

**Статистические  
методы обработки  
медико-биологических  
данных**

**Нормальный закон  
распределения**

# Тема: Статистические методы обработки медико-биологических данных

## Нормальный закон распределения.

### План лекции:

- Понятие случайных **дискретных** и **непрерывных** величин.
- Распределения и характеристики случайных величин.
- Нормальный закон **распределения**.  
**Кривая Гаусса** и ее особенности.  
Правило «трёх **сигм**».

**В медицине необходимо вести учет, анализ и прогноз различных массовых явлений. В целом, массовым явлениям присущи свои особые закономерности. К доктору обращаются пациенты с различными заболеваниями. Болезнь конкретного человека - случайное событие у врача. Но случайные события предсказуемы, например, в период эпидемии гриппа наиболее часто встречаются заболевания гриппом.**

**Закономерности **массовых**  
**случайных событий** -  
статистических данных,  
отражающих эти события, -  
изучаются с помощью  
математической статистики.**

**Типичная задача математической статистики - это приближенная оценка неизвестной вероятности случайного события по результатам наблюдений, экспериментов, когда событие может происходить или не осуществляться.**

# Случайной величиной

называется переменная  
величина, значение  
которой зависит от исхода  
некоторого испытания.

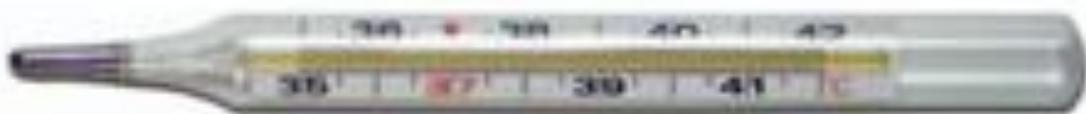
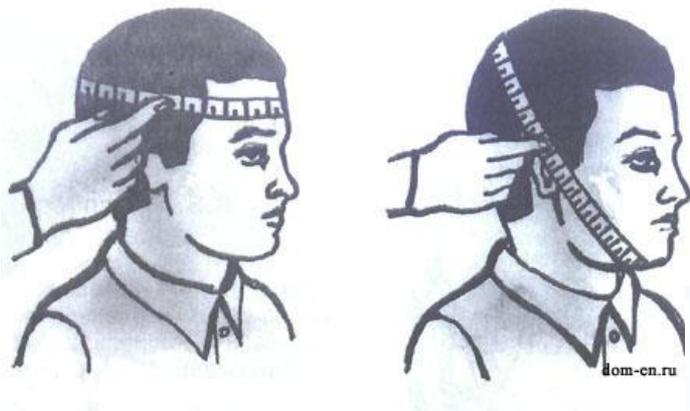
Дискретная

Непрерывная

Дискретной называется случайная величина, которая может принимать значения некоторой конечной или бесконечной числовой последовательности (число слов в тексте, студентов в аудитории, больных в клинике...)

**Непрерывной** называется  
случайная величина,  
которая может принимать  
любые значения внутри  
некоторого интервала  
(масса, температура,  
рост...)

# Дискретная ? или



# Непрерывная?

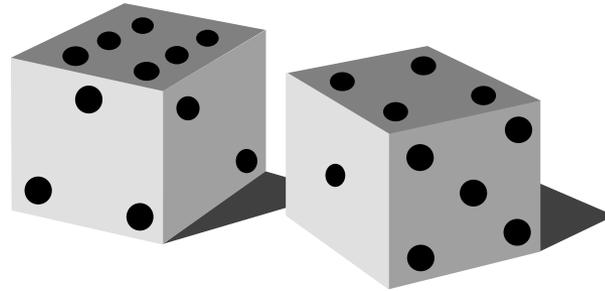


**Статистический ряд** – результаты измерений для статистического исследования, записанные последовательно по порядку их получения. Удобнее представить в таблице.

$X_1, X_2, X_3, \dots, X_n$

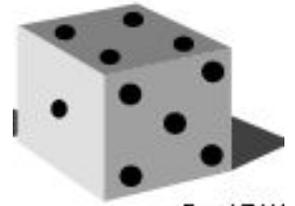
$X_1$	$X_2$	$X_3$	...	...	...	...	...
...	...	...	...	...	...	...	...
.....	....	.....	....	....	....	.....	....
.		.					
.....	....	..	....	....	....	.....	$X_n$

# Распределение дискретной случайной величины.



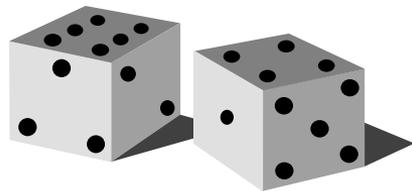
**Дискретная случайная величина считается заданной, если указаны ее возможные значения и соответствующие им вероятности**

Совокупность  $X$  и  $P$  называется распределением дискретной случайной величины.



Дискретные случайные величины $x_i$	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$
Вероятность $p_i$	$P_1$	$P_2$	$P_3$	$P_4$	$P_5$	$P_6$

$n$  – общее число случайных событий



Дискретные случайные величины $x_i$	$X_1$	$X_2$	$X_3$	...	...	$X_n$
Вероятность $p_i$	$P_1$	$P_2$	$P_3$	...	...	$P_n$

<p><b>Дискретные случайные величины</b></p> <p><math>X_i</math></p>	<p><math>X_1</math> <b>1</b></p>	<p><math>X_2</math> <b>2</b></p>	<p><math>X_3</math> <b>3</b></p>	<p><math>X_4</math> <b>4</b></p>	<p><math>X_5</math> <b>5</b></p>	<p><math>X_6</math> <b>6</b></p>
<p><b>Вероятность <math>p_i</math></b></p>	<p><math>P_1</math> <b>1/6</b></p>	<p><math>P_2</math> <b>1/6</b></p>	<p><math>P_3</math> <b>1/6</b></p>	<p><math>P_4</math> <b>1/6</b></p>	<p><math>P_5</math> <b>1/6</b></p>	<p><math>P_6</math> <b>1/6</b></p>



