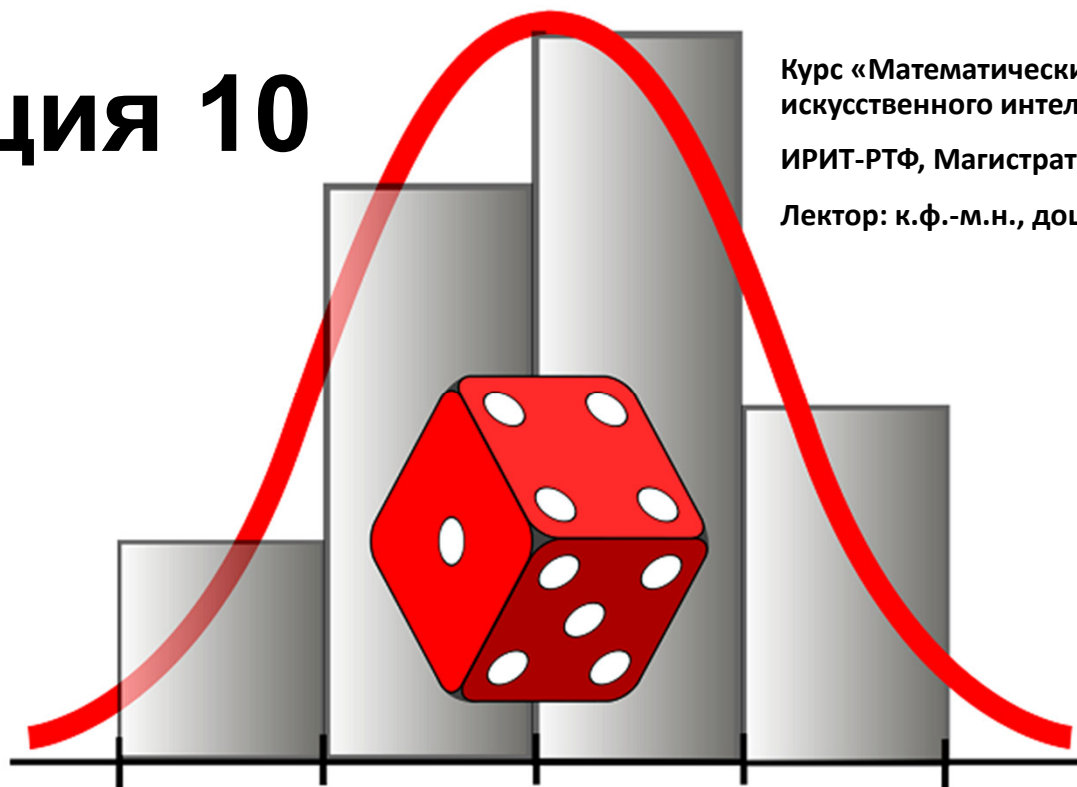


Лекция 10

Курс «Математические основы
искусственного интеллекта»

ИРИТ-РТФ, Магистратура (СпецВО)

Лектор: к.ф.-м.н., доцент Нагребецкая Ю.В.



Математические основы искусственного интеллекта Теория вероятностей

Лекция 10

Непрерывные случайные величины

1. Равномерное распределение.
2. Нормальное распределение.

Понятие равномерного распределения

Задача 3

Задача 3. Пусть трамваи по некоторому маршруту в данное время ходят с интервалом 10 минут. Человек приходит в некоторый момент времени на остановку и видит, что трамвая нет. Найти функцию распределения $F(x)$ и плотность распределения $f(x)$ случайной величины X — времени ожидания трамвая. Найти среднее время ожидания трамвая и вероятность того, что время ожидания трамвая будет более 5 минут.



Курс "Математические основы искусственного интеллекта", автор к.ф.-м.н., доцент ИЕиМ УрФУ
Нагребцкая Ю.В.

Задача 3. Решение

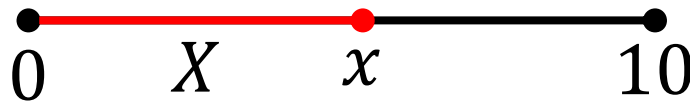
Решение. Найдем функцию распределения $F(x) = P(X < x)$.

Если $x \leq 0$, то, очевидно, $F(x) = 0$.

Если $x > 10$, то, $F(x) = 1$.

Если $0 < x \leq 10$, то, по опр. геометрической вероятности

$$F(x) = \frac{\text{длина отрезка } [0;x]}{\text{длина отрезка } [0;10]} = \frac{x}{10} \Rightarrow$$



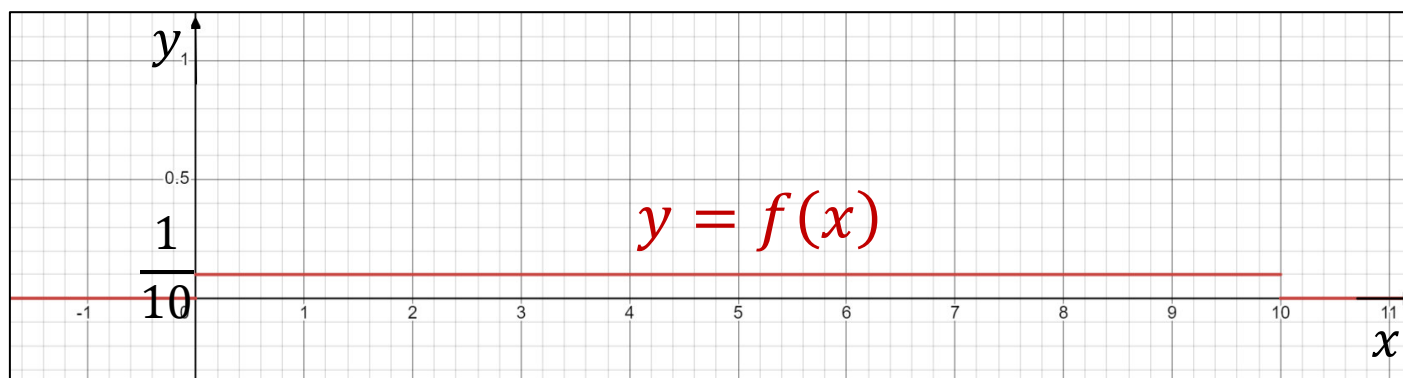
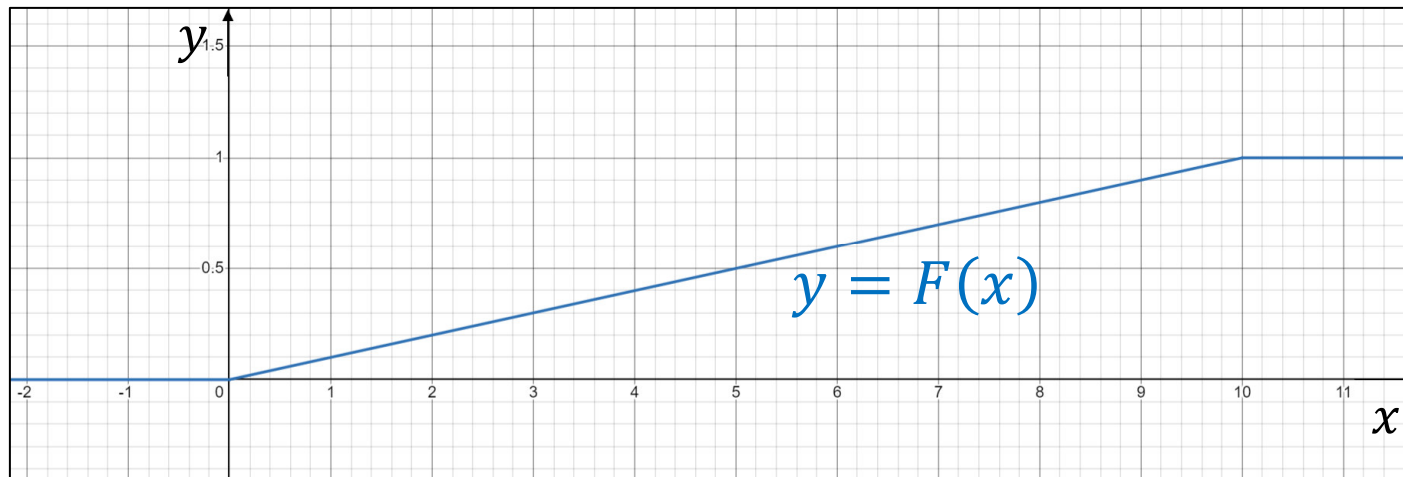
Задача 3. Решение (продолжение)

