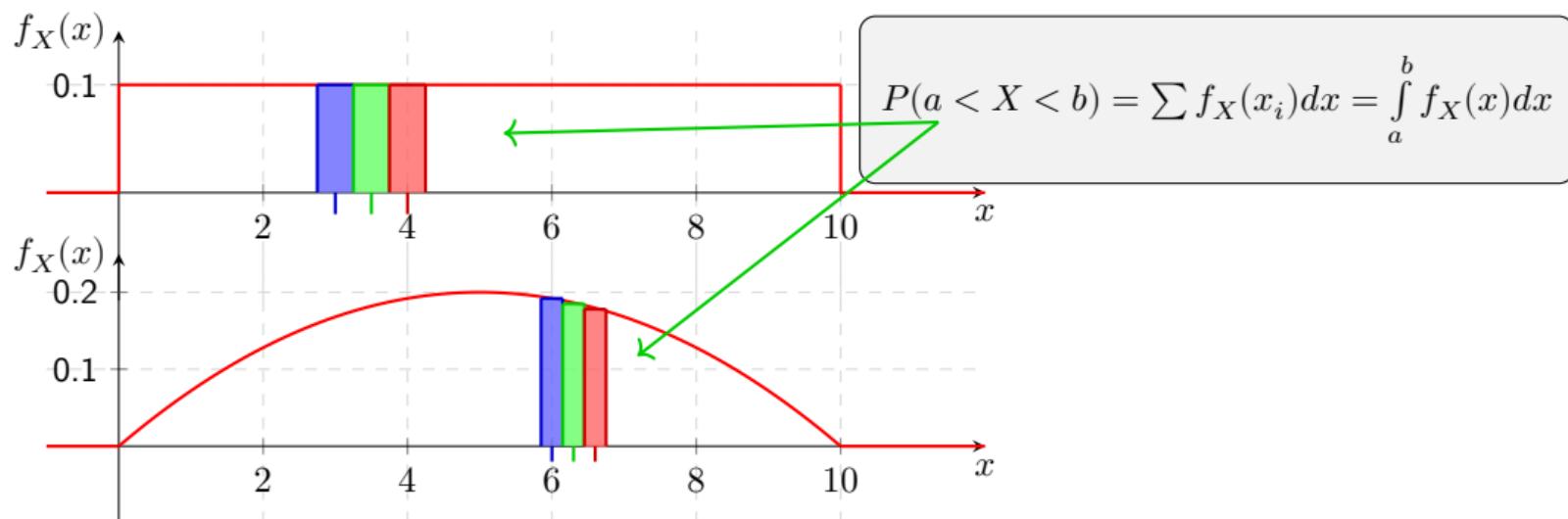
- По аналогии с физикой: само по себе значение функции плотности $f_X(x_0)$ не обозначает вероятность в точке x_0 . Смысл несет именно произведение плотности на длину интервала.



В прошлых сериях...

Вероятность попадания в интервал

- Последовательно складывая такие вероятности попадания в окрестность, получаем площадь под графиком функции плотности от точки a до точки b , которая и интерпретируется нами как $P(a < X < b)$.



Математическое ожидание

- Напомним, что для дискретной СВ мы вычисляем математическое ожидание как 'взвешенную сумму':

$$E[X] = \sum_{i=1}^n x_i P(X = x_i), \text{ вычисляется для } \forall x_i \in \Omega_X.$$

- Если у нас есть другая СВ G , являющаяся функцией от X : $G = g(X)$:

$$E[g(X)] = \sum_{i=1}^n g(x_i) P(X = x_i)$$

- Для непрерывного случая мы заменяем сумму на интеграл:

$$\sum_{i=1}^n x_i P(X = x_i) \Rightarrow \int_{-\infty}^{+\infty} x f_X(x) dx \quad \sum_{i=1}^n g(x_i) P(X = x_i) \Rightarrow \int_{-\infty}^{+\infty} g(x) f_X(x) dx$$

- Финально:

$$E[X] = \int_{-\infty}^{+\infty} x f_X(x) dx$$

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$$

Дисперсия

- Формула для дисперсии остается той же:

$$Var[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2$$

- Напомним доказательство:

$$\begin{aligned} Var[X] &= E[(X - E[X])^2] = E[X^2 - 2XE[X] + (E[X])^2] = \\ &= E[X^2] - 2E[X]E[X] + (E[X])^2 = E[X^2] - (E[X])^2. \end{aligned}$$

Равномерное распределение

- Равномерное распределение на интервале $[a, b]$ имеет постоянную плотность
- Функция плотности вероятности:

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{иначе} \end{cases}$$

- Функция распределения:

$$F_X(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases}$$

- Математическое ожидание и дисперсия:

$$E[X] = \frac{a+b}{2}, \quad Var[X] = \frac{(b-a)^2}{12}$$

Экспоненциальное распределение

- Экспоненциальное распределение моделирует время между событиями в пуассоновском процессе
- Функция плотности вероятности:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

- Функция распределения:

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

- Математическое ожидание и дисперсия:

$$E[X] = \frac{1}{\lambda}, \quad Var[X] = \frac{1}{\lambda^2}$$

Нормальное распределение

- Функция плотности вероятности параметризована двумя константами μ и σ^2 и записывается как:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \forall x \in \mathbb{R}$$

- Параметры влияют на вид функции плотности и, как следствие, на "поведение" нормальной случайной величины. Чтобы подчеркнуть конкретную нормальную величину, с конкретно рассматриваемым набором параметров, мы пишем:

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

i Характеристики нормальной величины

Имеет место интересное совпадение параметров и характеристик нормальной случайной величины:

$$E[X] = \mu, \quad Var[X] = \sigma^2$$

Будьте внимательны, в общем случае параметры и характеристики случайной величины совпадать не обязаны.

