6. Нормальный закон распределения

Стандартный нормальный закон

Функция Лапласа

Вероятность попадания нормальной СВ в заданный интервал

Правило «трех сигм»

Непрерывное распределение, которое занимает наиболее важное положение в теории и практике статистики – *распределение Гаусса*, или *нормальное распределение*

Это так, и этот ЗР действительно важен вот почему:

«Нормальный» можно понимать как выражение нормы, некоторого стандарта, «образца поведения» СВ

to be continued

1) Чаще всех в практических задачах («приложениях»)

2) Им часто аппроксимируют другие законы

3) Является пределом для других ЗР при некотором $n \to \infty$ (биномиальный при ∞ числе испытаний)

4) Занимает центральное положение в семействе ЗР (моделей распределений), центр симметрии (А, Е = 0)

Часто встречается в с
тем, что

Примеры:

Случайная величина X распределена норма
когда все ее значения x формируются под суммарным воздействием очень большого числа случайных факторов, эффекты каждого из них малы, сравнимы по величине и равновероятны по знаку

Ошибки измерений часто нормальны
(такие распределения обнаружили астрономы в 18 веке)

В статистике – распределение выборочного среднего стремится к нормальному, даже если отдельные наблюдения не нормальны

Некоторые характеристики живых организмов подчинены нормальному закону

В производстве и контроле качества – % брака, производительность, размеры деталей …

В сфере финансов, рынка, в деловой практике – отношение «цена / доход», годовая зарплата …

СВ распределена по нормальному закону если ее функция плотности равна

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

Тогда

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp[-\frac{(x-\mu)^2}{2\sigma^2}]$$

Функция распределения нормальной величины определяется как

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{x} \exp[-\frac{(u-\mu)^2}{2\sigma^2}]du$$

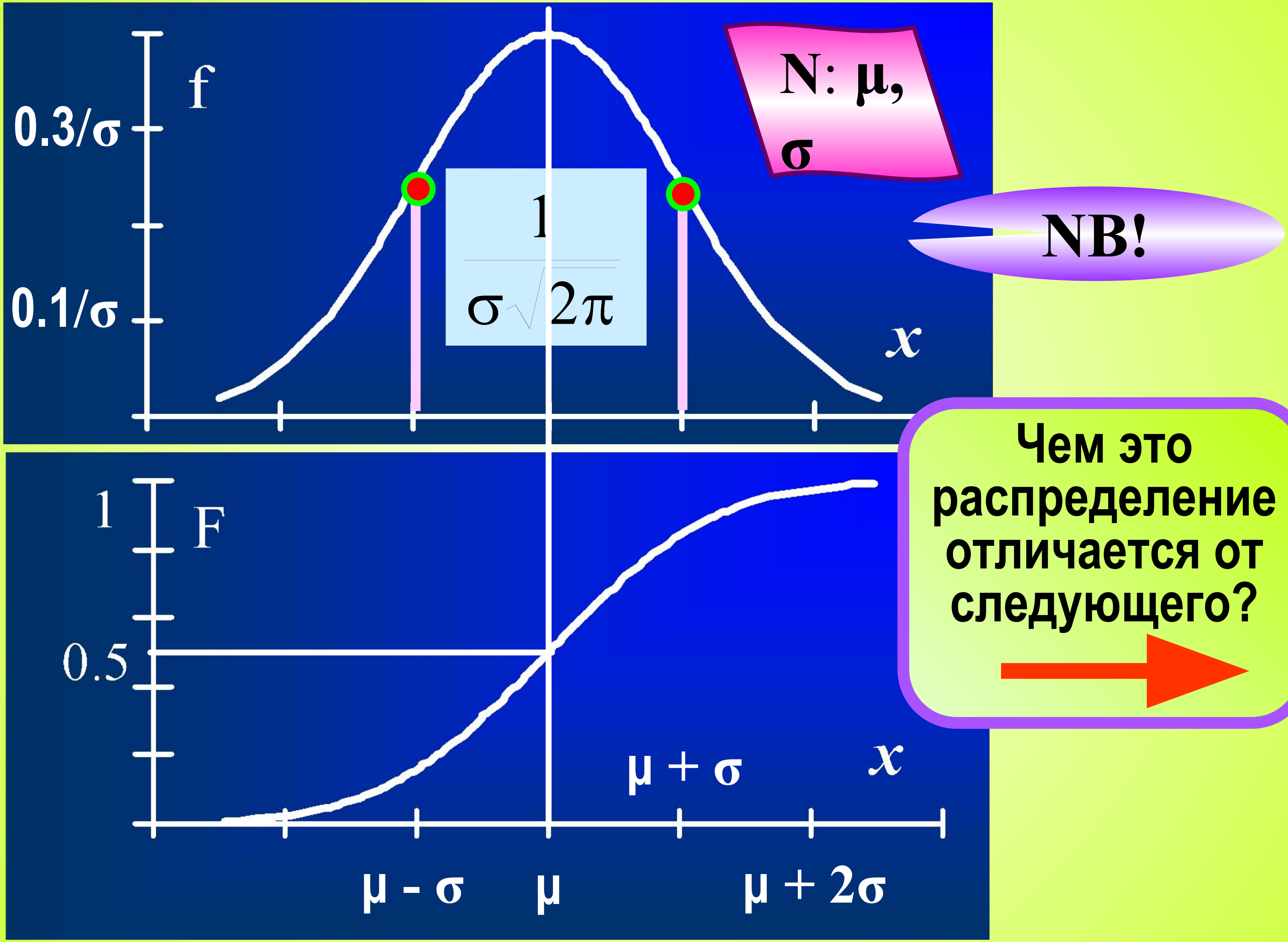
Параметры μ и σ^2 – матожидание и дисперсия