$$F(x) = \begin{cases} 0 & \text{при } x \leq 0, \\ \frac{x}{10} & \text{при } 0 < x \leq 10, \\ 1 & \text{при } x > 10, \end{cases}$$

$$f(x) = \begin{cases} 0 & \text{при } x \leq 0, \\ \frac{1}{10} & \text{при } 0 < x \leq 10, \\ 0 & \text{при } x > 10 \end{cases}$$

Такое распределение называется **равномерным**.

Задача 3. Решение (иллюстрация)



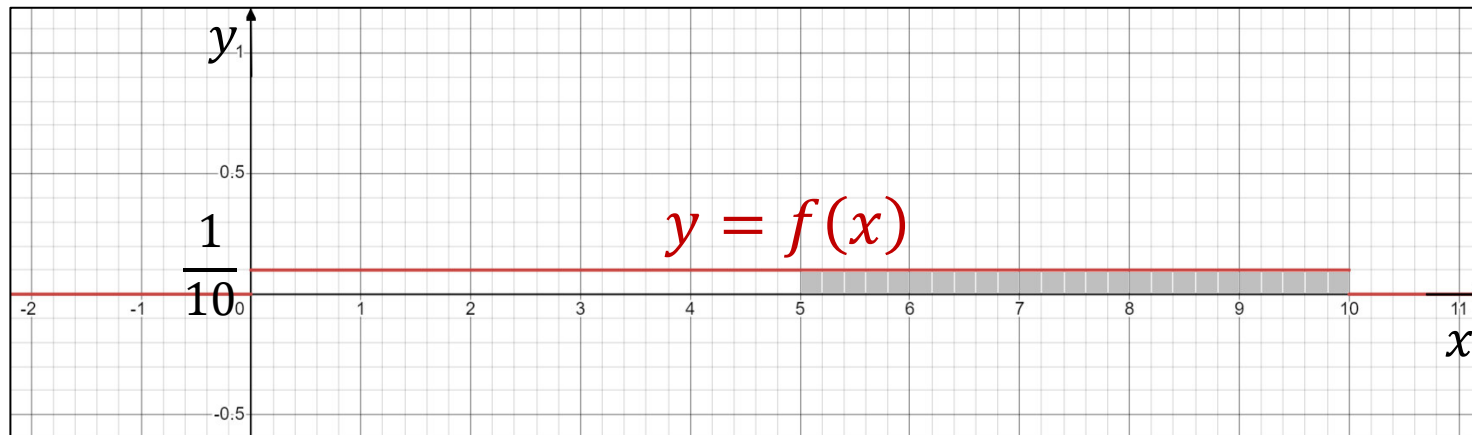
Задача 3. Решение (продолжение)

Тогда искомая вероятность равна

$$P(X > 5) = P(5 < X < +\infty) =$$
$$(I \text{ сп.}) = F(+\infty) - F(5) = 1 - \frac{5}{10} = 0.5$$

$$(II \text{ сп.}) = \int_5^{+\infty} f(x) dx = \int_5^{10} \frac{1}{10} dx = \left. \frac{x}{10} \right|_5^{10} =$$
$$= \frac{10}{10} - \frac{5}{10} = 0.5$$

Задача 3. Решение (иллюстрация)



Задача 3. Решение (продолжение)

Среднее значение - это матем. ожидание:

$$\begin{aligned} M(X) &= \int_{-\infty}^{+\infty} xf(x)dx = \\ &= \int_{-\infty}^0 xf(x)dx + \int_0^{10} xf(x)dx + \int_{10}^{+\infty} xf(x)dx = \\ &= \int_{-\infty}^0 0dx + \int_0^{10} x \frac{1}{10} dx + \int_{10}^{+\infty} 0dx = \\ &= 0 + \frac{1}{10} \int_0^{10} xdx + 0 = \end{aligned}$$

Задача 2. Решение (продолжение)

$$= \frac{1}{10} \int_0^{10} x dx = \frac{1}{10} \frac{x^2}{2} \Big|_0^{10} = 5$$

Найдем $D(X)$ по (4) теоремы 5.

Для этого найдем

$$\begin{aligned} M(X^2) &= \int_{-\infty}^{+\infty} x^2 f(x) dx = \\ &= \int_{-\infty}^0 x^2 f(x) dx + \\ &+ \int_0^{10} x^2 f(x) dx + \int_{10}^{+\infty} x^2 f(x) dx = \end{aligned}$$

Решение (окончание)

$$= \int_{-\infty}^0 0 dx + \int_0^{10} x^2 \frac{1}{10} dx + \int_{10}^{+\infty} 0 =$$

$$= \int_0^{10} x^2 \frac{1}{10} dx = \frac{1}{10} \frac{x^3}{3} \Big|_0^{10} = \frac{100}{3} \Rightarrow$$

$$D(X) = M(X^2) - (M(X))^2 = \frac{100}{3} - 5^2 = \frac{25}{3}$$

$$\sigma(X) = \sqrt{D(X)} = \sqrt{\frac{25}{3}} = \frac{5}{\sqrt{3}} \approx 2.89$$