Влияние параметров на функцию плотности

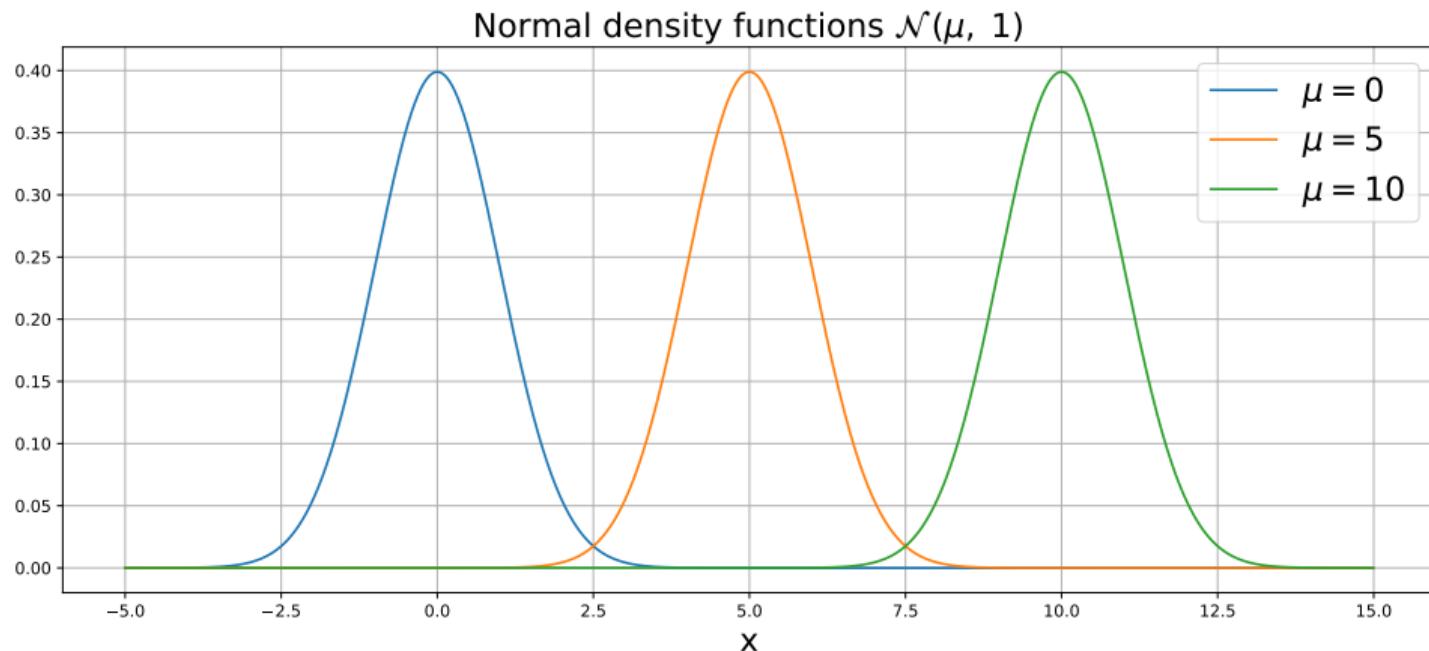


Рис. 1: Влияние параметра μ

Влияние параметров на функцию плотности

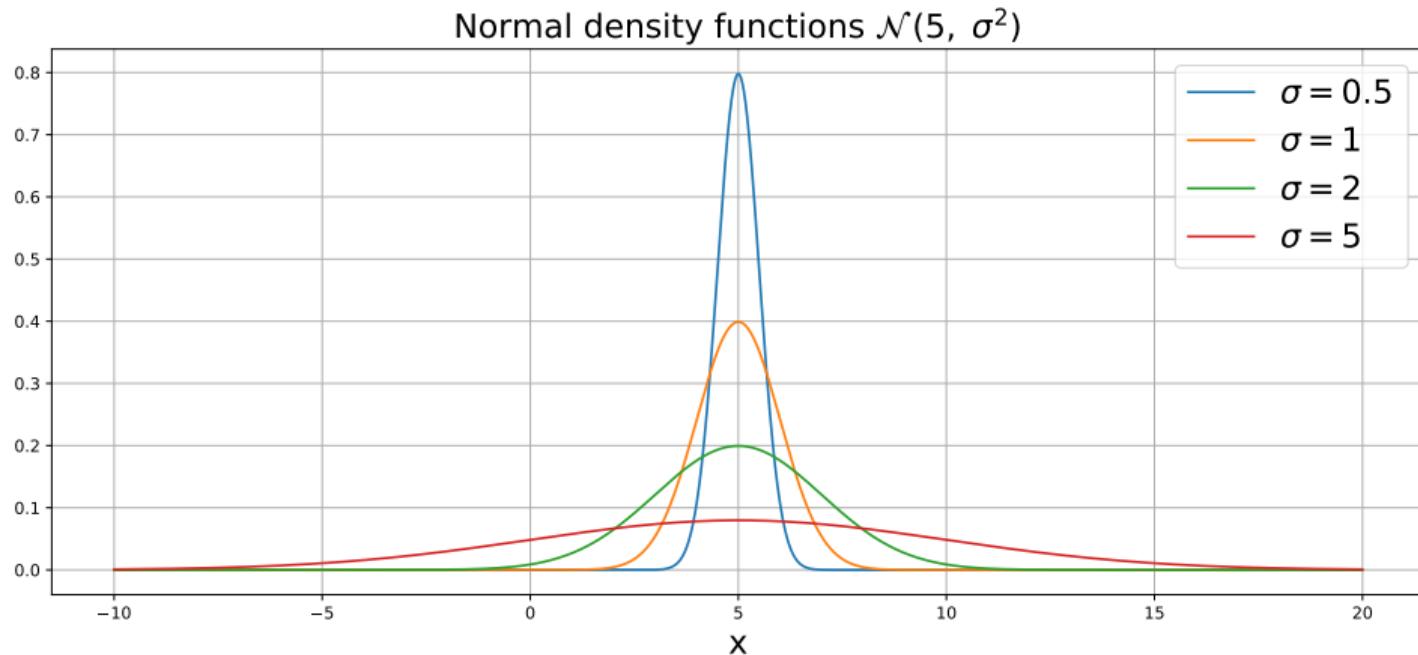


Рис. 2: Влияние параметра σ

Влияние параметров на функцию плотности

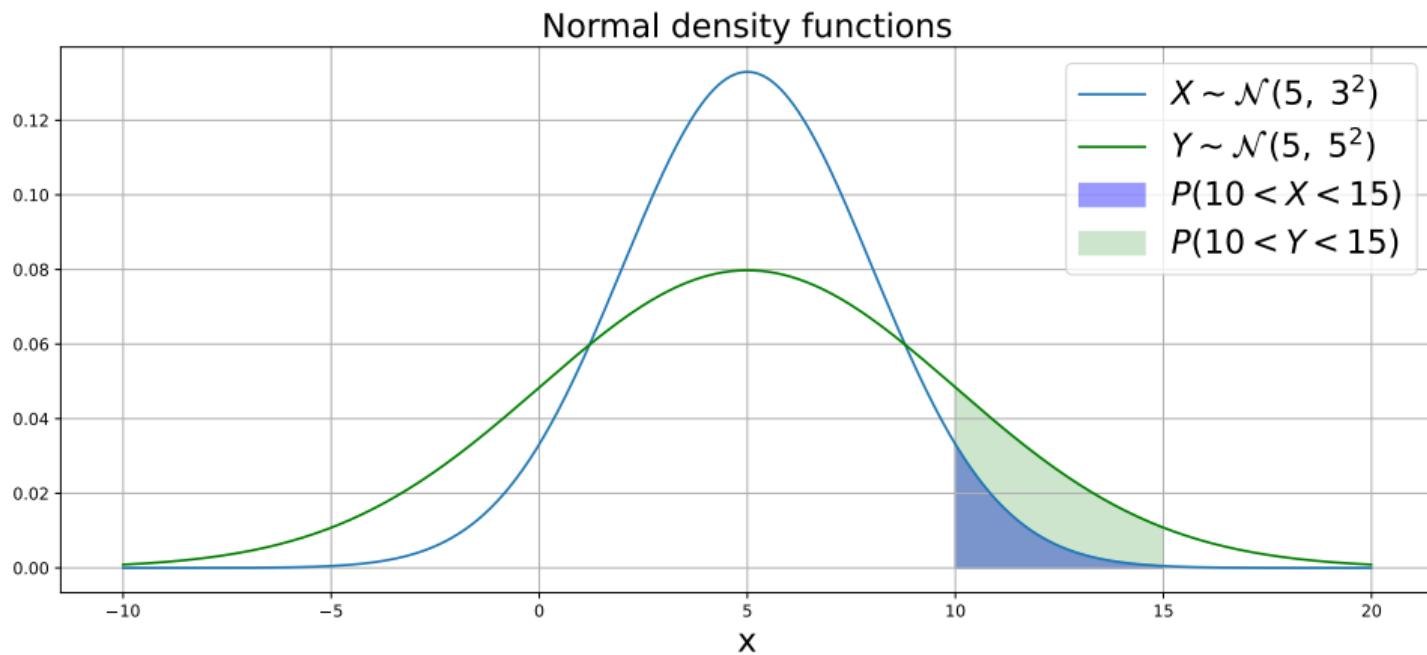


Рис. 3: Влияние параметра σ

Стандартная нормальная величина

Все нормальные равны, но одно равнее других

- Поскольку существует бесконечное число пар (μ, σ^2) , существует также бесконечное число различных нормальных распределений.
- Это потребовало бы вычисления десятков сложных вероятностных интегралов каждый раз.
- Что если вычисление вероятности **любой** нормальной СВ можно свести к знанию всего одной переменной?