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

**условие  
нормировки  
дискретных  
случайных  
величин**

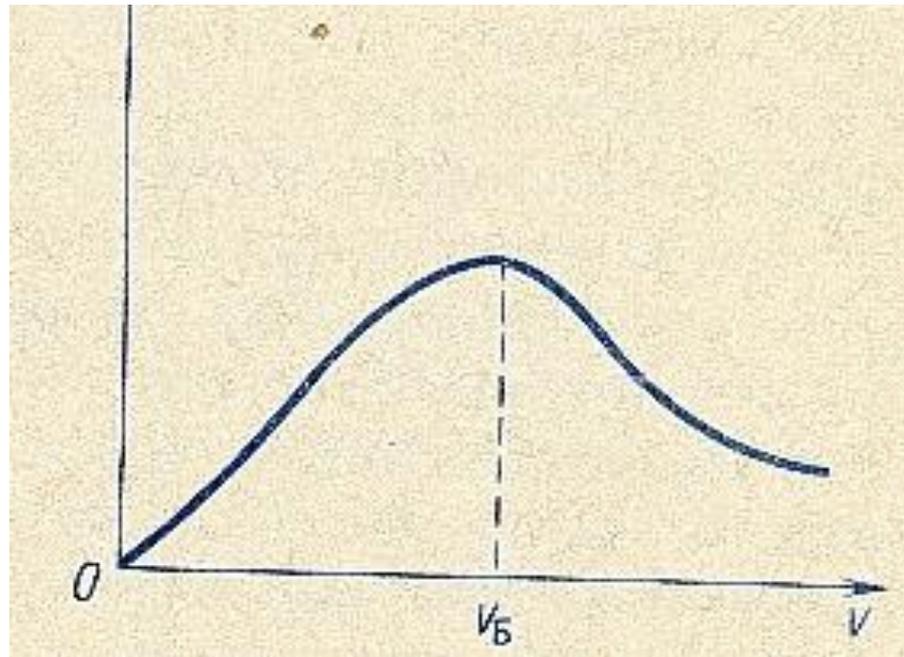
# Различные распределения

## 1. Биномиальное распределение

(позволяет определить  
вероятность того, что

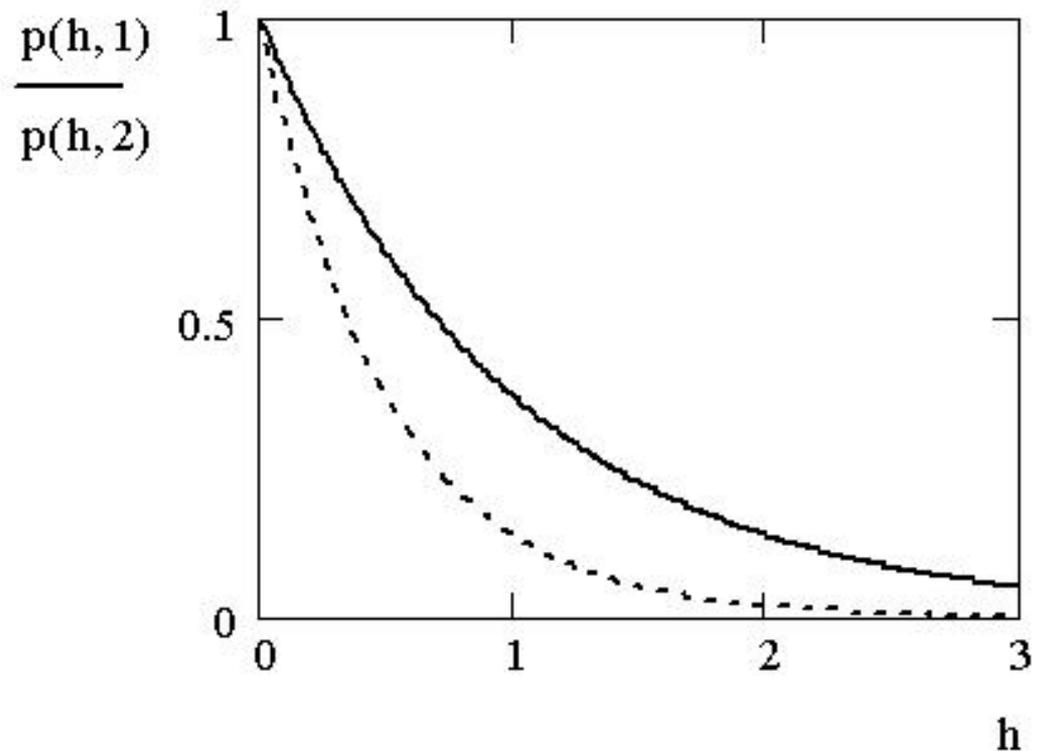
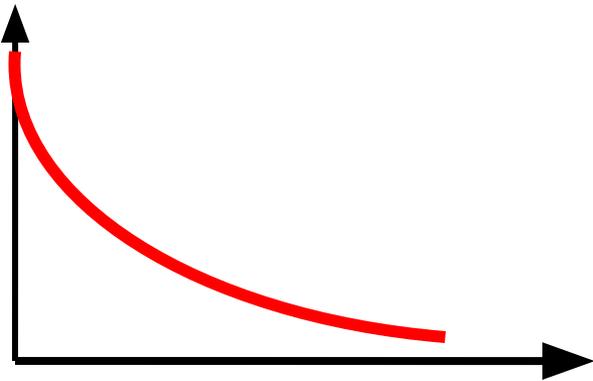
событие  $A$  произойдет  $m$  раз  
при  $n$  испытаниях).

**2. Распределение Максвелла**  
(распределение молекул газа по скоростям, кинетическим энергиям). График - кривая Максвелла.