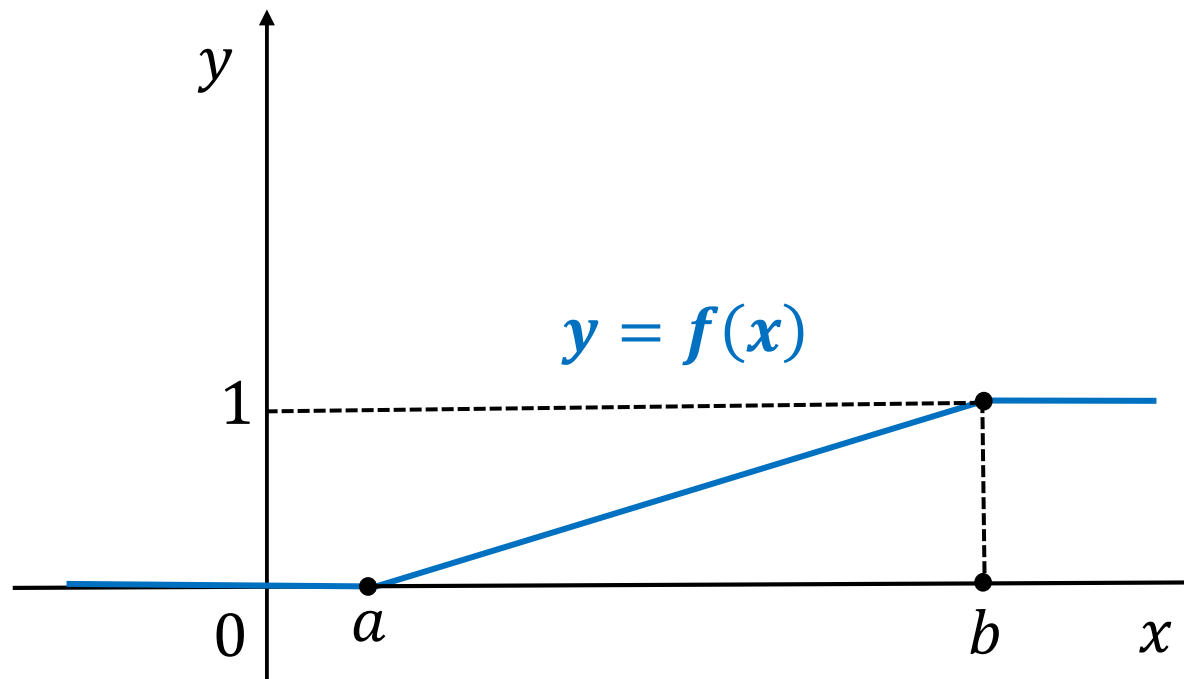
Равномерное распределение НСВ в общем случае

В общем случае **равномерное распределение** на отрезке $[a; b]$ непрерывной СВ X задается следующей функцией распределения

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

где $a < b$.

Равномерное распределение НСВ в общем случае. График $y = F(x)$



Равномерное распределение НСВ в общем случае

Тогда, очевидно,

$$f(x) = F'(x) = \begin{cases} 0 & \text{при } x < a, \\ \frac{1}{b-a} & \text{при } a < x < b, \\ 0 & \text{при } x > b, \end{cases}$$

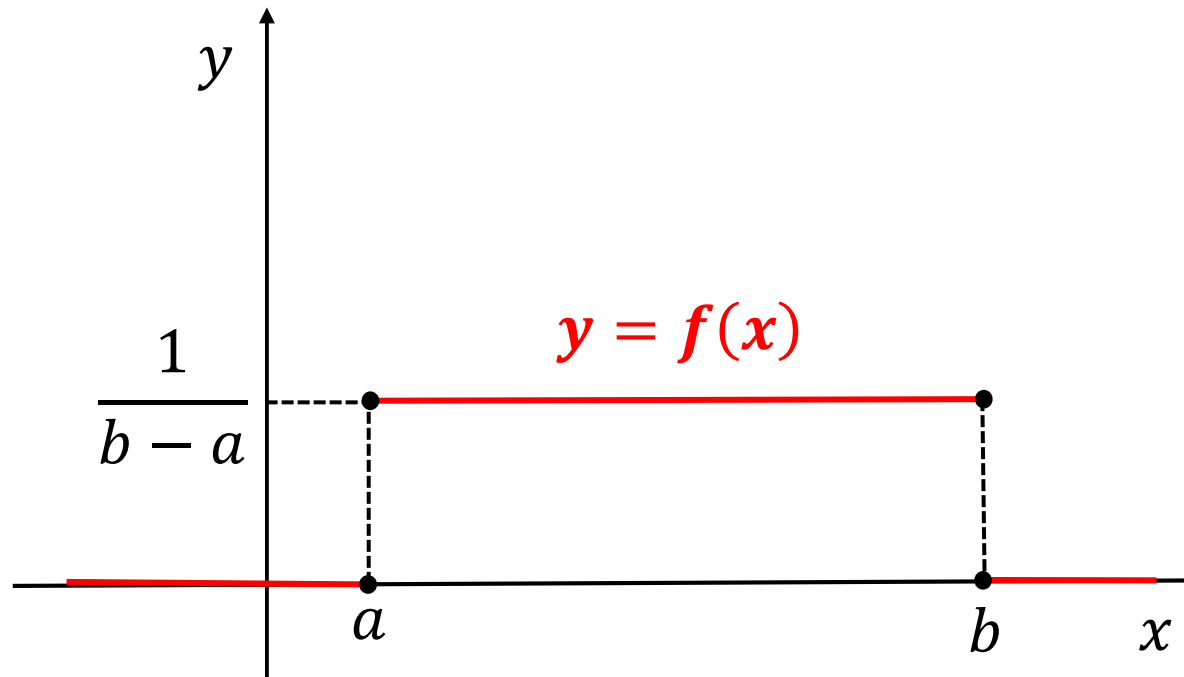
Упр. Доказать, что

$$M(X) = \frac{a+b}{2}$$

$$M(X^2) = \frac{b^2 + ab + a^2}{3}$$

$$D(X) = \frac{(b-a)^2}{12}$$

Равномерное распределение НСВ в общем случае. График $y = f(x)$



Равномерное распределение. Упр. 4

Упр.4. Школьник меряет длину учебника при помощи линейки, округляя до ближайшего целого. Какова вероятность, что точность его измерений меньше или равна 0,1 см.

Указание: Использовать равномерно распределенную на отрезке $\left[-\frac{1}{2}; \frac{1}{2}\right]$ случайную величину X — точность (абсолютную погрешность) его измерений.

Ответ: 0,2



Нормальное распределение НСВ

Самое распространенное в практических приложениях является нормальное распределение.