Стандартная нормальная величина

Все нормальные равны, но одно равнее других

- Встречайте стандартную нормальную случайную величину:

$$Z \sim \mathcal{N}(0, 1), \quad f_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \forall x \in \mathbb{R}$$

- Давным-давно математики вычислили десятки вероятностных интегралов и создали таблицу стандартного нормального распределения, которая показывает значения функции распределения $F_Z(x)$ для многих возможных значений x .
- Мы можем найти любую вероятность любой нормальной случайной величины, преобразовав её в стандартную нормальную.

Стандартная нормальная величина

Формула преобразования

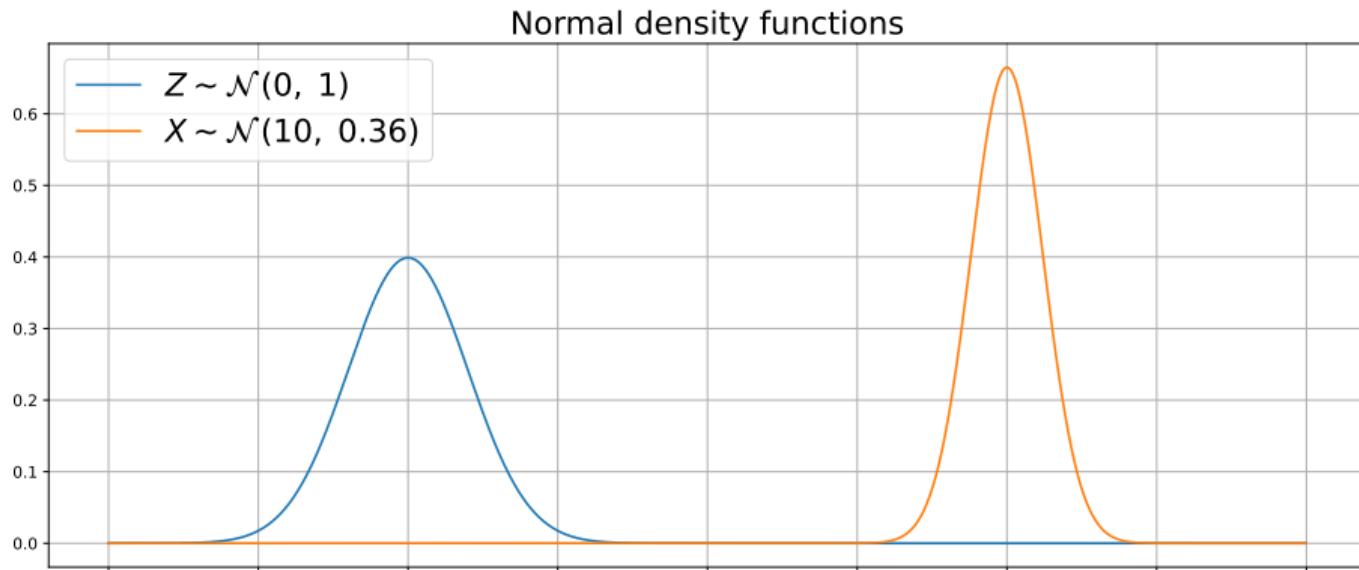
- Пусть X — нормальная случайная величина и $X \sim \mathcal{N}(\mu, \sigma^2)$.

Стандартная нормальная величина

Формула преобразования

- Пусть X — нормальная случайная величина и $X \sim \mathcal{N}(\mu, \sigma^2)$.
- Тогда следующая функция (преобразование) преобразует X в стандартную нормальную величину:

$$Z = \frac{X - \mu}{\sigma}$$



Стандартная нормальная величина

- Пусть X — нормальная случайная величина и $X \sim \mathcal{N}(\mu, \sigma^2)$.

Стандартная нормальная величина

- Пусть X — нормальная случайная величина и $X \sim \mathcal{N}(\mu, \sigma^2)$.
- Нас интересует некоторая вероятность $P(a < X < b)$, $\forall a, b : a \leq b$.