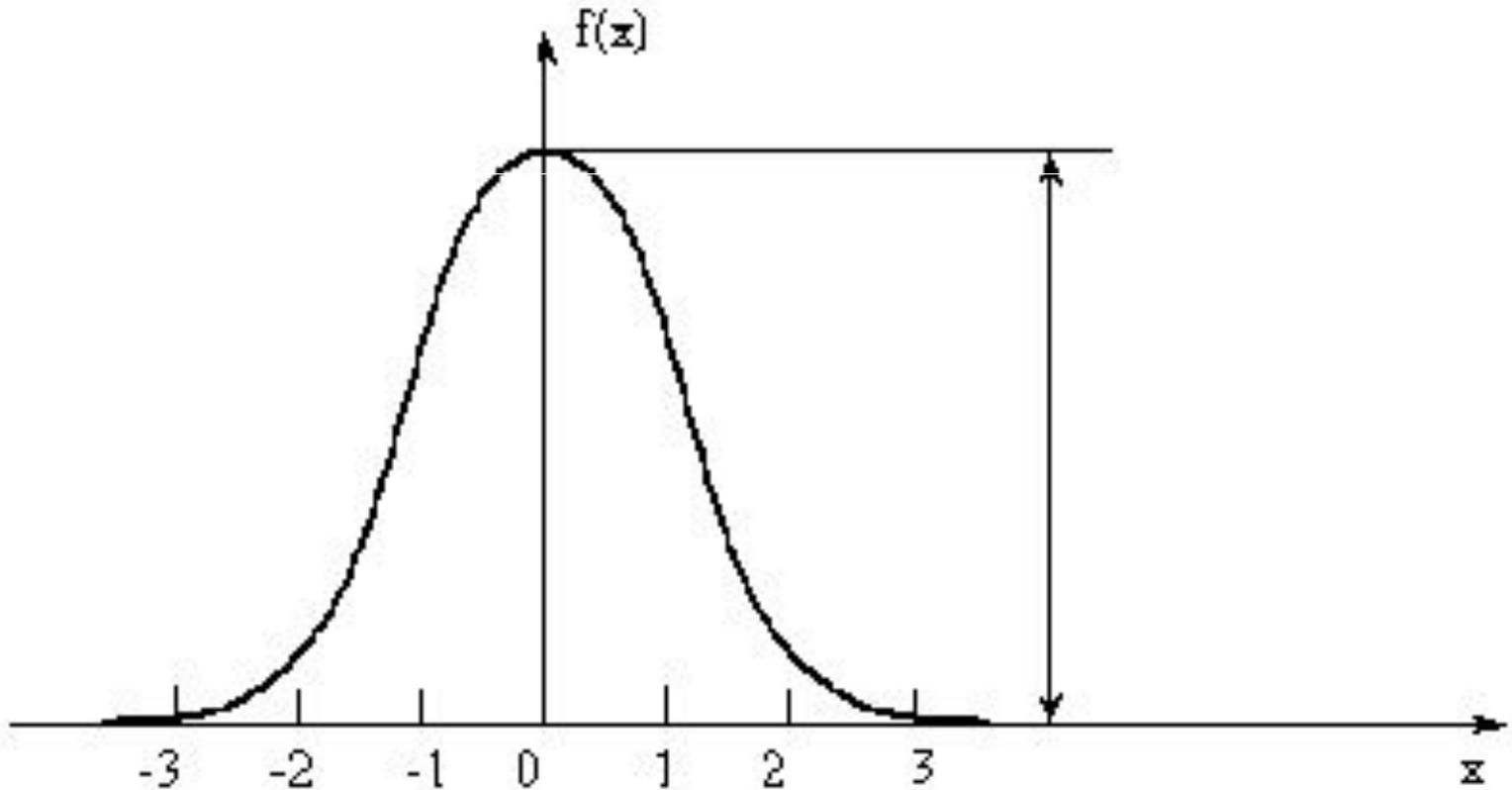


### 3. Распределение Больцмана (распределение частиц по потенциальным энергиям в силовых полях - гравитационном, электрическом).

График - экспонента



## 4. Нормальное распределение (график - кривая Гаусса)



# **5. Распределение Пуассона и др. ...**

**Нормальный закон распределения  
имеет важное практическое  
значение в естественных науках.**

**Оказывается, распределение  
роста, массы новорожденных и  
много других случайных событий  
физической и биологической  
природы описываются  
нормальным законом  
распределения и графически  
иллюстрируются кривой Гаусса.**

**Числовые характеристики**  
**дискретных случайных**  
**величин.**

**1. Математическое ожидание**  
случайной величины есть сумма произведений всех возможных ее значений на вероятности этих значений:

$$M(x) = \sum_{i=1} x_i p_i$$

$$M(x) = x_1 \cdot p_1 + x_2 \cdot p_2 + \dots + x_n \cdot p_n$$

## 2. Среднее арифметическое значение

$$\bar{X} = \frac{m_1 x_1 + m_2 x_2 + m_3 x_3 + m_4 x_4 \dots + \dots m_n x_n}{n}$$

$$= \frac{m_1}{n} x_1 + \dots + \frac{m_n}{n} x_n = p_1 x_1 + \dots + p_n x_n$$

**n – число измерений.**

$$\overline{X} \quad \langle X \rangle$$

Если  $n$  велико, то **относительные частоты**  $m/n = p$ ,  
а **среднее арифметическое значение**  
**практически равно математическому**  
**ожиданию.**

$$\langle X \rangle = M(x)$$

**Математическое ожидание часто**  
**отождествляют со средним значением**

### 3. Дисперсия случайной

**величины** – это математическое ожидание квадрата отклонения случайной величины от ее математического ожидания;

$$D(x) = M(x^2) - [M(x)]^2$$

$$D(x) = M[x - M(x)]^2$$

Дисперсия характеризует  
**рассеяние** случайных  
величин относительно  
математического ожидания.



