20.10.16
Занятие
4

2.4. Элементы комбинаторики

Перечислительная комбинаторика рассматривает задачи подсчёта количества различных конфигураций (например, перестановок, размещений, сочетаний и т.д.).

Это нужно для расчета m и n в формуле априорной вероятности.

2.4.1. Перестановки

Перестановкой из n элементов называется всякий упорядоченный набор этих n элементов.

Обозначим число перестановок n элементов $П_n$.

Пример:

1) 123
2) 132
3) 213
4) 231
5) 312
6) 321

$П_3 = 6$.

Теорема 1. $П_n = n!$

2.4.2. Сочетания

Пусть $k \le n$.

Сочетанием из n элементов по k называется всякий неупорядоченный набор k элементов, выбранных из n данных элементов.

Замечание: 2 сочетания являются одинаковыми, если имеют одинаковый состав элементов. При этом они могут иметь разный порядок этих элементов. Например, 12 = 21.

Обозначим число сочетаний из n элементов по k: C_n^k.

Пример: из элементов 1,2,3 берём сочетания по два элемента:

1) 12
2) 23
3) 13

$C_3^2 = 3$.

Теорема 2. $C_n^k = \dfrac{n!}{k!(n-k)!}$

Частные случаи сочетаний:

$$C_n^0 = \frac{n!}{n!} = 1$$

$$C_n^1 = \frac{n!}{(n-1)!} = n$$

$$C_n^n = \frac{n!}{n!} = 1$$

2.4.3. Размещения

Пусть $k \leq n$.

Размещением из n элементов по k называется всякий упорядоченный набор k элементов, выбранных из n данных элементов.

Замечание: 2 размещения являются разными, если имеют не только разный состав элементов, но и разный порядок этих элементов. Например, $12 \neq 21$.

Обозначим число размещений из n элементов по k: A_n^k.

Пример: из элементов 1,2,3 берём размещения по два элемента:

1) 12 4) 21
2) 23 5) 32
3) 13 6) 31

$A_3^2 = 6$.

Теорема 3. $A_n^k = \dfrac{n!}{(n-k)!}$

2.5. Случайные величины и их распределения

Случайной величиной называется величина, которая в результате испытания примет одно и только одно возможное значение, заранее неизвестное и зависящее от множества причин, которые заранее не могут быть учтены.

Каждый исход испытания характеризуется случайной величиной.

$X = \{x_1, x_2, \ldots\}$

Случайные величины:

ДСВ – дискретные

НСВ - непрерывные

Дискретная случайная величина

ДСВ – СВ, которая принимает отдельные изолированные возможные значения.

Число возможных значений может быть конечным или бесконечным.

Непрерывная случайная величина

НСВ – СВ, которая может принимать любые значения из конечного или бесконечного промежутка.

Число возможных значений всегда бесконечно.

2.5.1. Числовые характеристики распределений

Закон распределения СВ полностью характеризует случайную величину.

Но для решения многих задач достаточно использовать числовые характеристики случайной величины:

- математическое ожидание M;
- дисперсия D;
- среднеквадратическое отклонение (СКО) σ.

1. Математическое ожидание

это наиболее вероятное, усредненное значение СВ.

ДСВ: $$M(X) = \sum_{i=1}^{n} p_i x_i$$

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

2. Дисперсия

это мера разброса СВ, то есть её усреднённое отклонение от математического ожидания.

ДСВ: $$D(X)=\sum_{i=1}^{n}\left(x_i-M(X)\right)^2 p_i$$

НСВ: $$D(X)=\int_{-\infty}^{+\infty}\left(x-M(x)\right)^2 f(x)dx$$

Связь между математическим ожиданием и дисперсией

для ДСВ и НСВ:

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

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

3. Среднеквадратическое отклонение

это также мера разброса СВ, но в отличие от дисперсии СКО измеряется в тех же единицах, что и сама СВ.

$$\sigma = \sqrt{D(X)}$$

Статистическое определение M, D

Если закон распределения СВ неизвестен, но имеется выборка значений СВ объёмом n, то можно приблизительно оценить математическое ожидание и дисперсию:

$$M(X) = \frac{\sum_{i=1}^{n} x_i}{n} \qquad D(X) = \frac{\sum_{i=1}^{n} \left(x_i - M(X)\right)^2}{n-1}$$

Почему в формуле дисперсии в знаменателе n-1?

Потому что входящая в формулу величина мат.ожидания М сама зависит от элементов выборки.

Если бы в формуле ещё одна величина была функцией элементов выборки, то пришлось бы взять n-2 и т.д.

Альтернативная формула для определения дисперсии

Дисперсию можно рассчитать статистически, не зная мат.ожидания:

$$D(X) = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}\right)$$

Центрированная СВ

ЦСВ – это отклонение СВ от её математического ожидания

$\dot{X} = X - M$

$M(\dot{X}) =$

$D(\dot{X}) =$

Центрированная СВ

ЦСВ – это отклонение СВ от её математического ожидания

$$\dot{X} = X - M$$

$$M(\dot{X}) = 0$$

$$D(\dot{X}) = D(X)$$

Нормированная СВ

это ЦСВ, выраженная в долях СКО.

$Z = \dot{X}/\sigma$

$M(Z) =$

$D(Z) =$

Нормированная СВ

это ЦСВ, выраженная в долях СКО.

$Z = \dot{X}/\sigma$

$M(Z) = 0$

$D(Z) = 1$

2.5.2. Законы распределения вероятностей ДСВ

ДСВ задаётся:

- рядом распределения;
- функцией распределения (интегральный закон)

а) Ряд распределения

это совокупность всех возможных значений x_i дискретной СВ X и соответствующих им вероятностей p_i.

$\mathbf{x_i}$	x_1	x_2			x_n
$\mathbf{p_i}$	p_1	p_2			p_n

Замечание: события x_i образуют группу гипотез =>

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

Графически эту таблицу задают гистограммой или полигоном

б) Функция распределения (интегральный закон)

это функция F(x), равная вероятности того, что СВ примет значение, не превышающее x.

$$F(x) = P(X < x) = \sum_{x_i \le x} p_i$$

Пример

Из 100 изделий, среди которых 10 дефектных, выбирают случайным образом 5 изделий.