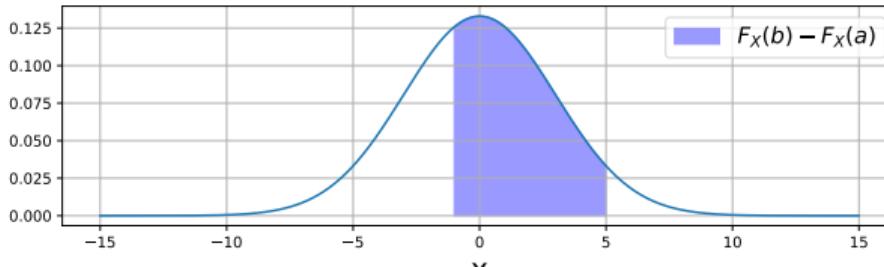
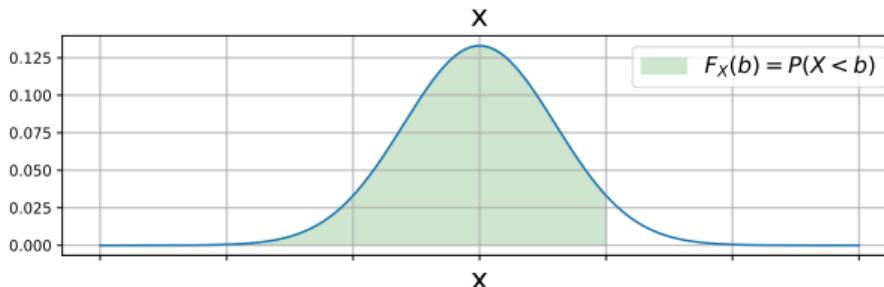
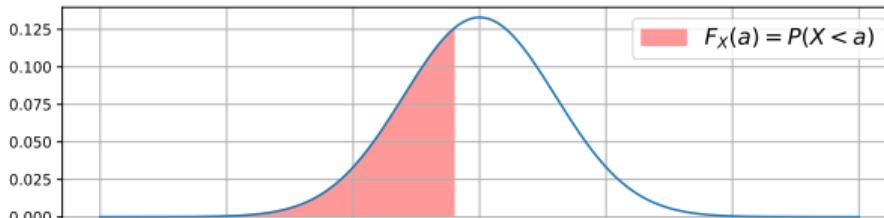
Стандартная нормальная величина

- Пусть X — нормальная случайная величина и $X \sim \mathcal{N}(\mu, \sigma^2)$.
- Нас интересует некоторая вероятность $P(a < X < b)$, $\forall a, b : a \leq b$.
- Применим преобразование к каждой части неравенства, чтобы не нарушить его:

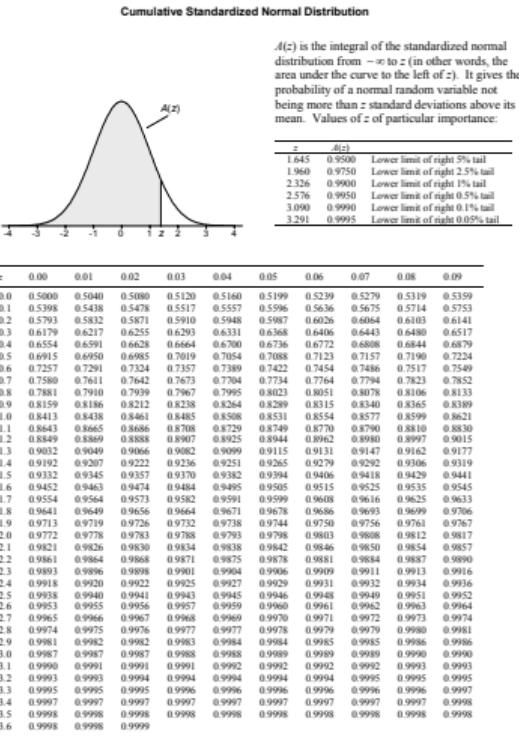
$$P(a < X < b) = P\left(\underbrace{\frac{a - \mu}{\sigma}}_{\tilde{a}} < \underbrace{\frac{X - \mu}{\sigma}}_Z < \underbrace{\frac{b - \mu}{\sigma}}_{\tilde{b}}\right) = P(\tilde{a} < Z < \tilde{b}) = F_Z(\tilde{b}) - F_Z(\tilde{a}).$$

Стандартная нормальная величина

Calculation of probability via CDF



Стандартная нормальная величина



Пример

Если X — нормально распределенная случайная величина со средним 6 и дисперсией 25, найти:

1. $P(6 \leq X \leq 12),$
2. $P(0 \leq X \leq 8),$
3. $P(-2 < X \leq 0),$
4. $P(X > 21),$
5. $P(|X - 6| < 5).$