120

150

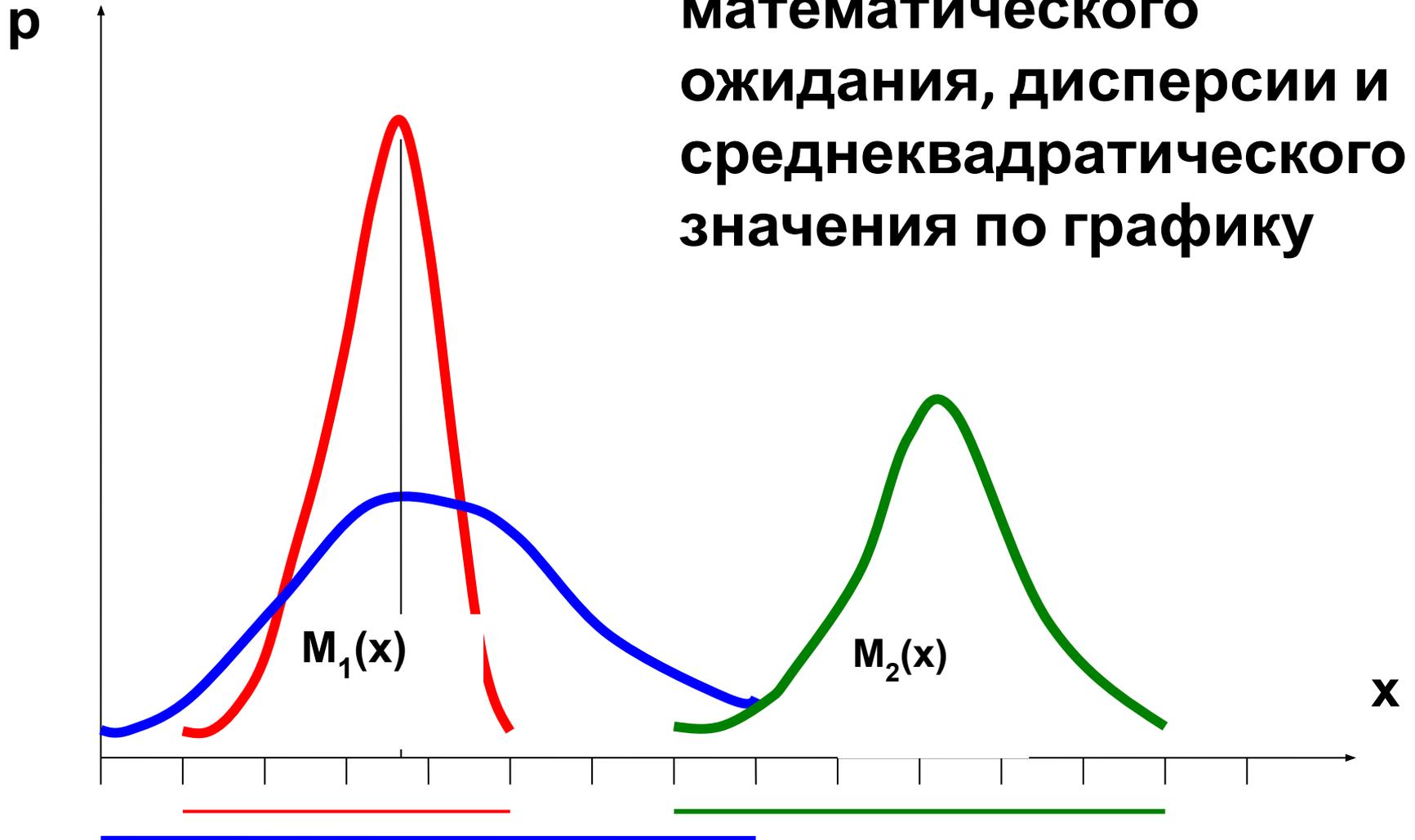
180

**Размерность дисперсии -  
квадрат размерности  
случайной величины, поэтому  
введена величина**

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

**4.  $\sigma$  - среднеквадратическое  
отклонение, которое имеет  
размерность случайной  
величины.**

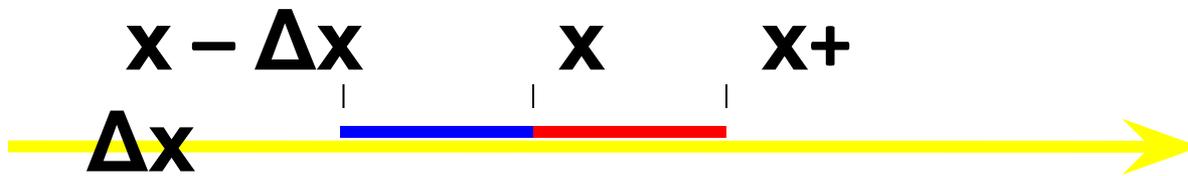
# Сравнительный анализ значений математического ожидания, дисперсии и среднеквадратического значения по графику



**Числовые**  
**характеристики**  
**непрерывных случайных**  
**величин.**

непрерывная случайная величина  
X принимает значения между  $x$  и  
 $x \pm \Delta x$

$$dP = f(x)dx$$



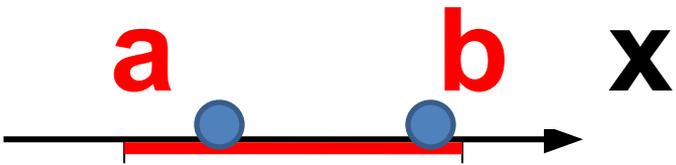
$dP = f(x)dx$ , где  $f(x)$  - **плотность вероятности** или **функция распределения вероятности**.

**функция распределения вероятности**  
**показывает, как изменяется**  
**вероятность, отнесенная к интервалу**  
 **$dx$  случайной величины в**  
**зависимости от значения самой этой**  
**величины:  $f(x) = dP/dx$**

$$P_{ab} = \int_a^b f(x) dx$$

$$P_{ab} = \int_a^b f(x) dx$$

- вероятность того, что случайная величина принимает значения в интервале **(ab)**.





**Какова вероятность того, что случайная величина находится в данном интервале?**

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

**условие  
нормировки для  
непрерывной  
случайной  
величины**

# 1. Математическое ожидание

$M(x)$ :

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

## 2. Дисперсия $D(x)$ :

$$D(x) = \int_{-\infty}^{+\infty} [x - M(x)]^2 \cdot f(x) dx$$

**3.**

**Среднеквадратическое  
отклонение,  
которое имеет размерность  
случайной величины.**

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

# Нормальный закон распределения:

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

exp - экспонента;

$e^{\pm x} = \exp(\pm x)$ ;