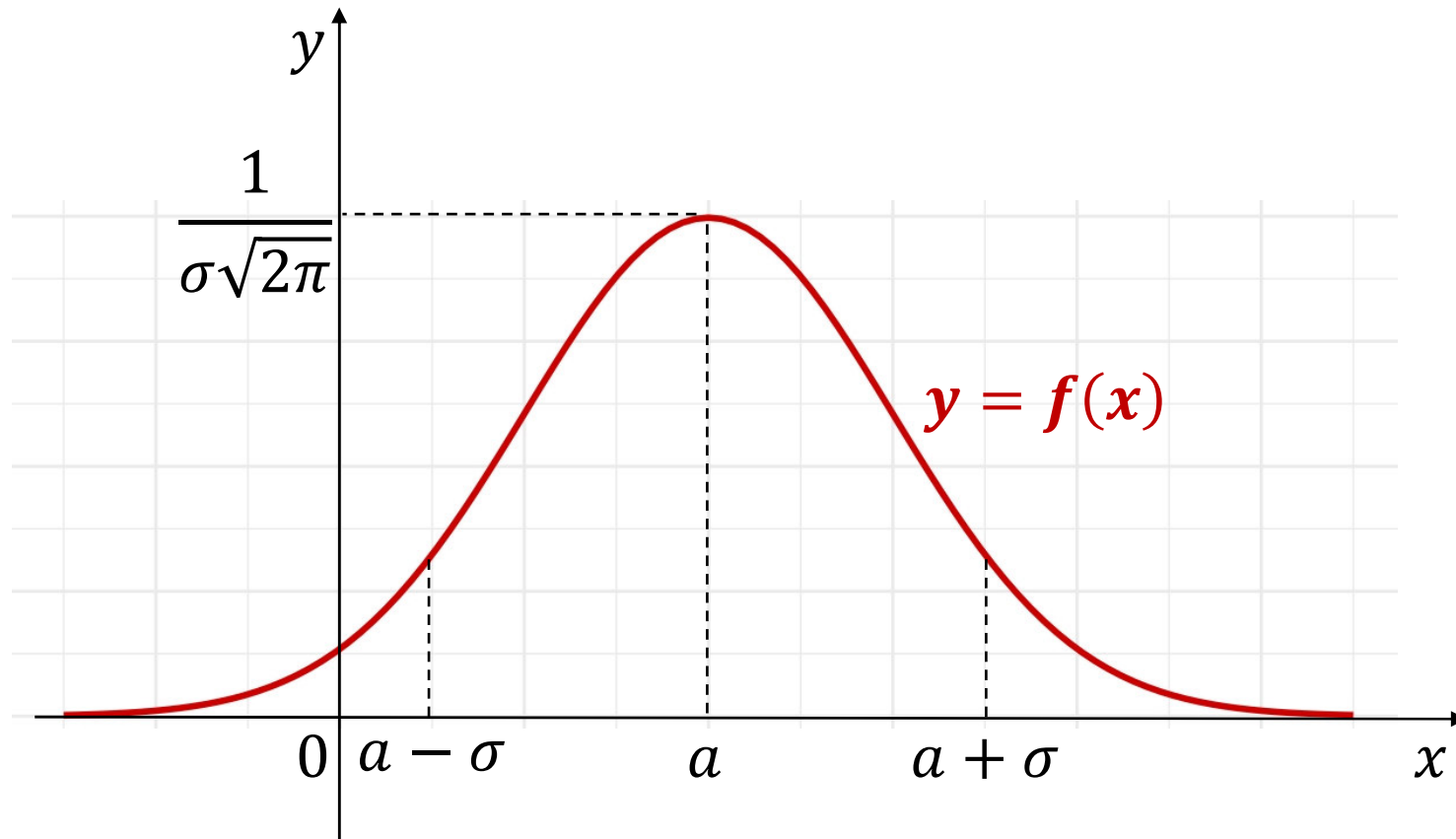
Опр. Нормальное распределение (распределение Гаусса) НСВ – это распределение с плотностью

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

При этом заметим, что $\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-a)^2}{2\sigma^2}} dt = 1$