Построить ряд распределения дефектных изделий в данной выборке.

Решение

$$P(X = x_i) = \frac{C_{10}^{x_i} C_{90}^{5-x_i}}{C_{100}^{5}}$$

$$p_1 = P(X = 0) = 0{,}583;$$

$$p_2 = P(X = 1) = 0{,}340;$$

$$p_3 = P(X = 2) = 0{,}070;$$

$$p_4 = P(X = 3) = 0{,}007;$$

$$p_5 = p_6 = 0;$$

$$F(0) = 0{,}583;$$

$$F(1) = 0{,}923;$$

$$F(2) = 0{,}993;$$

$$F(3) = 1{,}0.$$

2.5.3. Законы распределения вероятностей НСВ

Задать НСВ таблицей нельзя.

НСВ задают:

- функцией распределения $F(x)$ (интегральный закон);
- плотностью распределения $f(x)$ (дифференциальный закон).

а) Интегральный закон распределения

Функция распределения – это вероятность того, что НСВ примет значение, меньшее х.

$$F(x) = P(X < x)$$

Свойства:

1) $P(a < X < b) = F(b) - F(a)$
2) $F(x_1) < F(x_2) \quad <=> \quad x_1 < x_2$ (функция F не убывает)
3) $\lim\limits_{x \to \infty} F(x) = 1$
4) $\lim\limits_{x \to -\infty} F(x) = 0$

б) Дифференциальный закон распределения

Плотность распределения:

$$f(x) = \lim_{\Delta x \to 0} \frac{P(x < X < x + \Delta x)}{\Delta x}$$

Плотность распределения связана с функцией распределения:

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

б) Дифференциальный закон распределения

Свойства

$$1)\ f(x) \geq 0$$

$$2)\ f(x) = F'(x)$$

$$3)\ \int_{-\infty}^{\infty} f(x)dx = 1$$

$$4)\ P(a \leq X \leq b) = \int_{a}^{b} f(x)dx$$

Некоторые дискретные распределения

Рассмотрим следующие распределения ДСВ:

- биномиальное (закон Бернулли);
- Пуассона (закон редких событий)

1) Биномиальное распределение

Теорема 4.

Пусть p – вероятность события A.

Тогда вероятность того, что
из n независимых испытаний
ровно k исходов будут благоприятны, равна:

$$p_n(k) = C_n^{\ k} p^k (1-p)^{n-k}$$

- формула Бернулли.

1) Биномиальное распределение

У биноминального распределения достаточно просто рассчитываются M и D:

$$M = np$$

$$D = np(1 - p)$$

Пример на биномиальное распределение

Энергосистема имеет 150 генераторных блоков.

Вероятность отказа одного блока равна 0,06.

а) Определить вероятность того, что в данный момент не работают ровно 2 блока.

б) При каком числе k вероятность отказа одновременно k блоков будет максимальной?

Определить эту вероятность.

Решение

p = 0,06

1 – p = 1 – 0,06 = 0,94

$$p_{150}(2) = C_{150}{}^{2} \cdot 0{,}06^{2} \cdot 0{,}94^{150-2} = 0{,}00424.$$