**График нормального  
закона - кривая Гаусса.**

Учитывая, что  $\sigma = \sqrt{D}$

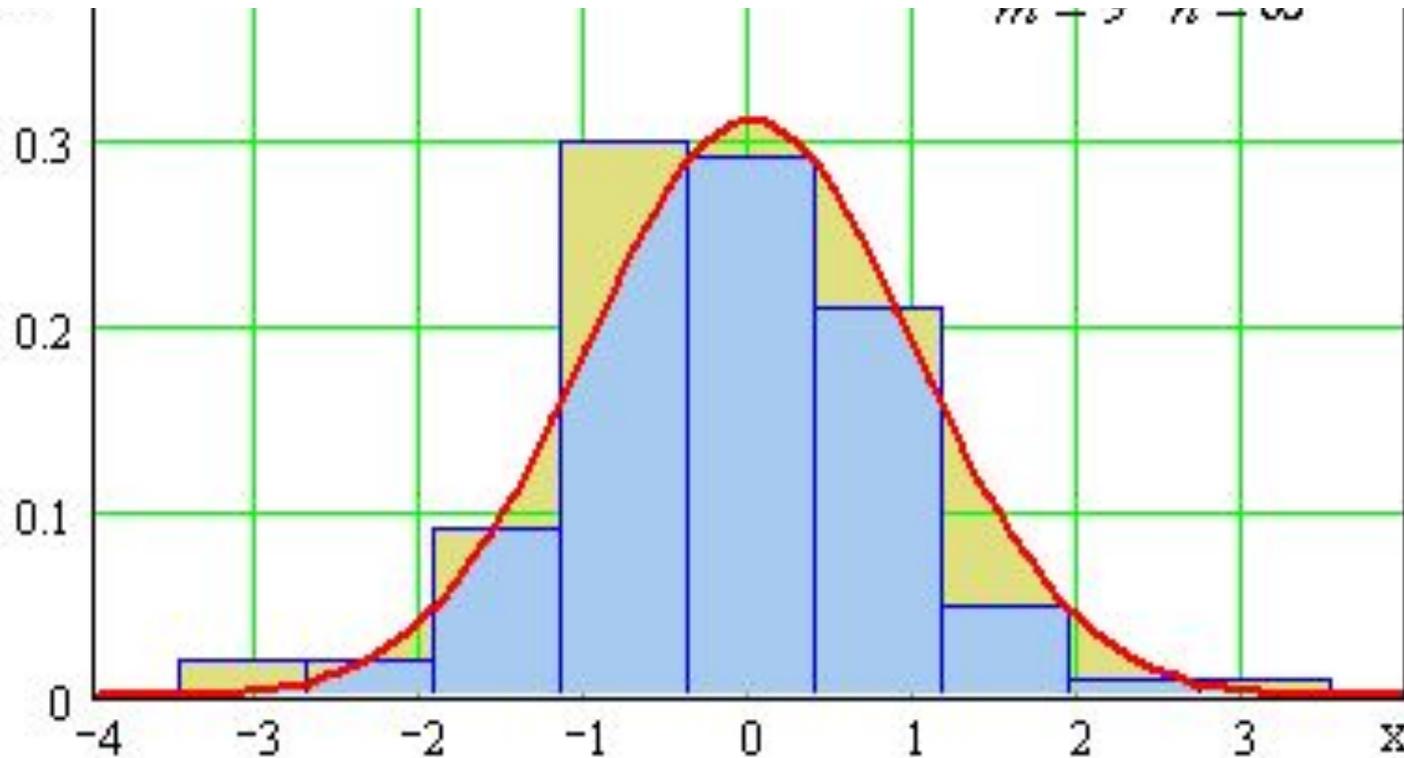
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left[-\frac{[x-M(x)]^2}{2\sigma^2}\right]}$$

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

$$f(x) = \frac{1}{\sqrt{2\pi D}} \exp\left[-\frac{[x - M(x)]^2}{2D}\right]$$

# Особенности кривой Гаусса

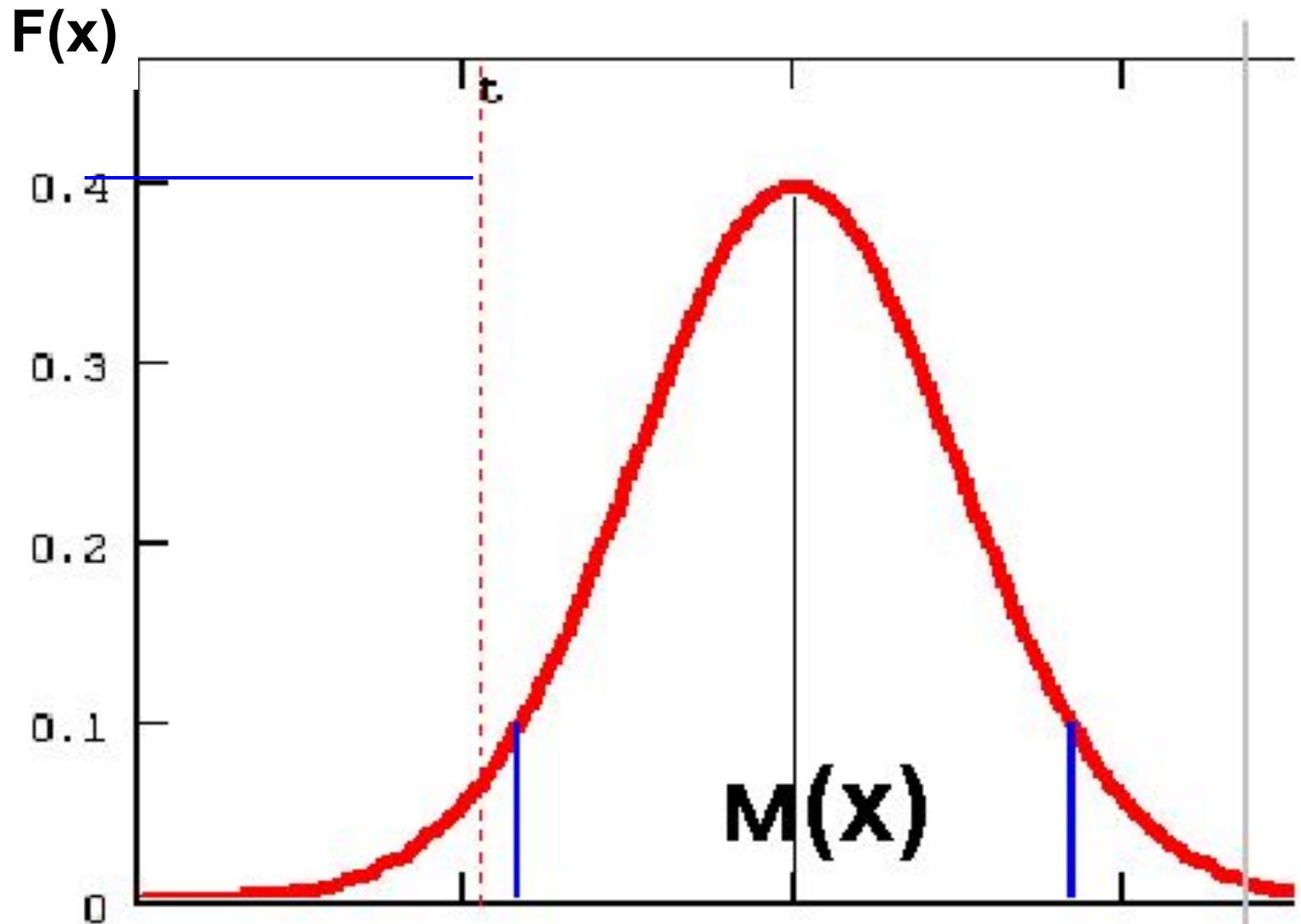
## колоколообразная форма



ветви – экспоненты (**возрастающая** и **убывающая**)

• симметрия относительно  $M(X)=x$ .