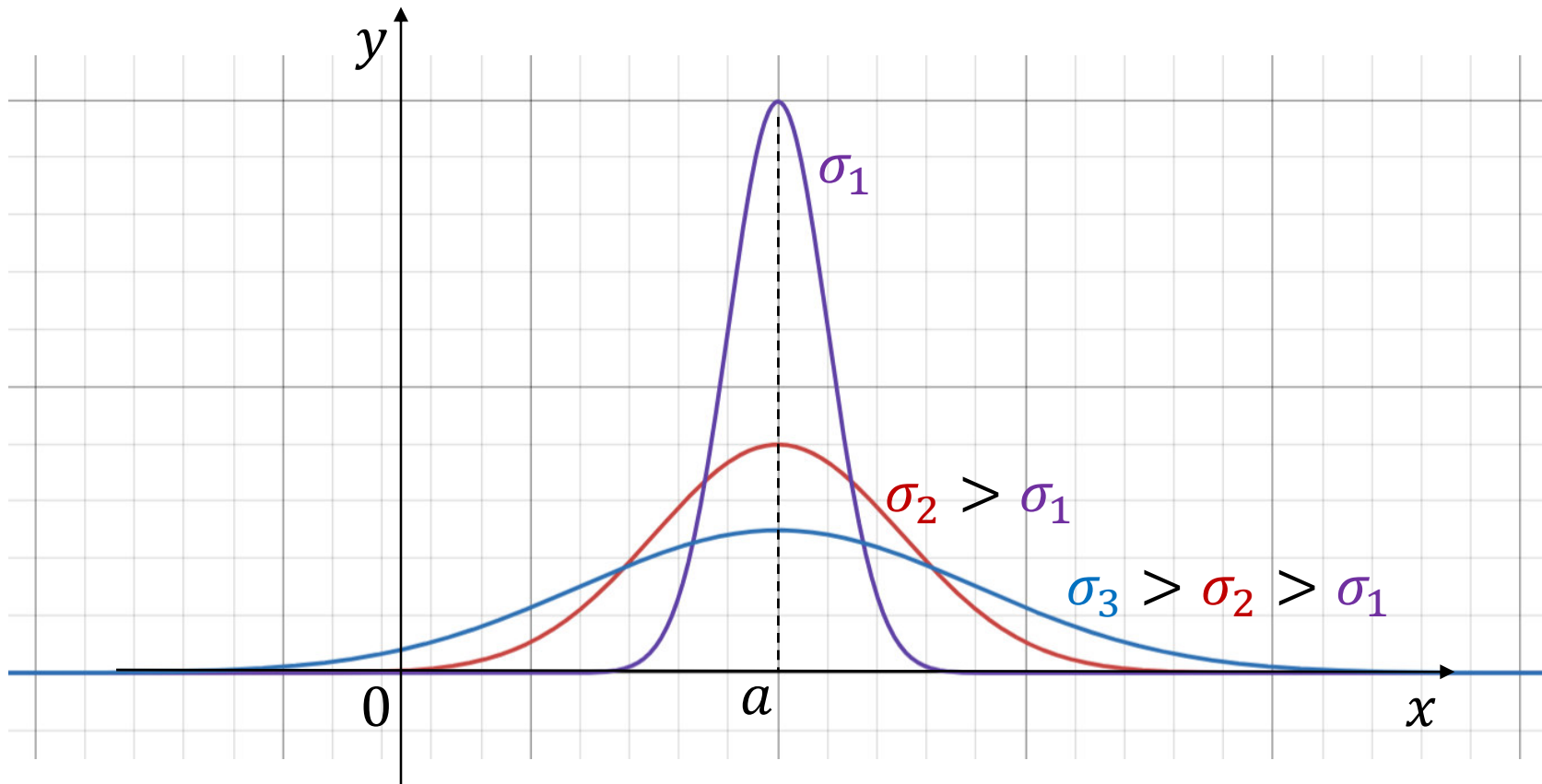
Нормальное распределение

График $y = f(x)$



Нормальное распределение

График $y = f(x)$ для разных сигм



Нормальное распределение НСВ

Функция распределения

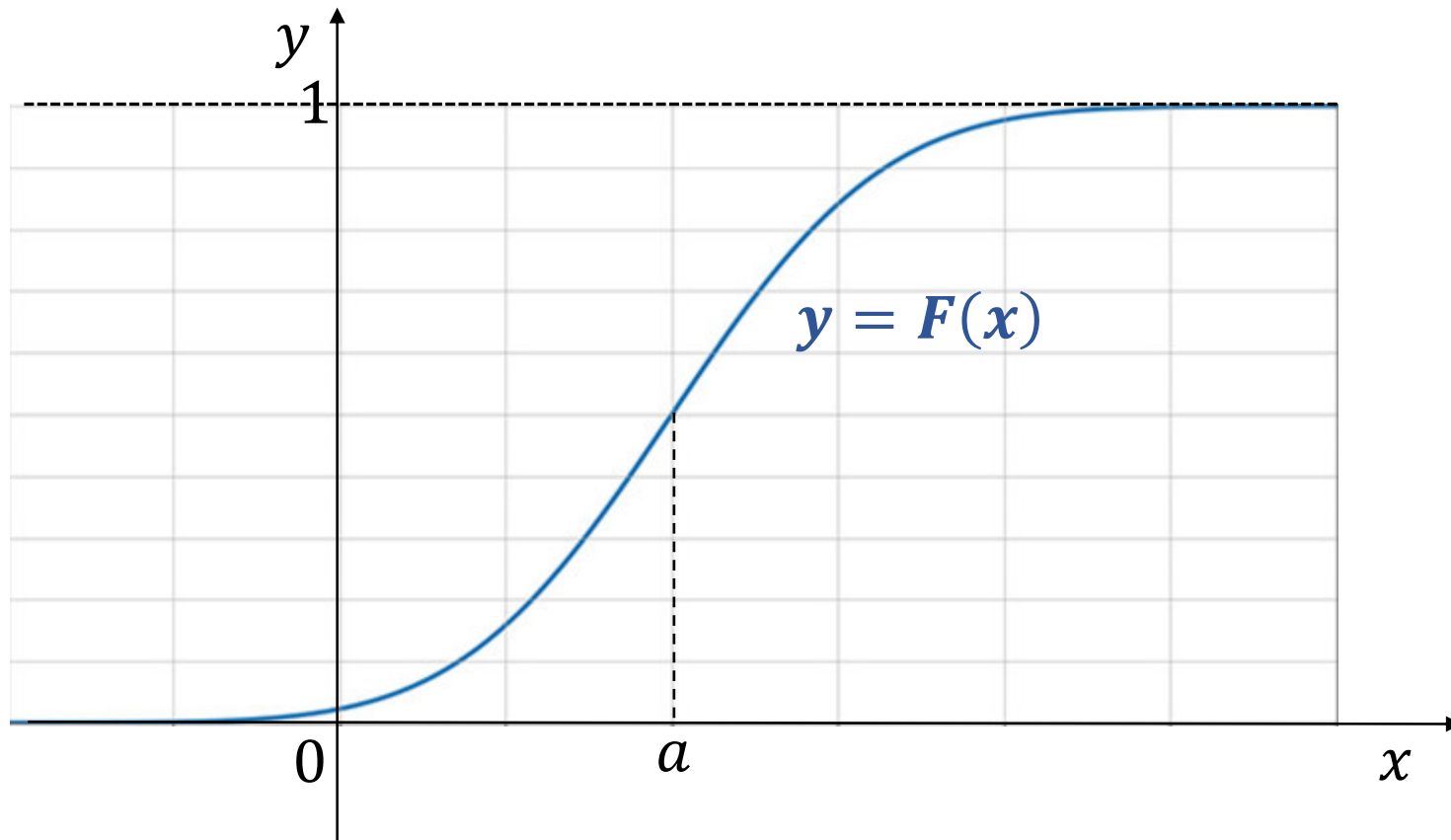
Функция распределения нормального распределения по (2) теоремы 3 (часть I лекции) тогда равна

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-a)^2}{2\sigma^2}} dt$$

Обозначение для нормального распределения:
 $N(a, \sigma)$.

Нормальное распределение НСВ

График $y = F(x)$



Стандартное нормальное распределение

Опр. Если $a = 0, \sigma = 1$, то нормальное распределение называется **стандартным**.

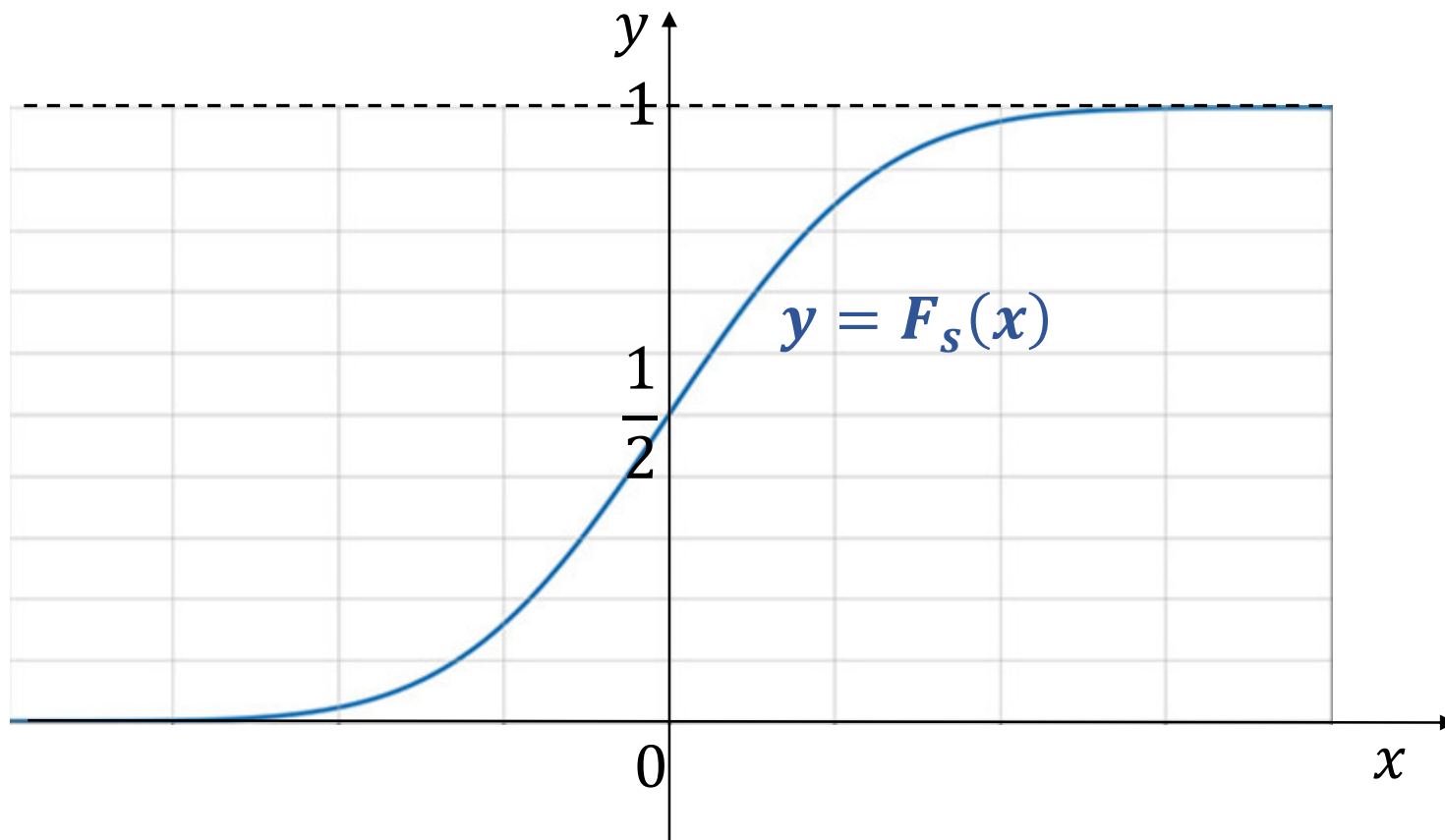
Тогда функция плотности и функция распределения соответственно равны

$$f_s(x) = \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$F_s(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