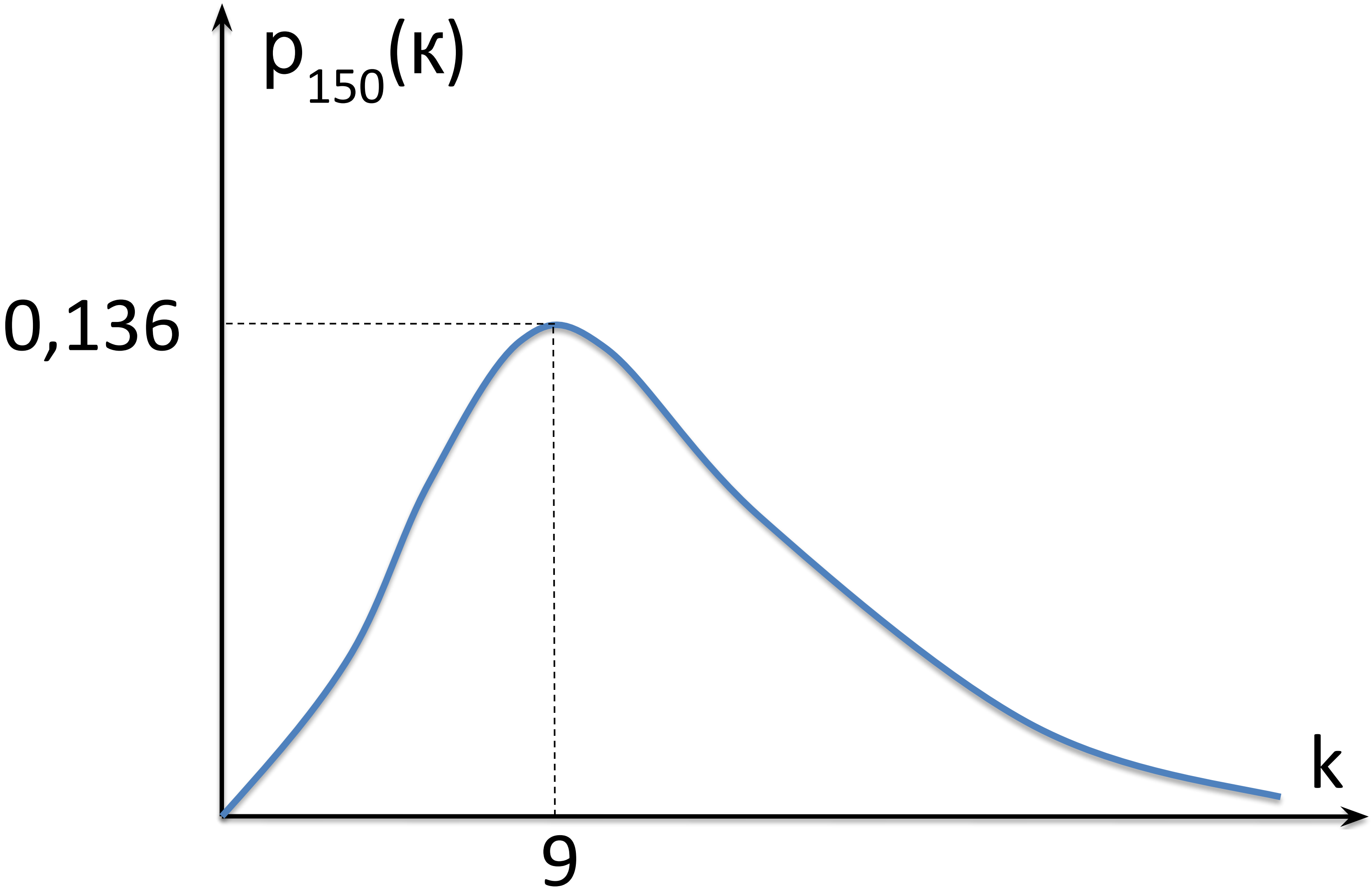
M = 150·0,06 = 9 = k

D = 9·0,94 = 8,46

σ = 2,91

$$p_{150}(9) = C_{150}{}^{9} \cdot 0{,}06^{9} \cdot 0{,}94^{150-9} = 0{,}136.$$

Распределение $p_{150}(к)$



2) Распределение Пуассона

Теорема 5.

Пусть p – вероятность события A.

При этом p – очень малое число.

Проводится серия из n испытаний.

Среднее число появления события A не меняется в различных сериях испытаний.

a = np = const.

Тогда вероятность появления k событий A равна

$$P_n(k) = \frac{a^k e^{-a}}{k!}$$

2) Распределение Пуассона

У распределения Пуассона достаточно просто рассчитываются М и D:

$$M = D = a$$

Пример

Завод производит реле с вероятностью дефекта 0,01.

Покупаем 200 реле.

Найти вероятность того, что среди купленных реле:

- не будет дефектных реле;
- будет 1 дефектное реле;
- будет 2 дефектных реле и т.д.

Построить ряд распределения числа дефектных реле среди купленных.

Решение

$$P_{200}(0) = \frac{2^0 e^{-2}}{0!} = 0{,}135$$

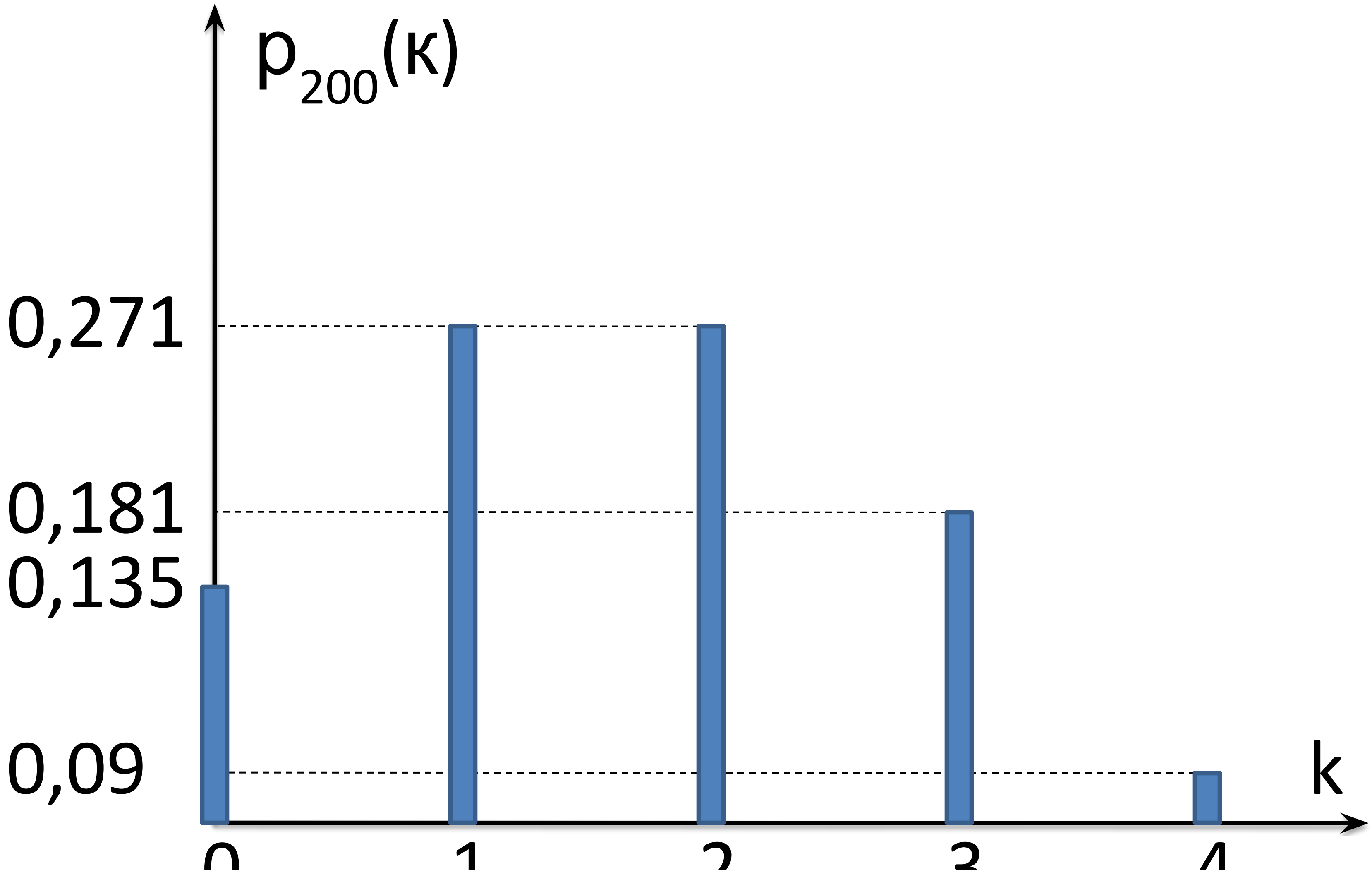
$$P_{200}(1) = \frac{2^1 e^{-2}}{1!} = 0{,}271$$

$$P_{200}(2) = \frac{2^2 e^{-2}}{2!} = 0{,}271$$

$$P_{200}(3) = \frac{2^3 e^{-2}}{3!} = 0{,}181$$

$$P_{200}(1) = \frac{2^4 e^{-2}}{4!} = 0{,}09$$

Распределение $p_{200}(к)$



Некоторые непрерывные распределения

Рассмотрим следующие распределения НСВ:

- экспоненциальное;
- нормальное.