$M(X)$  - центр рассеивания



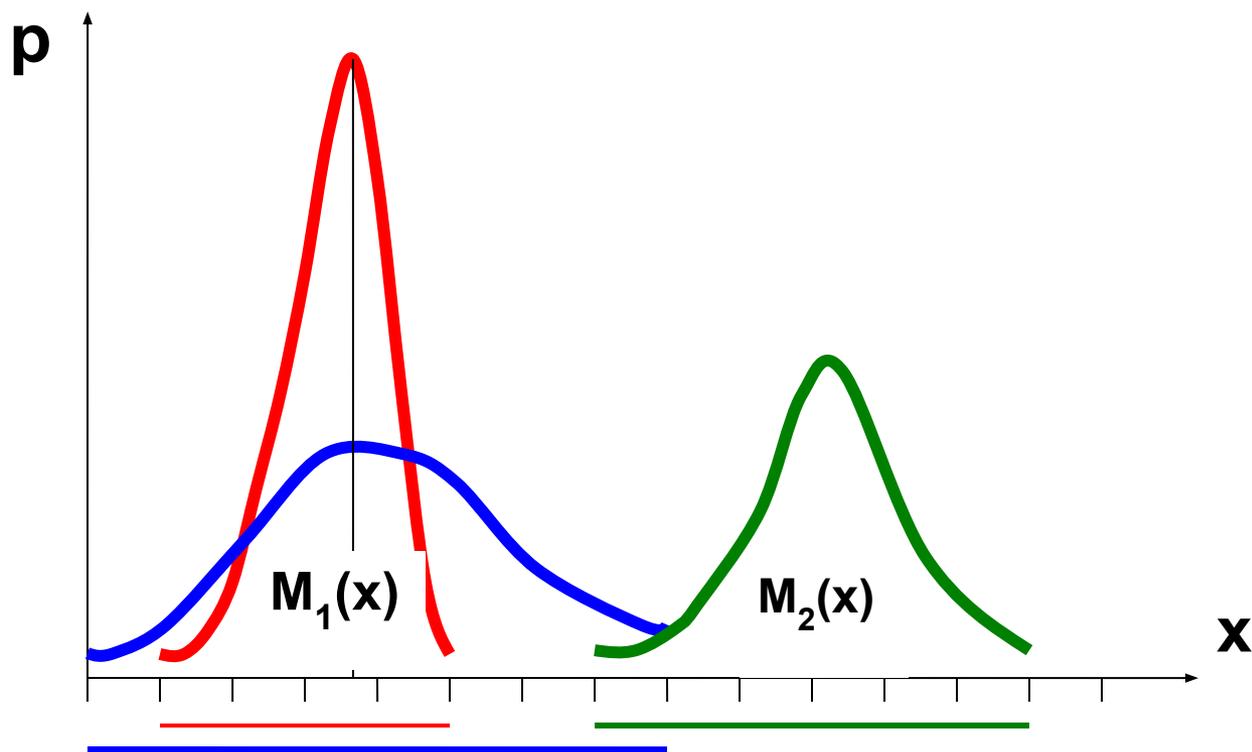
по данной формуле определяем координаты вершины кривой Гаусса, когда  $x = M(x)$ .

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left[-\frac{[x-M(x)]^2}{2\sigma^2}\right]}$$