Пример

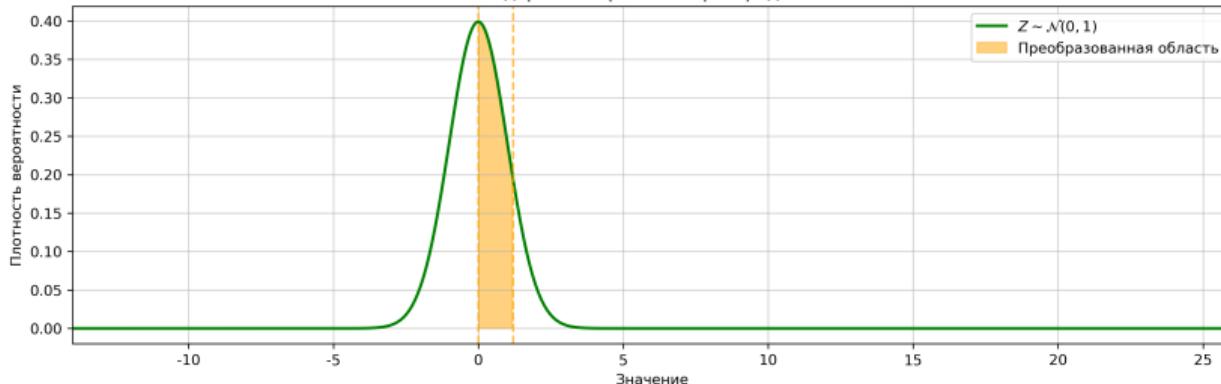
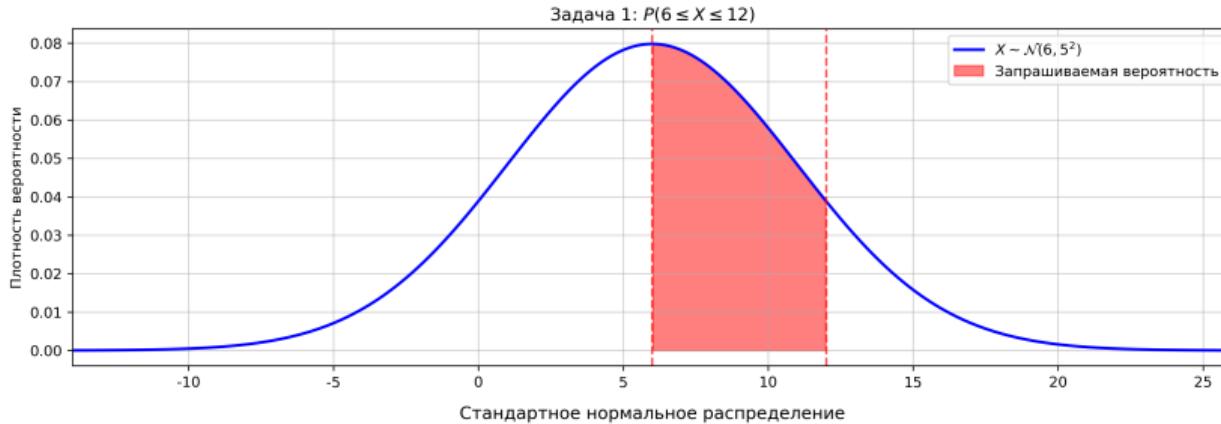
Вопрос 1

$$P(6 \leq X \leq 12)$$

- $X \sim \mathcal{N}(6, 25)$, значит $\mu = 6, \sigma = 5$
- Преобразование: $Z = \frac{X-6}{5}$
- $P(6 \leq X \leq 12) = P\left(\frac{6-6}{5} \leq Z \leq \frac{12-6}{5}\right) = P(0 \leq Z \leq 1.2)$
- $= F_Z(1.2) - F_Z(0) = 0.8849 - 0.5 = 0.3849$

Пример

Визуализация вопроса 1



Пример

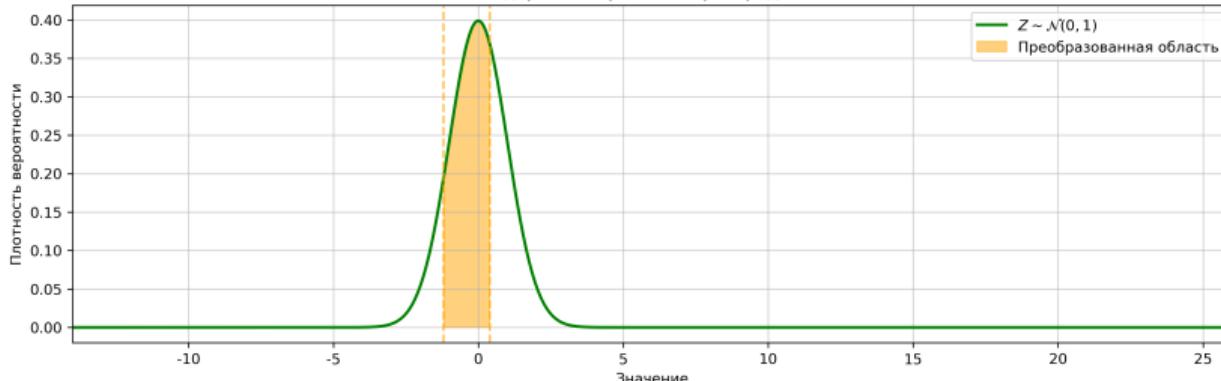
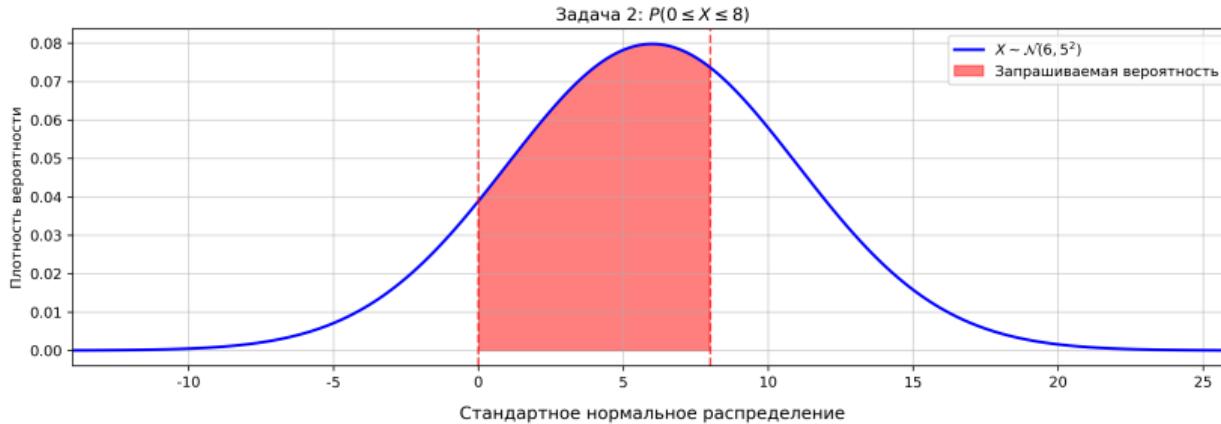
Вопрос 2