1) Экспоненциальное распределение

Задаётся плотность распределения:

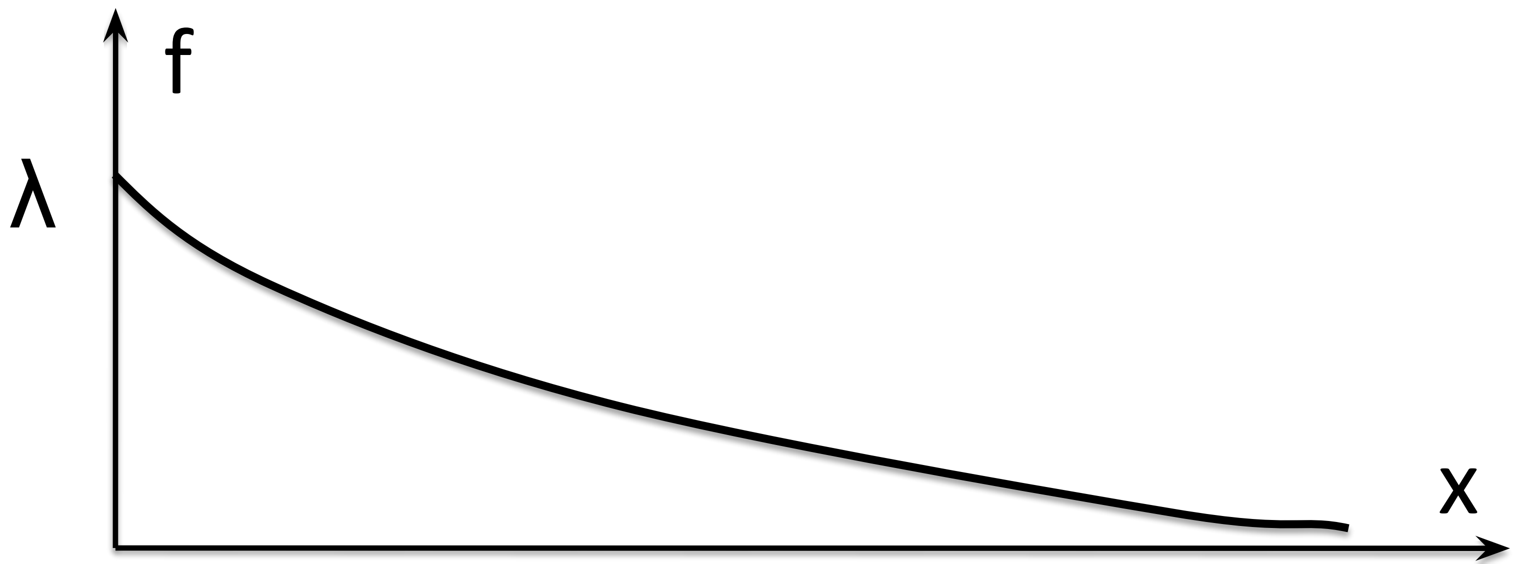
$f(x) = \lambda \exp(-\lambda x)$,

где $\lambda = \mathrm{const} > 0$ – единственный параметр распределения;

$x \geq 0$

Это распределение моделирует **время** между двумя последовательными совершениями одного и того же события.

1) Экспоненциальное распределение



В этом распределении
x – время (например, в ч);
λ – средняя интенсивность события (например, ч$^{-1}$)