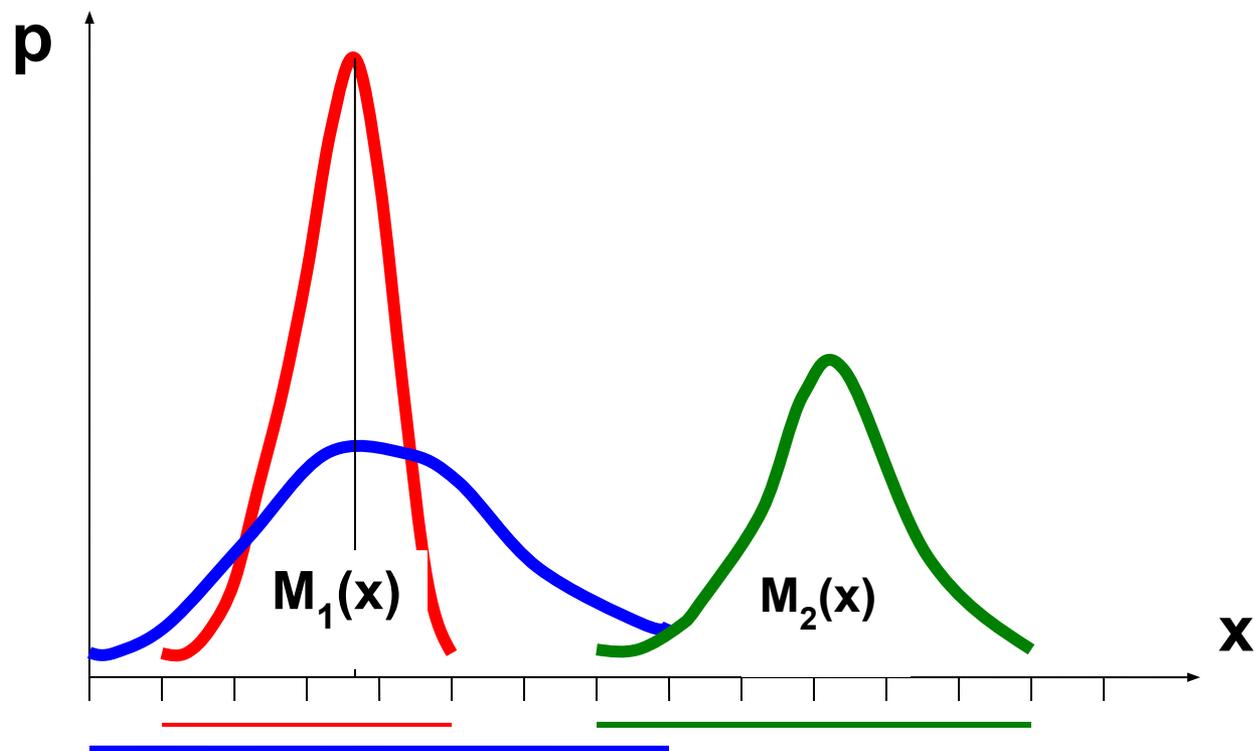
# Вершина графика

$$f(x)_{\max} = \frac{1}{\sigma \sqrt{2\pi}}$$

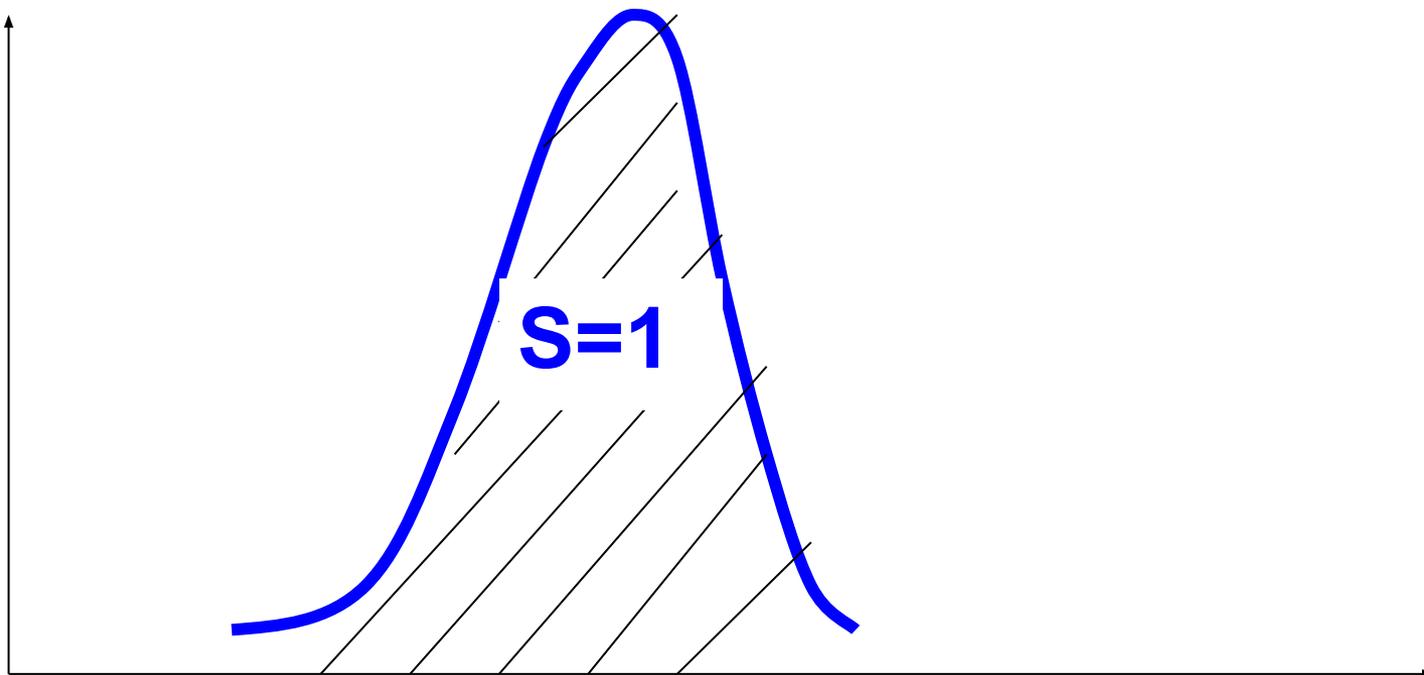
Ветви асимптотически приближаются к оси  $x$ . Чем больше  $\sigma$ , тем **менее острая** вершина.