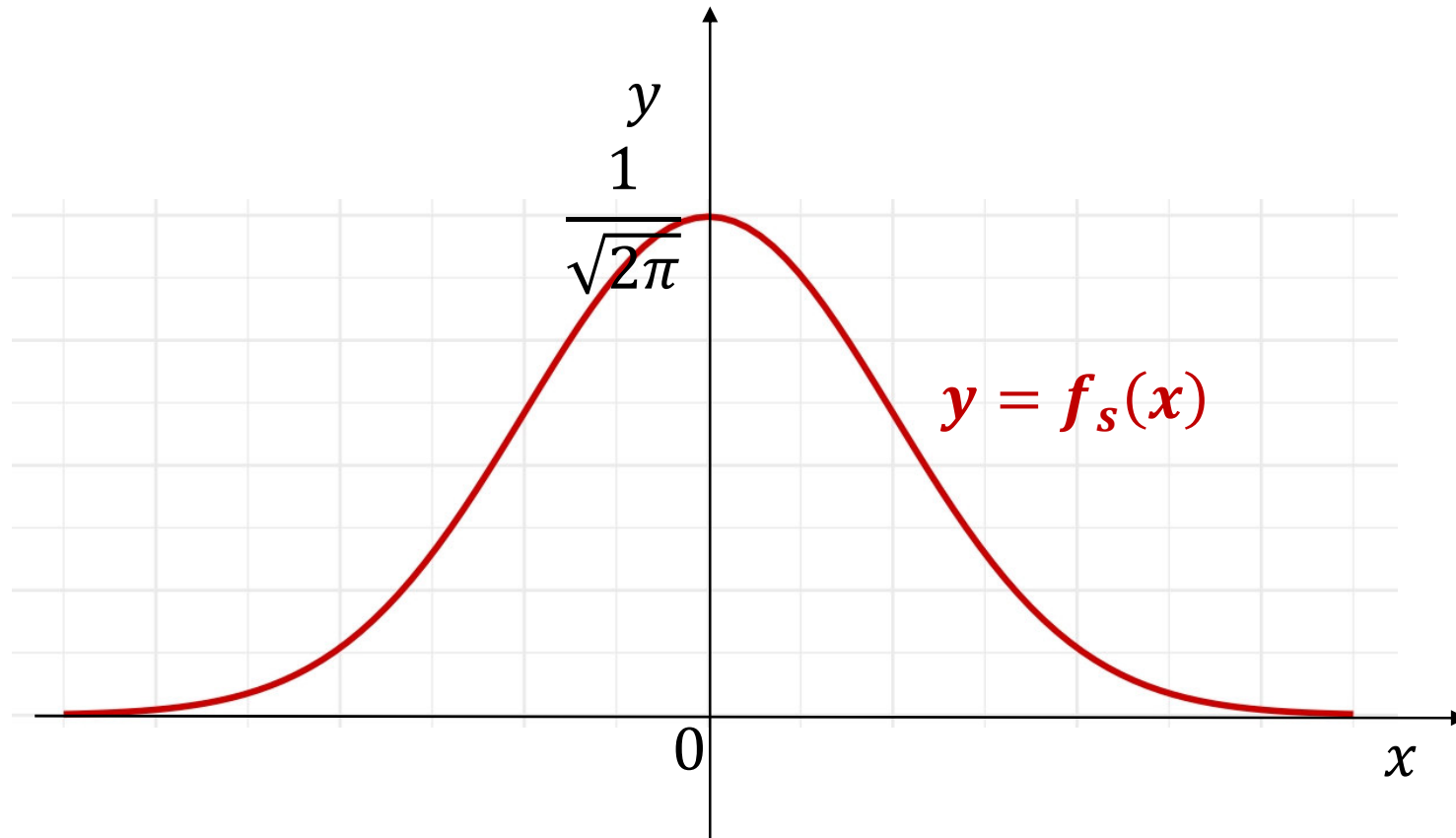
Стандартное нормальное распределение

График $y = F_s(x)$



Стандартное нормальное распределение

График $y = f_s(x)$



Стандартное нормальное распределение

Из условия нормальности

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-\frac{t^2}{2}} dt = \frac{1}{2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = \frac{1}{2} \cdot 1 = \frac{1}{2}$$

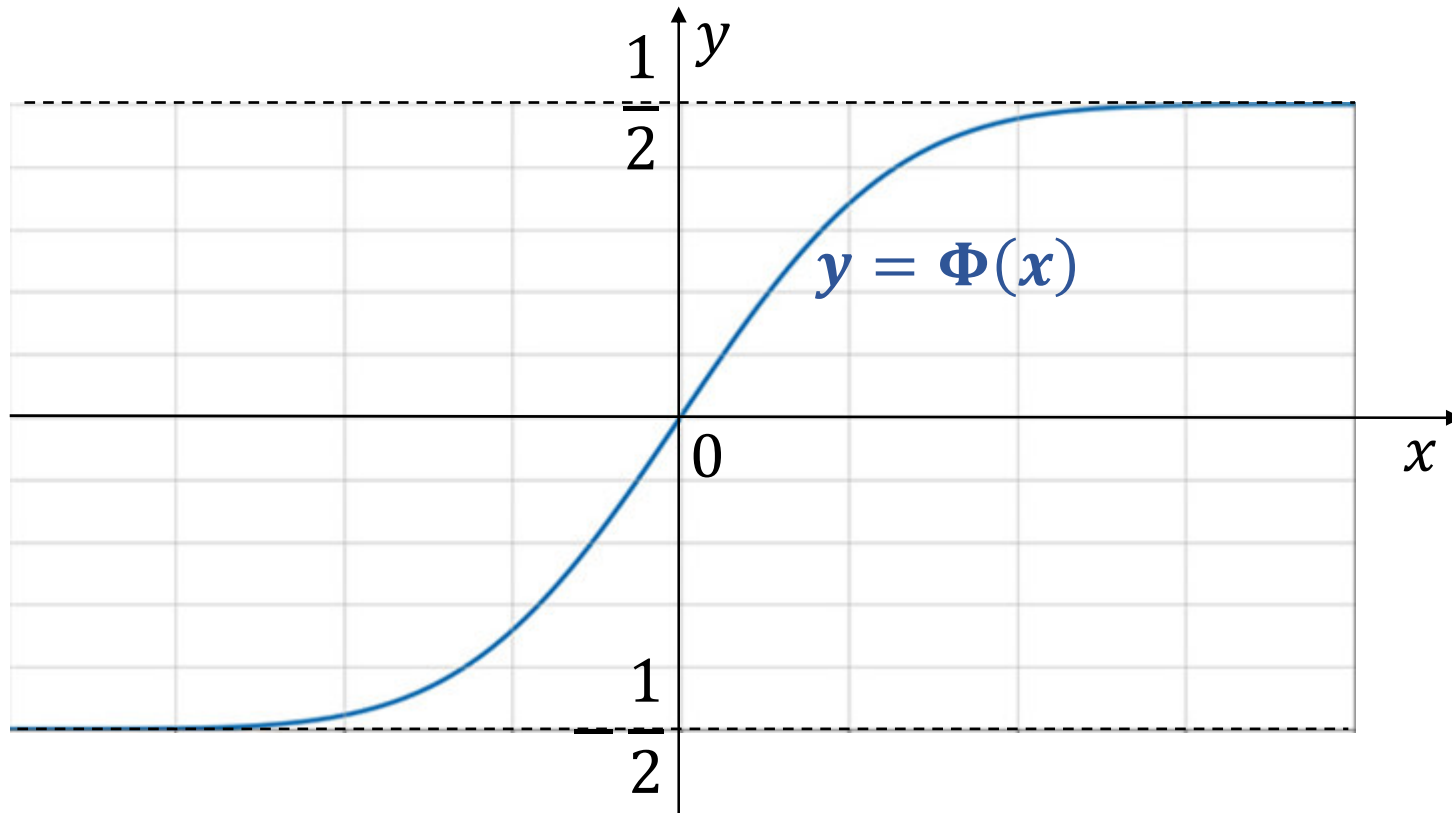
$$\Rightarrow F_s(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-\frac{t^2}{2}} dt + \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt =$$