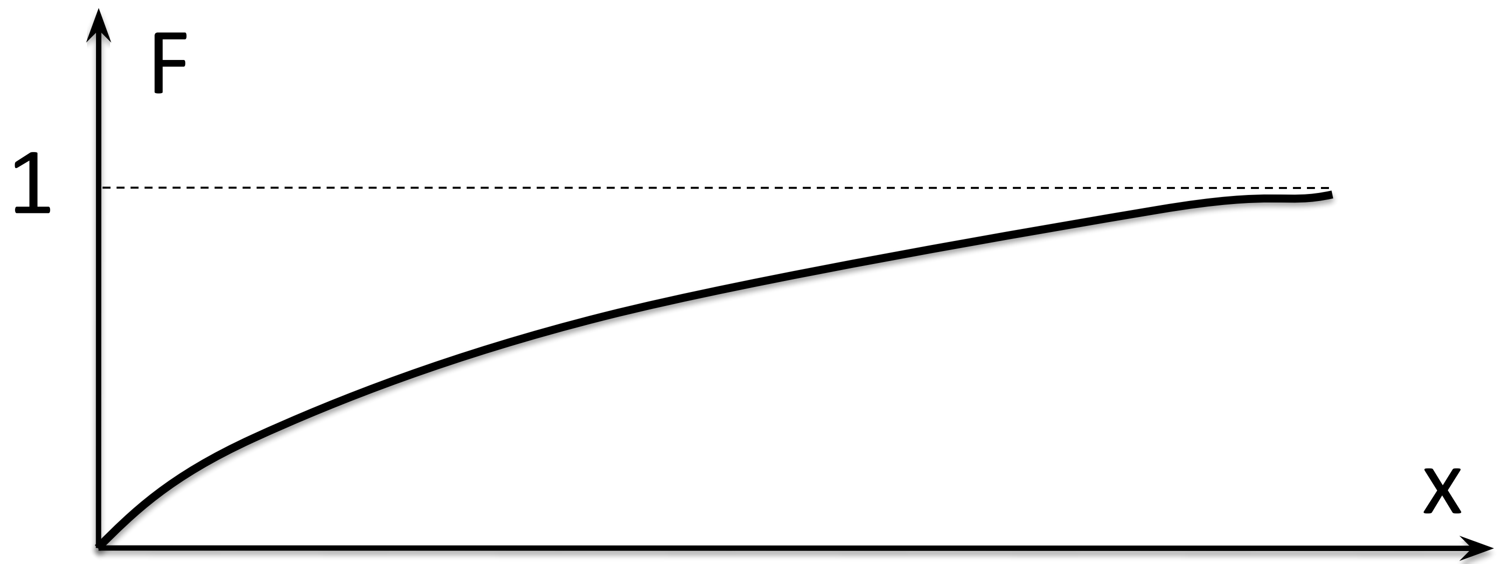
1) Экспоненциальное распределение

$F(x) = 1 - \exp(-\lambda x)$,

$M(X) = 1/\lambda$

$D(X) = 1/\lambda^2$

1) Экспоненциальное распределение



Функция распределения

Пример на экспоненциальное распределение

В среднем выключатель отказывает раз в 20 лет ($\lambda = 1/20$).

Тогда вероятность отказа выключателя:

- за 10 лет: $1 - \exp(-10/20) = 0{,}39$;
- за 20 лет: $1 - \exp(-20/20) = 0{,}63$;
- за 40 лет: $1 - \exp(-40/20) = 0{,}86$;
- за 60 лет: $1 - \exp(-60/20) = 0{,}95$.

2) Нормальное распределение

Плотность распределения:

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

Функция распределения:

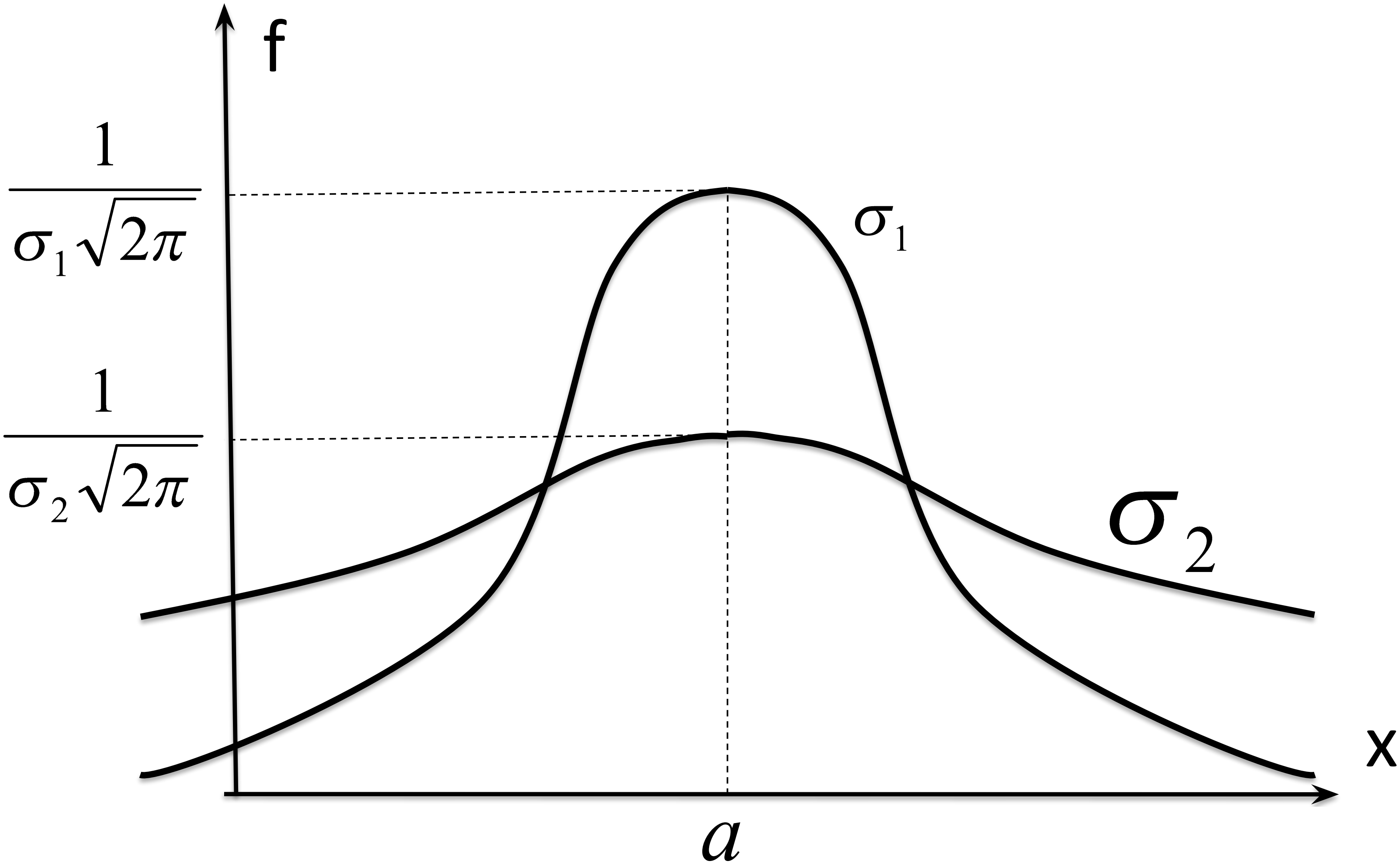
$$F(x) = \int_{-\infty}^{x} f(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{(x-a)^2}{2\sigma^2} \right) dx$$