$$P(0 \leq X \leq 8)$$

- Преобразование: $Z = \frac{X-6}{5}$
- $P(0 \leq X \leq 8) = P\left(\frac{0-6}{5} \leq Z \leq \frac{8-6}{5}\right) = P(-1.2 \leq Z \leq 0.4)$
- Используем свойство симметрии: $F_Z(-1.2) = 1 - F_Z(1.2) = 1 - 0.8849 = 0.1151$
- $= F_Z(0.4) - F_Z(-1.2) = 0.6554 - 0.1151 = 0.5403$

Пример

Визуализация вопроса 2



Пример

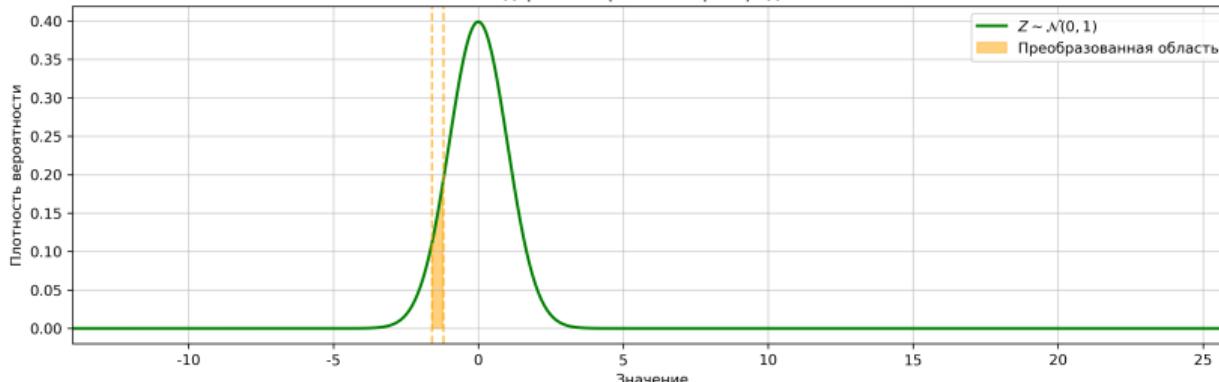
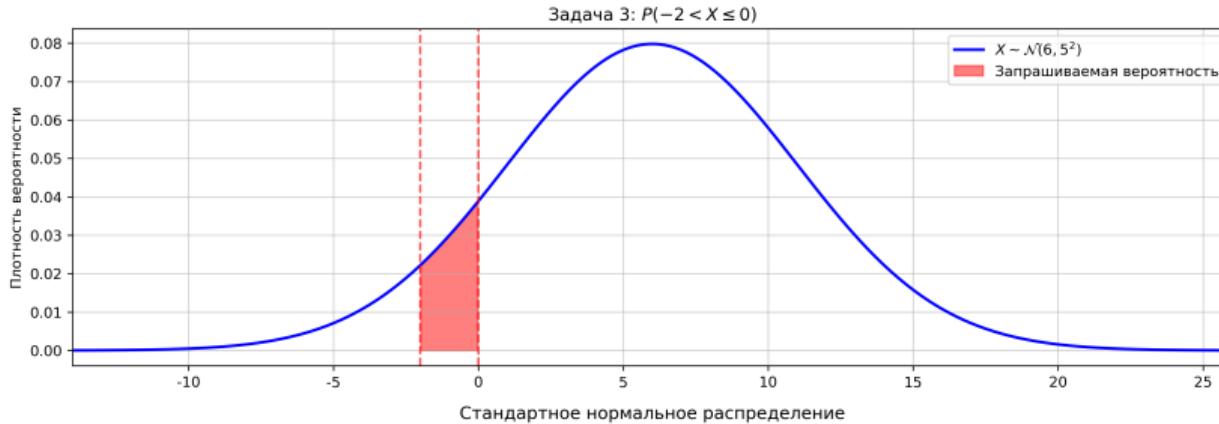
Вопрос 3

$$P(-2 < X \leq 0)$$

- Преобразование: $Z = \frac{X-6}{5}$
- $P(-2 < X \leq 0) = P\left(\frac{-2-6}{5} < Z \leq \frac{0-6}{5}\right) = P(-1.6 < Z \leq -1.2)$
- Используем свойство симметрии: $F_Z(-1.2) = 1 - F_Z(1.2) = 1 - 0.8849 = 0.1151$
- $F_Z(-1.6) = 1 - F_Z(1.6) = 1 - 0.9452 = 0.0548$
- $= F_Z(-1.2) - F_Z(-1.6) = 0.1151 - 0.0548 = 0.0603$