Это двухпараметрический закон →

1) **если известно, что распределение нормально, знание μ и σ дает полное описание СВ**
2) **<u>все нормальные СВ отличаются только μ и σ</u>**

!

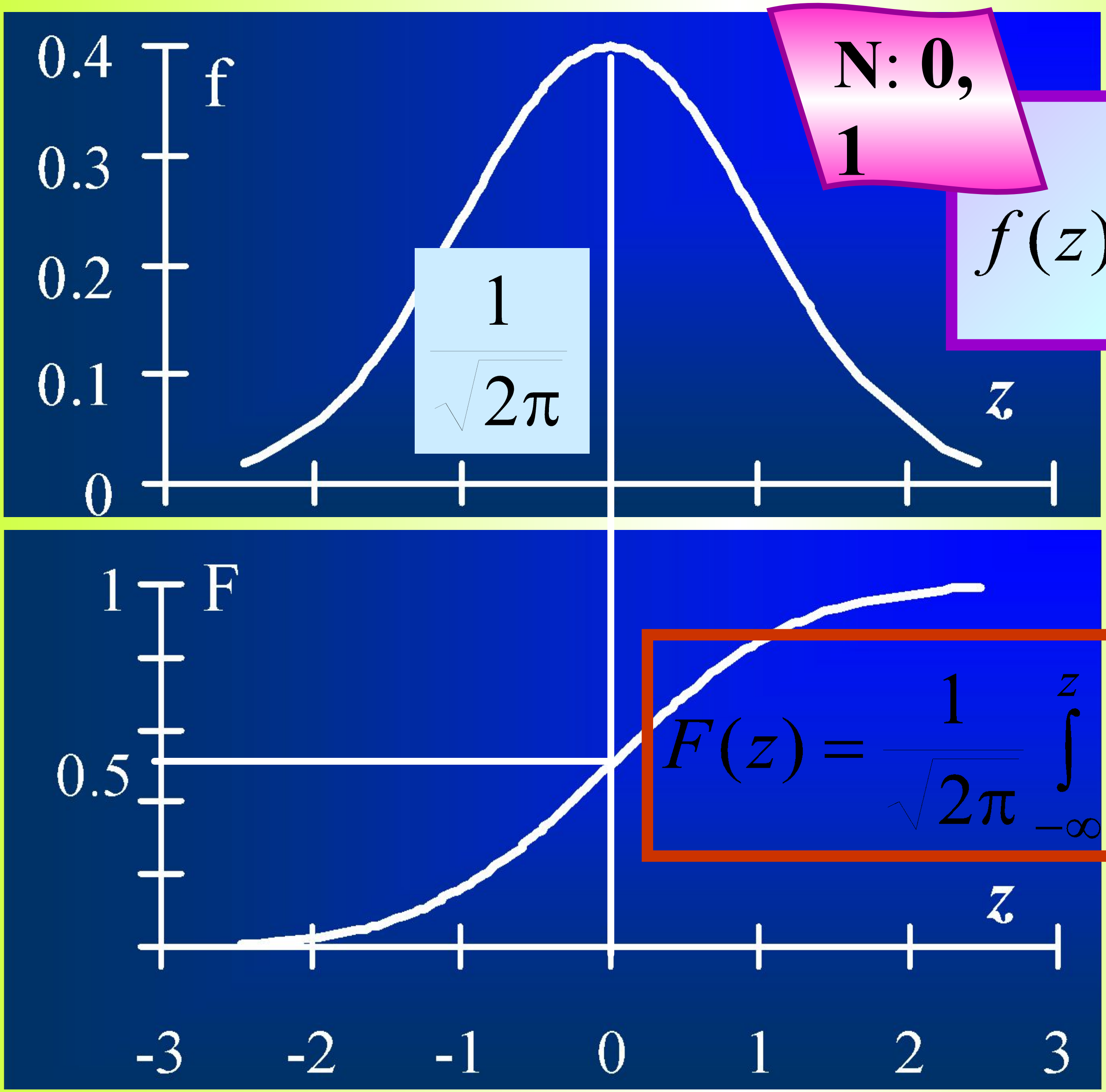
Используют специальное обозначение нормальных величин – N: μ, σ