$$= \frac{1}{2} + \Phi(x), \text{ где } \Phi(x) - \text{функция Лапласа}$$

Стандартное нормальное распределение

График $y = \Phi(x)$



Теорема о числовых характеристиках нормального распределения

Теорема 6. Пусть НСВ X распределено нормально с параметрами a и σ .

Тогда $M(X) = a, D(X) = \sigma^2, \sigma(X) = \sigma$

Док-во равенства $M(X) = a$ для простоты проведем для $N(0; 1)$. Остальные - без док-ва.

$$\begin{aligned} M(X) &= \int_{-\infty}^{+\infty} x f(x) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x e^{-\frac{x^2}{2}} dx = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \text{нечетная функция} dx = 0 = a \end{aligned}$$

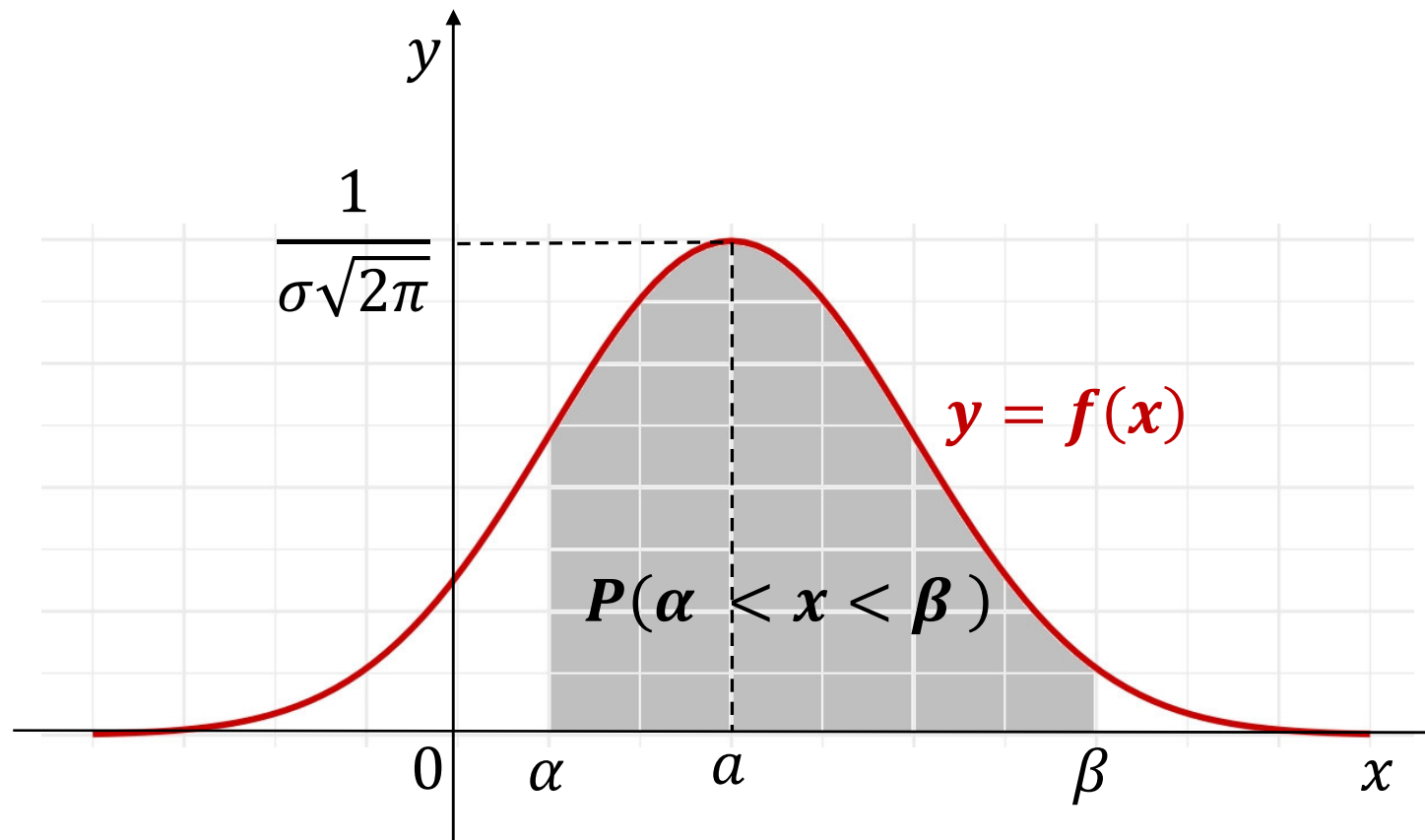
Нормальное распределение. Теорема о вычислении вероятности

Теорема 7. Пусть НСВ X имеет нормальное распределение $N(a; \sigma)$.

Тогда

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right)$$

Нормальное распределение. Теорема о вычислении вероятности Геометрическая иллюстрация



Теорема о вычислении вероятностей

Доказательство

Доказательство. По (3) теоремы 3

$$\begin{aligned} P(\alpha < X < \beta) &= \int_{\alpha}^{\beta} f(x) dx = \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \left[\begin{array}{l} t = \frac{x-a}{\sigma}, dt = \frac{dx}{\sigma} \\ dx = \sigma dt \\ x = \alpha \rightarrow t = \frac{\alpha-a}{\sigma} \\ x = \beta \rightarrow t = \frac{\beta-a}{\sigma} \end{array} \right] = \end{aligned}$$

Теорема о вычислении вероятностей

Доказательство

$$\begin{aligned} &= \frac{1}{\sqrt{2\pi}} \int_{\frac{\alpha-a}{\sigma}}^{\frac{\beta-a}{\sigma}} e^{-\frac{t^2}{2}} \sigma dt = \int_{\frac{\alpha-a}{\sigma}}^{\frac{\beta-a}{\sigma}} f_s(t) dt = \\ &= F_s \left(\frac{\beta-a}{\sigma} \right) - F_s \left(\frac{\alpha-a}{\sigma} \right) = \\ &= \left(\Phi \left(\frac{\beta-a}{\sigma} \right) + \frac{1}{2} \right) - \left(\Phi \left(\frac{\alpha-a}{\sigma} \right) + \frac{1}{2} \right) = \\ &= \Phi \left(\frac{\beta-a}{\sigma} \right) - \Phi \left(\frac{\alpha-a}{\sigma} \right). \blacksquare \end{aligned}$$

Следствие 1 из теоремы о вычислении вероятностей

Следствие 1. Пусть НСВ X имеет нормальное распределение.