Пример

Визуализация вопроса 3



Пример

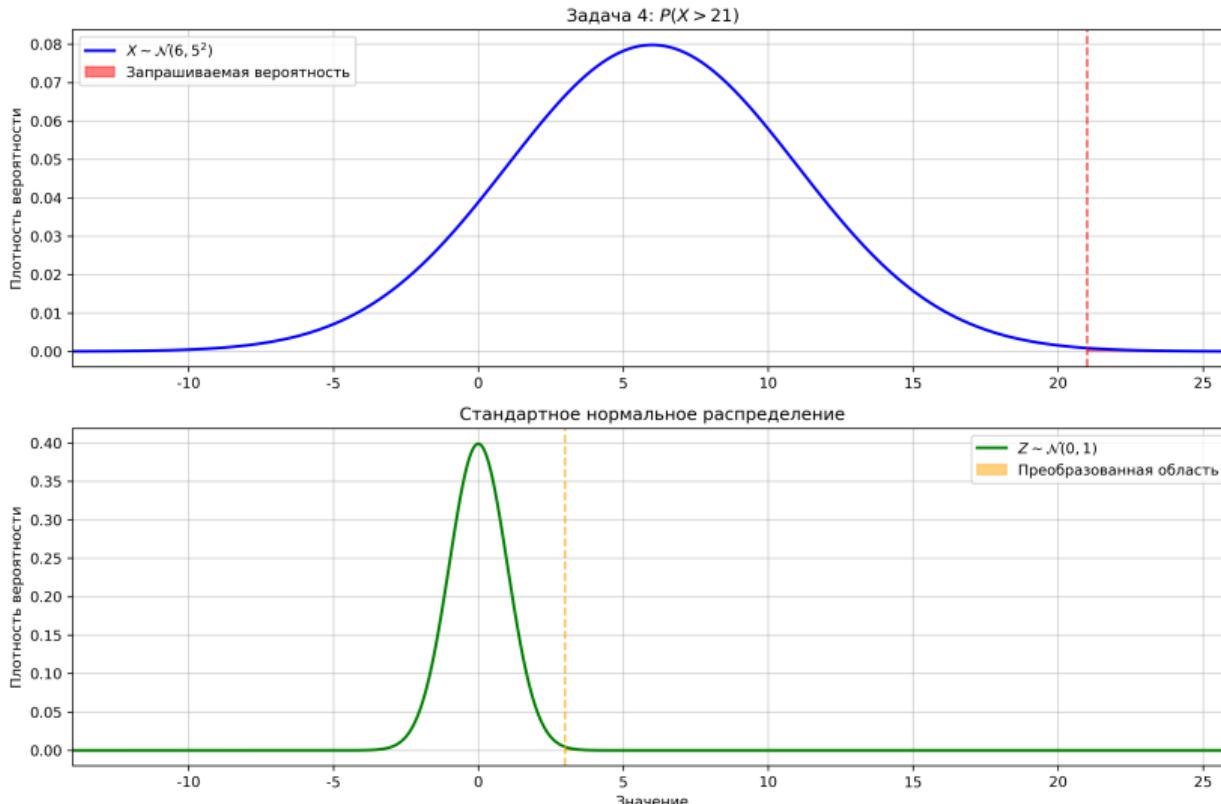
Вопрос 4

$$P(X > 21)$$

- Преобразование: $Z = \frac{X-6}{5}$
- $P(X > 21) = P\left(Z > \frac{21-6}{5}\right) = P(Z > 3)$
- $= 1 - F_Z(3) = 1 - 0.9987 = 0.0013$

Пример

Визуализация вопроса 4



Пример

Вопрос 5

$$P(|X - 6| < 5)$$

- Это означает: $P(-5 < X - 6 < 5) = P(1 < X < 11)$
- Преобразование: $Z = \frac{X-6}{5}$
- $P(1 < X < 11) = P\left(\frac{1-6}{5} < Z < \frac{11-6}{5}\right) = P(-1 < Z < 1)$
- Используем свойство симметрии: $F_Z(-1) = 1 - F_Z(1) = 1 - 0.8413 = 0.1587$
- $= F_Z(1) - F_Z(-1) = 0.8413 - 0.1587 = 0.6826$

Пример

Визуализация вопроса 5

Задача 5: $P(|X - 6| < 5) = P(1 < X < 11)$



Стандартное нормальное распределение