Чем это распределение отличается от следующего?

N: 0, 1

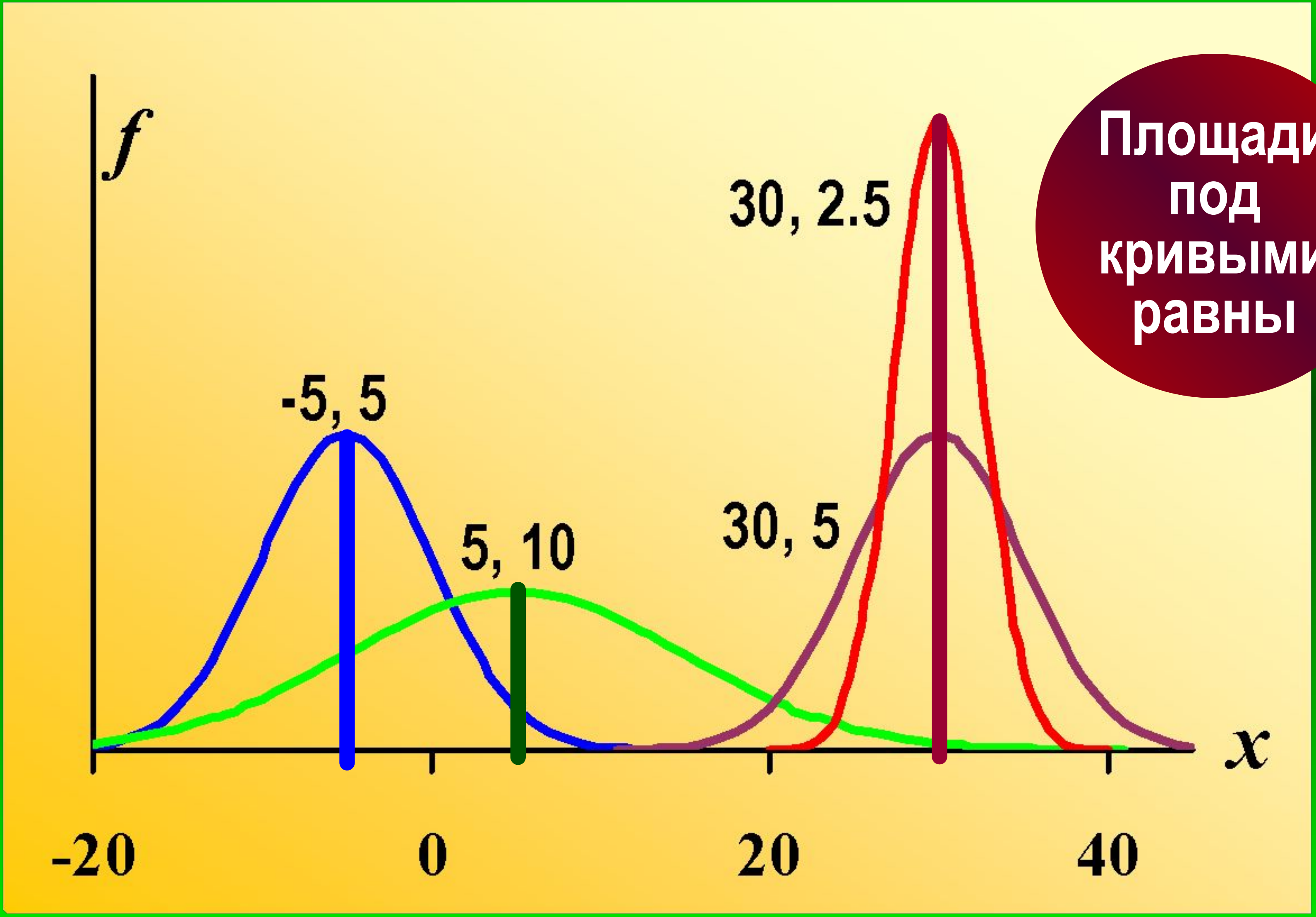
$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$