Тогда

$$P(|X - a| < \varepsilon) = 2\Phi\left(\frac{\varepsilon}{\sigma}\right)$$

Доказательство.

$$\begin{aligned} P(|X - a| < \varepsilon) &= P(a - \varepsilon < X < a + \varepsilon) = \\ &= \Phi\left(\frac{\varepsilon}{\sigma}\right) - \Phi\left(-\frac{\varepsilon}{\sigma}\right) = \Phi\left(\frac{\varepsilon}{\sigma}\right) + \Phi\left(\frac{\varepsilon}{\sigma}\right) = 2\Phi\left(\frac{\varepsilon}{\sigma}\right). \blacksquare \end{aligned}$$

Правило трех сигм

Следствие 2 (правило трех сигм). Пусть НСВ X имеет нормальное распределение $N(a; \sigma)$.

Тогда отклонение X от a более, чем на 3σ маловероятно (практически невозможно), т.е.

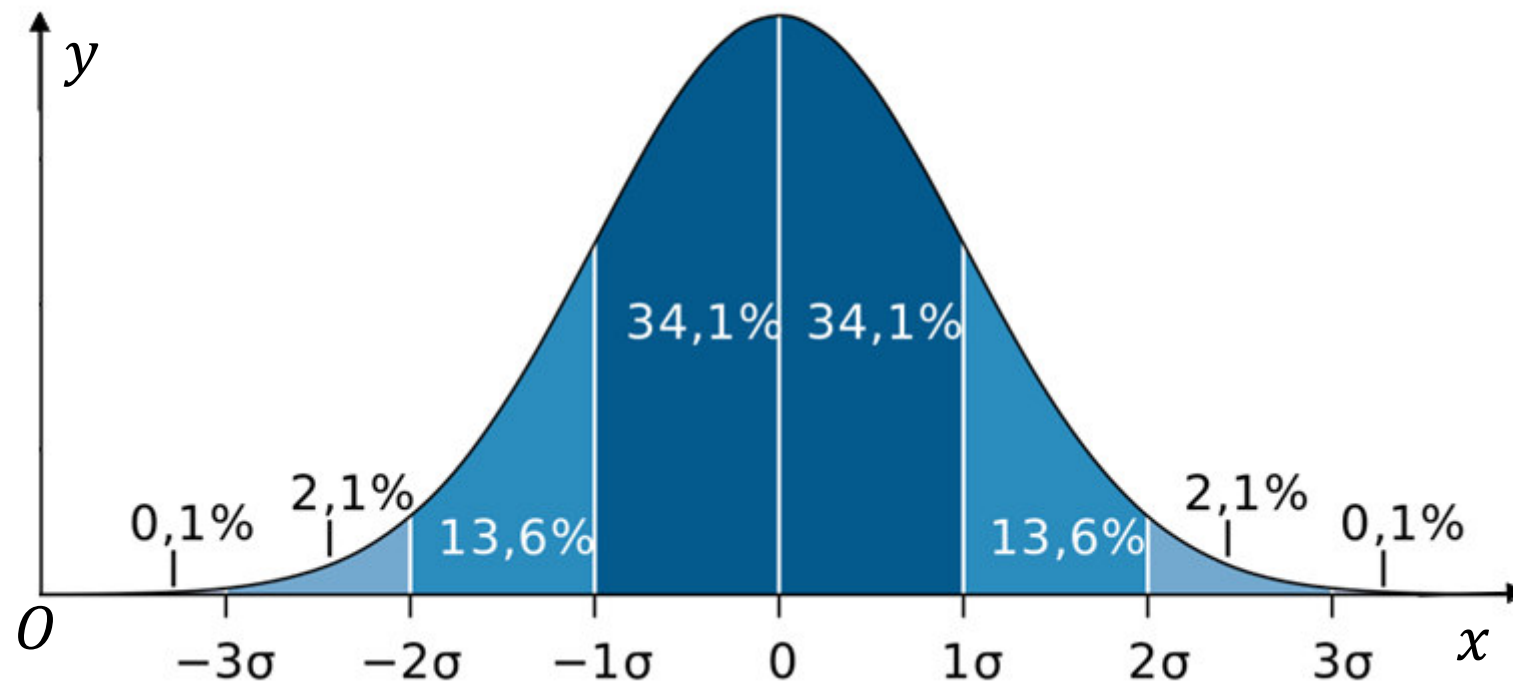
$$P(|X - a| \geq 3\sigma) \approx 0$$

Доказательство.

$$P(|X - a| < 3\sigma) = 2\Phi\left(\frac{3\sigma}{\sigma}\right) = 2\Phi(3) \approx 0.9973$$

$$\begin{aligned} \Rightarrow P(|X - a| \geq 3\sigma) &= \\ &= 1 - P(|X - a| < 3\sigma) \leq 0.0027 \blacksquare \end{aligned}$$

Правило трех сигм. Геометрическая иллюстрация



Нормальное распределение. Задача 4

Коробки с шоколадом упаковываются автоматически; их средняя масса равна 1.06 кг. Найти вероятность того, что коробка имеет массу меньше 1 кг, если масса коробок распределена по нормальному закону со средним квадратичным отклонением, равным 0,03 кг.



Нормальное распределение

Задача 4. Решение

Решение.

X – масса коробок - имеет нормальное распределение.

$$\mu = 1.06; \sigma = 0.03$$

$$P(X < 1) = ?$$

По теореме 7

$$\begin{aligned} P(X < 1) &= P(-\infty < X < 1) = \\ &= \Phi\left(\frac{1 - 1.06}{0.03}\right) - \Phi\left(\frac{-\infty - 1.06}{0.03}\right) = \end{aligned}$$

Нормальное распределение

Задача 4. Решение (продолжение)

$$\begin{aligned}\Phi\left(\frac{-0.06}{0.03}\right) - \Phi(-\infty) &= -\Phi(2) + \Phi(+\infty) = \\ &= [\text{по свойствам } \Phi(x)] = \\ &= -\Phi(2) + 0.5 \approx -0.48 + 0.5 = 0.02\end{aligned}$$

Нормальное распределение. Упр 5.

Упр.5. Автомат изготавливает шарики. Шарик считается годным, если отклонение X его диаметра от проектного размера по абсолютной величине меньше 0.7 мм. Считая, что случайная величина X нормально распределена с $a = 0$ (случайная ошибка) и $\sigma = 0.4$ мм, найти, сколько в среднем будет годных шариков из 100 случайно выбранных?

Указание: найти $100 \cdot P$, где $P = P(|X - a| < \varepsilon)$.

Ответ: ≈ 92 шт.