2) Нормальное распределение

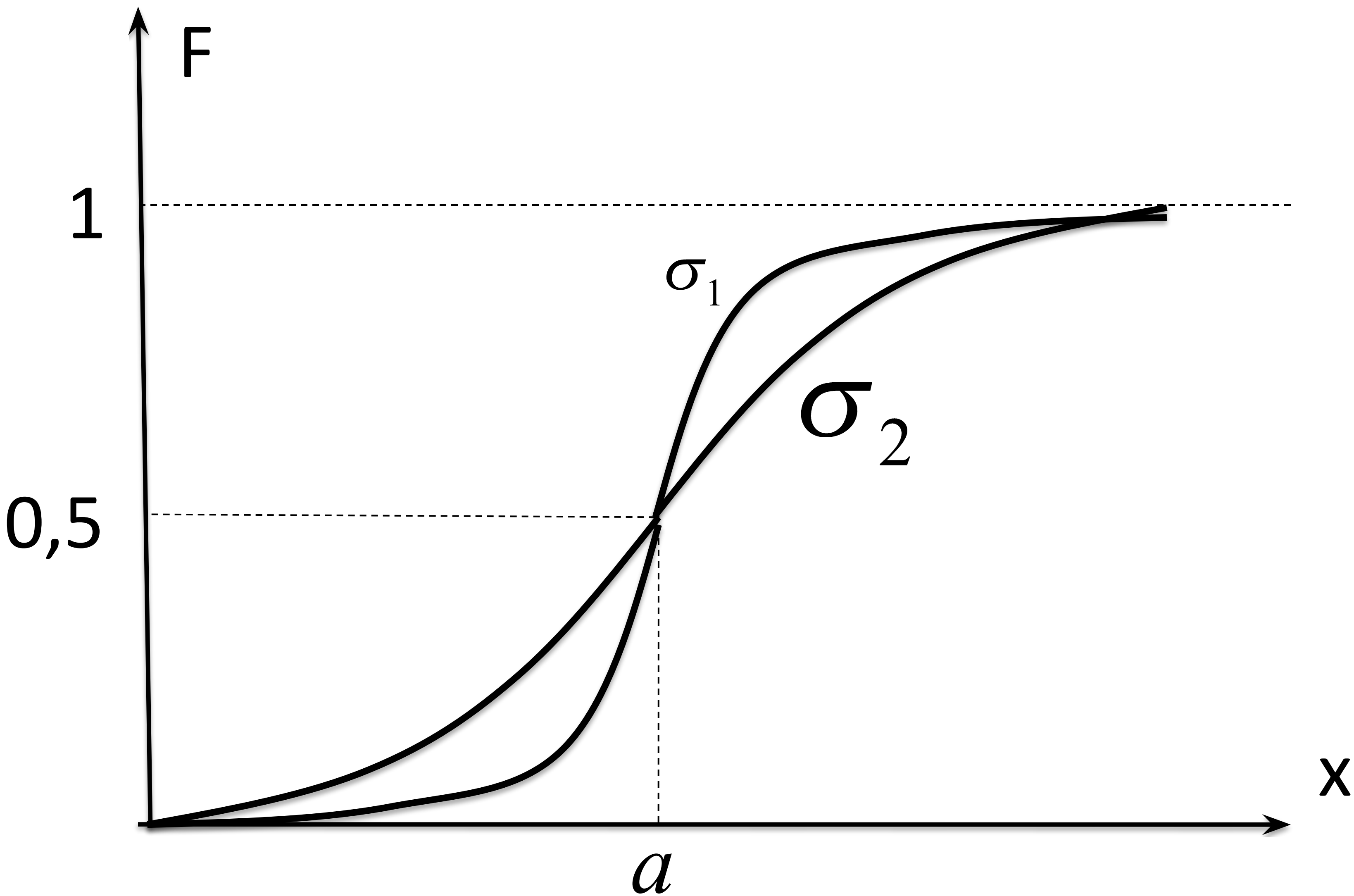
В отличие от экспоненциального распределения (с единственным параметром λ), характеризуется двумя параметрами:

- математическое ожидание a;
- СКО σ.

2) Нормальное распределение



2) Нормальное распределение



2) Нормальное распределение

Видно, что:

- график f(x) симметричен относительно оси x = a;
- график F(x) симметричен относительно точки (a; 0,5).

Отсюда – идея центрировать эти функции:

- f(x), чтобы она стала чётной;
- F(x), чтобы она стала нечётной.

2) Нормальное распределение

Пусть $z = (x - a)/\sigma$.

Этот аргумент – безразмерный, т.к. x, a, σ имеют одинаковые размерности.

То есть функцию не только центрируют, но и нормируют.

2) Нормальное распределение

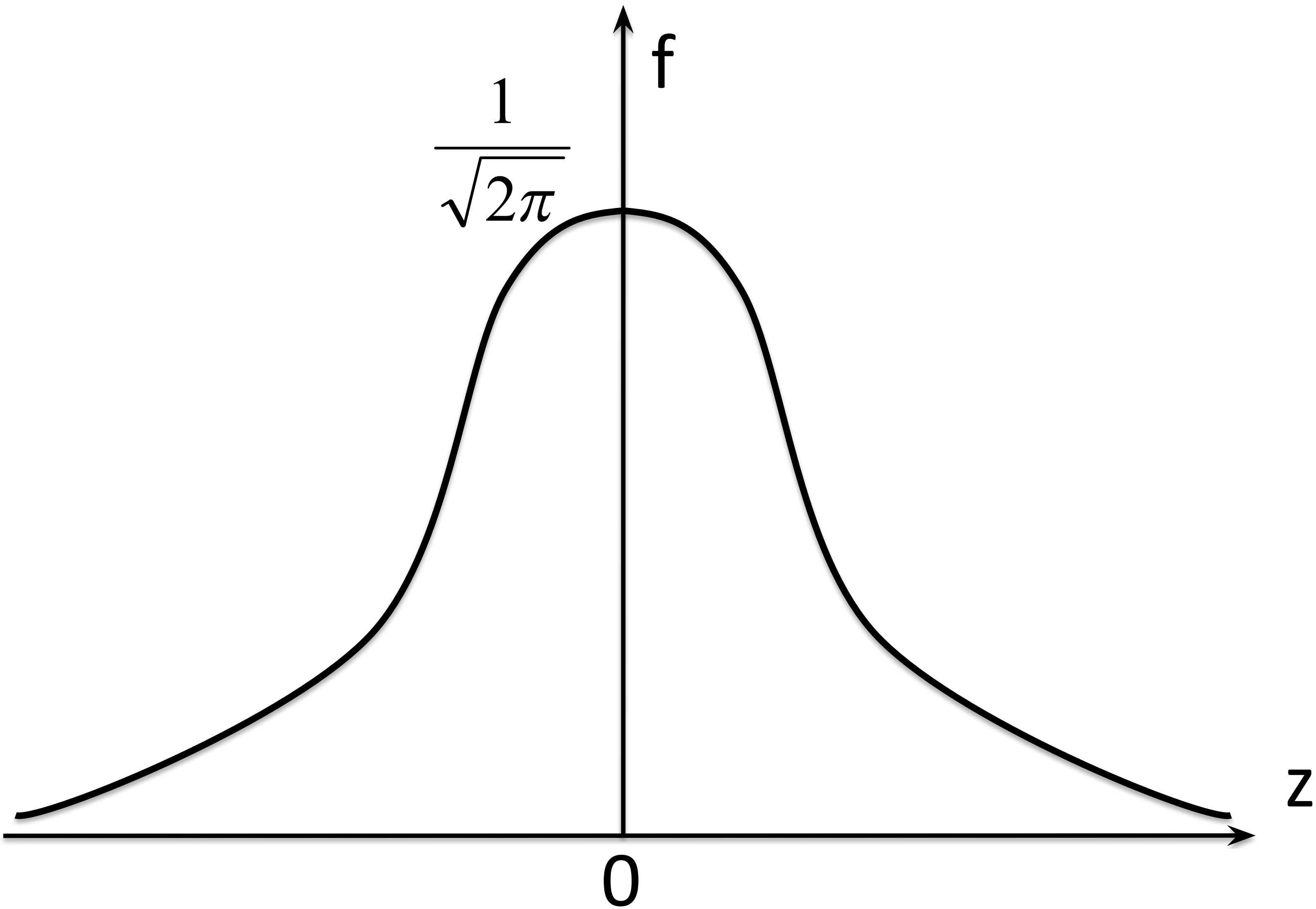
Плотность распределения:

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2} \right)$$

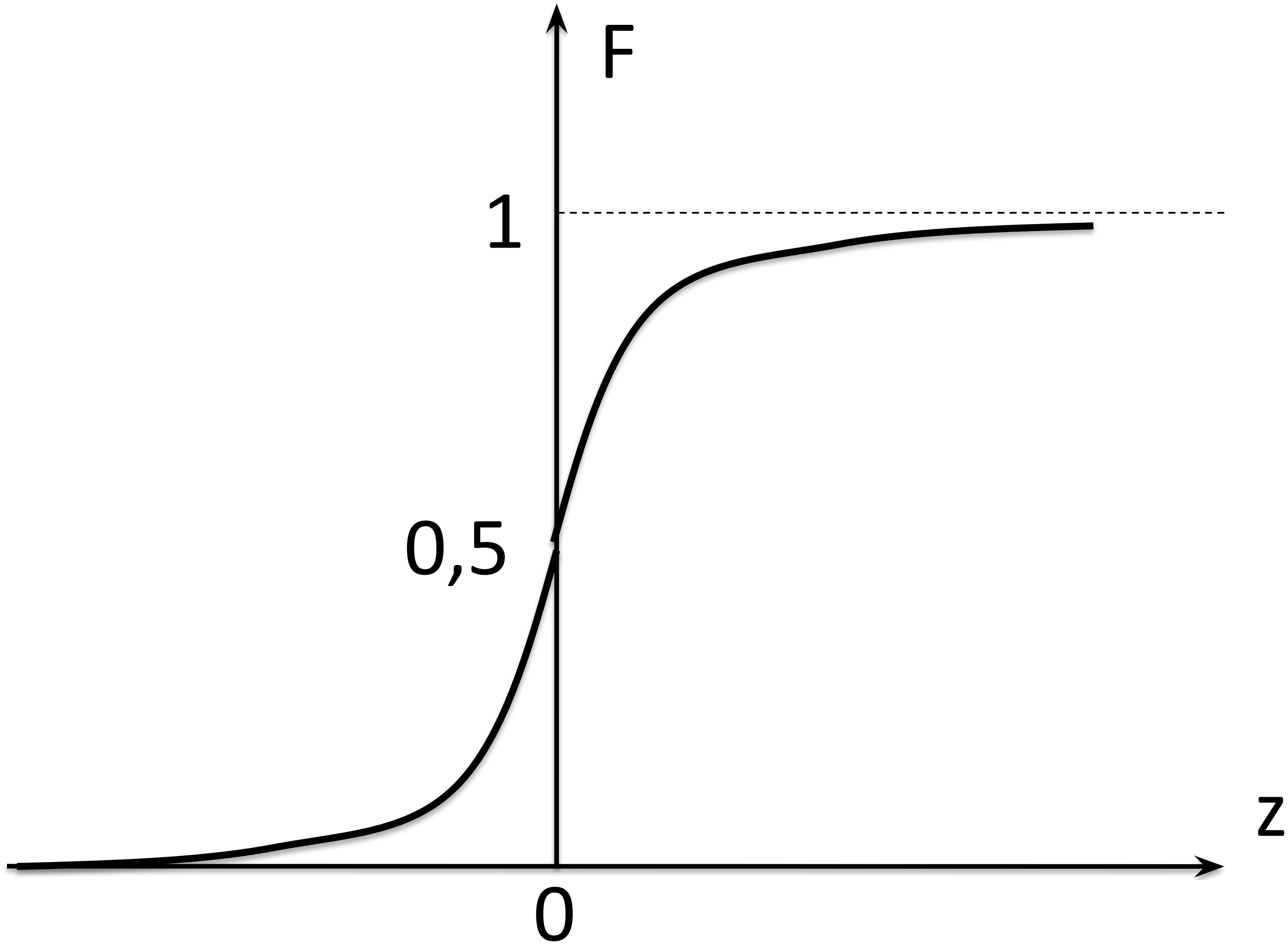
Функция распределения:

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(-\frac{z^2}{2} \right) dz$$

2) Нормальное распределение



2) Нормальное распределение



2) Нормальное распределение

Функция F(z) по-прежнему неудобна, т.к.:

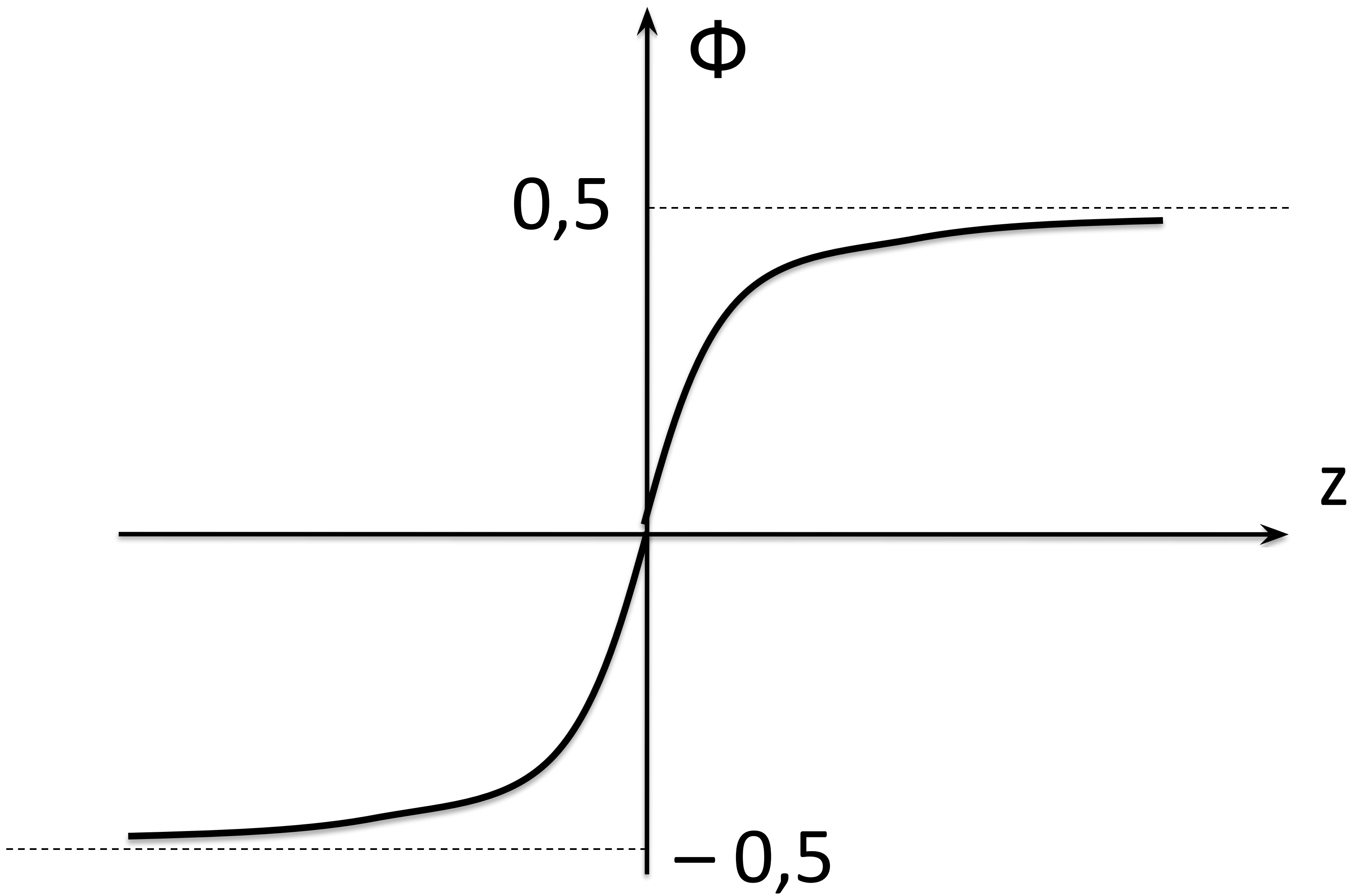
1) она не является ни чётной, ни нечётной;
2) интеграл $\int_{-\infty}^{z} \exp\left(-\frac{z^2}{2}\right) dz$ не берётся.

Введём функцию Лапласа:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{0}^{z} \exp\left(-\frac{z^2}{2}\right) dz$$

Докажем, что F(z) = 0,5 + Ф(z)

2) Нормальное распределение



2) Нормальное распределение

Функция Лапласа Ф(z) нечётная, поэтому её можно задать только при z ≥ 0.

Интеграл $\int_0^z \exp\left(-\frac{z^2}{2}\right)dz$ также не берётся, но

его значения можно задать таблицей.

z	0	0,2	0,4	0,6	0,8	1	1,2
Ф(z)	0	0,0793	0,1554	0,2257	0,2881	0,3413	0,3849

z	1,4	1,6	1,8	2	2,2	2,4	2,6
Ф(z)	0,4192	0,4452	0,4641	0,4772	0,4861	0,4918	0,4953

2) Нормальное распределение

С помощью функции Лапласа можно вычислить вероятность того, что случайная величина примет значение между x_1 и x_2:

$P(x_1 < X < x_2) = F(x_2) – F(x_1) =$

$= 0,5 + \Phi(z_2) – 0,5 – \Phi(z_1) = \Phi(z_2) – \Phi(z_1)$,

где $z_{1,2} = (x_{1,2} – a) / \sigma$

Пример на вычисление вероятности для нормального распределения

Рассматривается НСВ – мощность нагрузки, МВт.

Данная НСВ имеет нормальное распределение с мат. ожиданием 10 МВт и СКО 2 МВт.

Найти вероятность того, что мощность нагрузки примет значение от 12 до 14 МВт.

z	0	0,2	0,4	0,6	0,8	1	1,2
Ф(z)	0	0,0793	0,1554	0,2257	0,2881	0,3413	0,3849

z	1,4	1,6	1,8	2	2,2	2,4	2,6
Ф(z)	0,4192	0,4452	0,4641	0,4772	0,4861	0,4918	0,4953

Пример на вычисление вероятности для нормального распределения

Решение:

$z_1 = (x_1 - a) / \sigma = (12 - 10) / 2 = 1;$

$z_2 = (x_2 - a) / \sigma = (14 - 10) / 2 = 2;$

$P(12 < X < 14) = \Phi(2) - \Phi(1) = 0{,}4772 - 0{,}3413 =$

$= 0{,}1359.$