$$\frac{1}{\sqrt{2\pi}}$$

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

Коллекция нормальных распределений



Площади под кривыми равны

У нормальных распределений с разными μ и σ

Различия:

- привязаны к разным точкам μ числовой оси
- чем < σ, тем > площадь под кривой вблизи μ, > вероятность значений вблизи центра

Общее:

* унимодальные, высшая ордината в точке μ = Mo = Me, хвосты → 0 в обоих направлениях (ПР → 0 при x → ± ∞)

* симметричные, равноудаленные от μ меньшие и большие x имеют равные p

* колоколообразная форма кривой, с точкой перегиба на расстоянии σ от μ

Стандартный нормальный закон

Это N: 0, 1 – нормальный закон с $\mu = 0$ и $\sigma = 1$

Будучи стандартом для других ЗР, нормальное распределение имеет свой

Узнаете стандартизованное (единичное) отклонение?

Измеряет в «сигмах» отклонения x от центра распределения

Любое нормальное распределение можно записать в стандартной форме с помощью нормализованной переменной

$$z = \frac{x - \mu}{\sigma}$$

Зачем нужна нормализация и стандартный нормальный ЗР?

Смысл есть, весьма утилитарный!

Дело в том, что

!

Из $x = \sigma z + \mu$ – $f(z)/\sigma$, $dx = \sigma\, dz$

и !

$$\mathrm{F}(x) = F(z)$$

Тогда

$$P\{X < x\} = P\{Z < z\} = \mathrm{F}[z = (x - \mu) / \sigma]$$

$$P\{x_1 < X < x_2\} = F[z_2 = (x_2 - \mu)/\sigma] - F[z_1 = (x_1 - \mu)/\sigma]$$

Вместо расчета интегралов всякий раз для разных μ и σ, можно использовать раз и навсегда рассчитанные значения стандартной нормальной ФР

?

называется *интеграл вероятности*

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

Однако,

есть таблицы значений

обычно **при расчете вероятностей значений нормальной величины в том или ином интервале**

вместо интеграла вероятности **используется функция Лапласа**

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_0^z \exp\left(-\frac{u^2}{2}\right) du$$

У нее следующие свойства:

$$\Phi(0) = 0; \quad \Phi(\infty) = 0.5;$$

это нечетная функция, $\Phi(-z) = -\Phi(z)$

Поскольку $F(z) = 0.5 + \Phi(z)$, **то**

$$P\{x_1 < X < x_2\} = \Phi(z_2) - \Phi(z_1)$$

Удобное практическое правило

Вероятность того, что нормальная величина примет значение из некоторого интервала равна разности значений функции Лапласа для нормализованных верхней и нижней границ этого интервала

Пример

Производительность за смену (Y) **распределена нормально, μ** = 160 изд., **σ** = 20.

Экономическая целесообразность требует, чтобы выпускалось не более 200 и не менее 150 шт.

$P\{150 < Y < 20$

$\Phi[z_2 =$

$\Phi(2) + \Phi(0.5) =$

Для более надежного выполнения требований необходимо:

статистически ? –

увеличить μ,

снизить σ !

организационно ???

только 67% производственных ситуаций отвечают требованиям

Важный пример – с обобщением

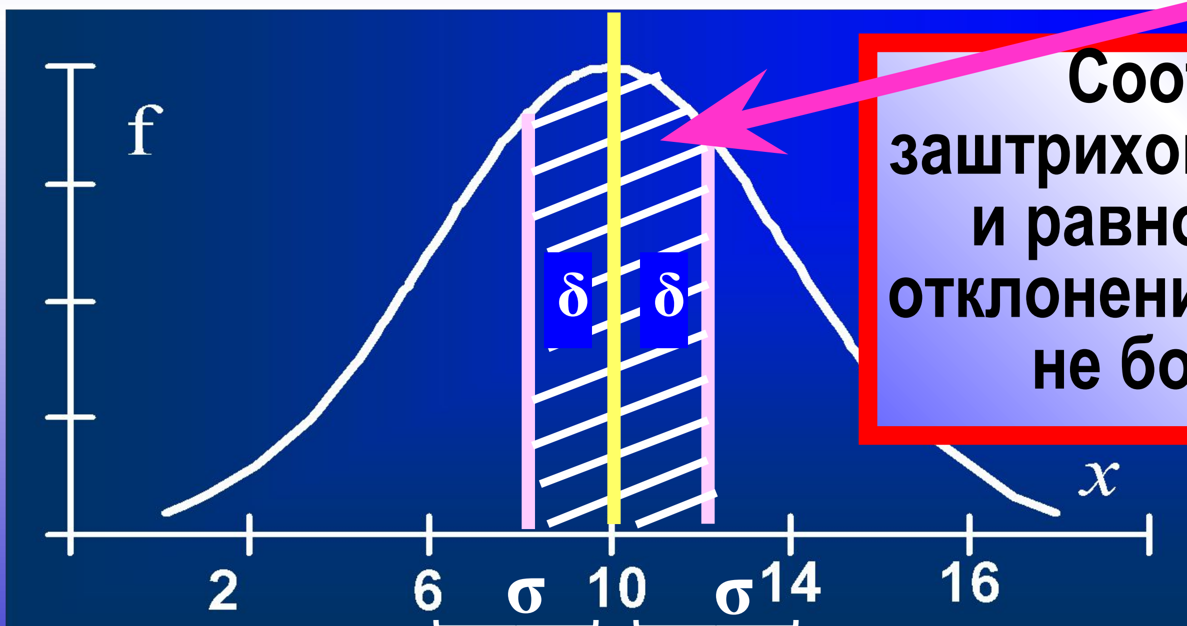
... ена нормально, μ = 10, σ = 4

т.е., вероятности попасть в интервал симметричный относительно μ

$$\ldots 2) = \Phi\left(\frac{12-10}{4}\right) - \Phi\left(\frac{8-10}{4}\right) =$$

$$\ldots (\delta = 2)/\sigma) = 2 \cdot 0.1915 = 0.383$$

?



Соответствует заштрихованной площади и равно вероятности отклонений от μ не более, чем на δ = 2

Общее правило

$$P(\mu - \delta < X < \mu + \delta) =$$

$$P(|X - \mu| < \delta) = \quad 2\Phi(z = \frac{\delta}{\sigma})$$

Вероятность того, что
(при измерении, управлении, производстве …)
отклонение от μ
(неизвестного истинного, предписанного … значения)
в обоих направлениях не превысит
максимально допустимого δ
равна удвоенному значению функции Лапласа
от «δ / σ стандартных отклонений»

Очевидно!

$$P(|X-\mu|>\delta)=1-2\Phi(\frac{\delta}{\sigma})$$

определяет !!!

шансы отклонений, превышающих заданное δ

!

риск выйти за нормативные границы

Примеры с важным обобщением

$\delta = \sigma \to z = \delta/\sigma = 1 \to 2\Phi(1) = 0.6826$

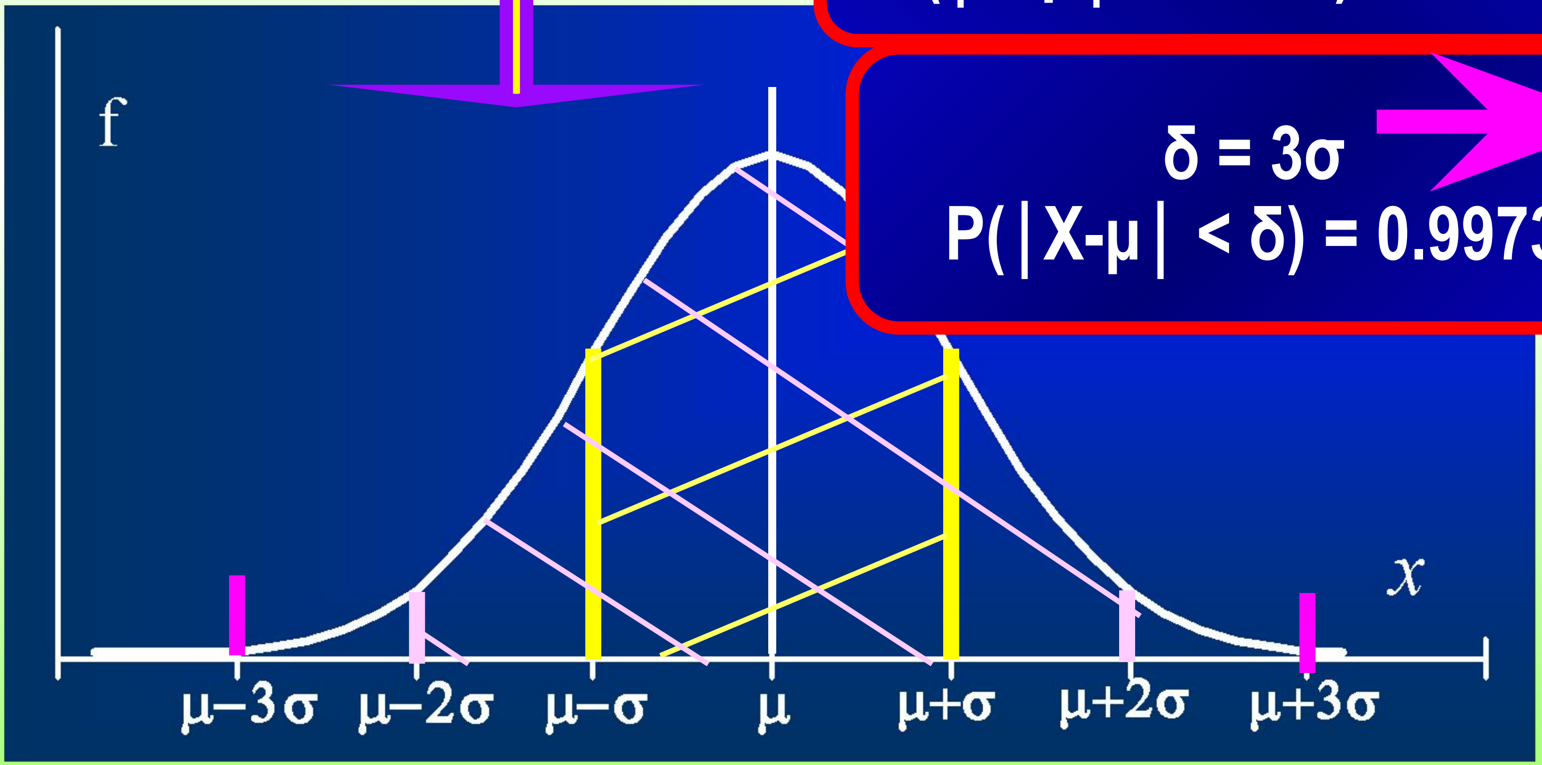
68.3% значений величины X в интервале $(\mu - \sigma, \mu + \sigma)$

$\delta = 2\sigma \to z = 2 \to$

$P(|X-\mu| < \delta=2\sigma) = 0.9545$

$\delta = 3\sigma$

$P(|X-\mu| < \delta) = 0.9973$



99.7% значений нормально распределенной величины попадают в интервал ($\mu \pm 3\sigma$)

Тогда

$$P(|X-\mu| > 3\sigma) = 1 - 0.9973 = 0.0027$$

только 27 из 10000 можно ожидать дальше от среднего, чем 3σ

Вспомнив про уровень значимости, можно считать это невозможным событием

«К

Если СВ нор
ее отклоне

The End

The End

<u>В примере с производительностью:</u>

количество производимой за смену продукции может находиться в пределах от 100 до 220 шт.