**Изменение математического ожидания  $M(X)$  сдвигает влево или вправо вершину кривой Гаусса**

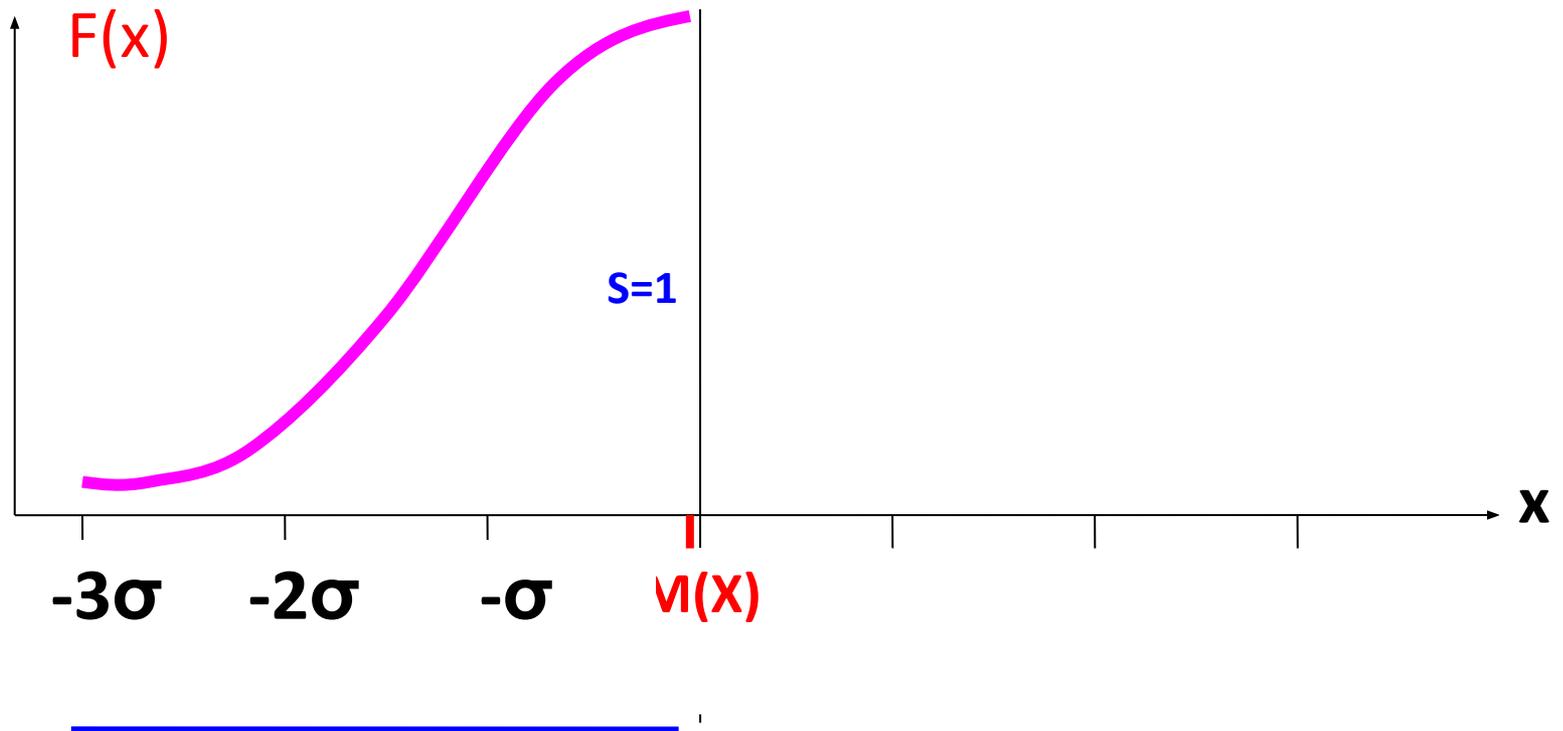


Площадь, заключенная под кривой равна **1** (условие нормировки)

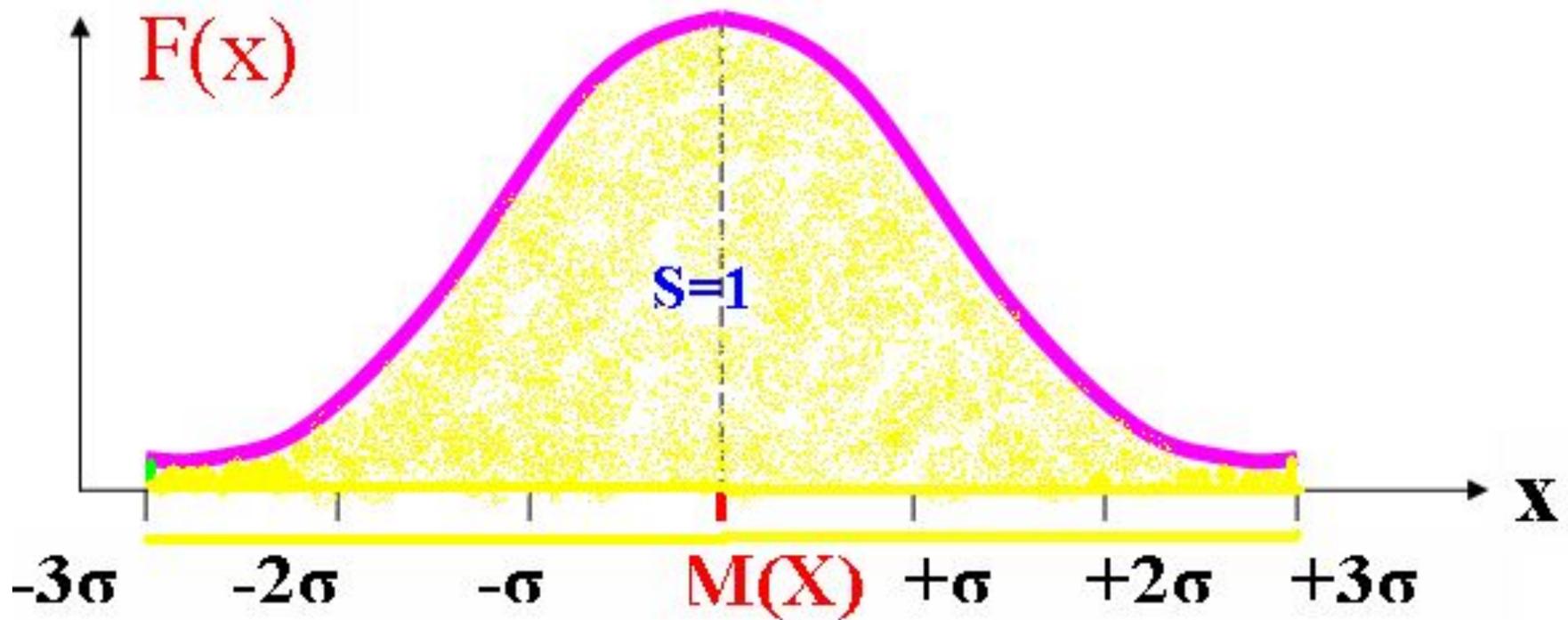


• выполняется правило "трёх сигм".

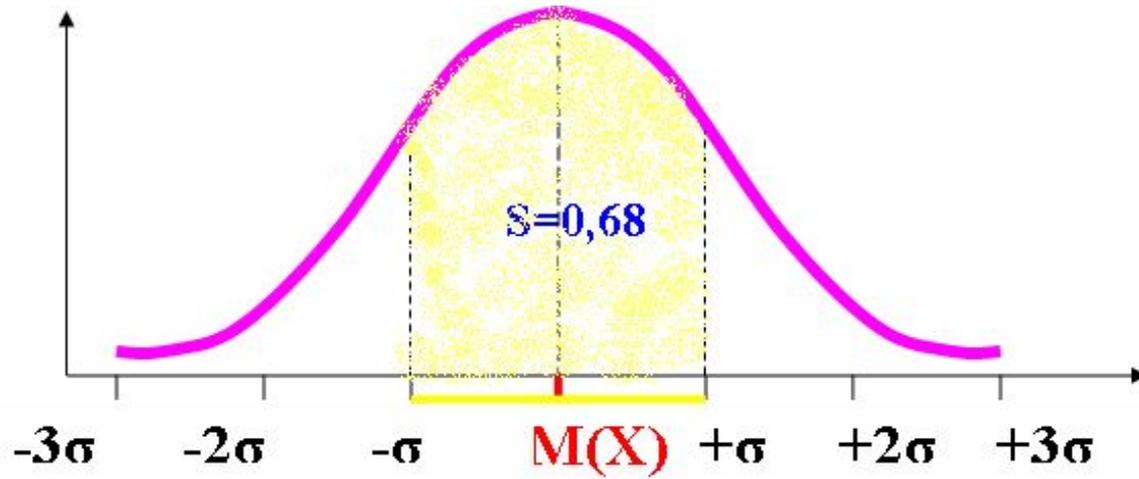
ГРАФИЧЕСКАЯ ИЛЛЮСТРАЦИЯ ПРАВИЛА «**3  $\sigma$** ».



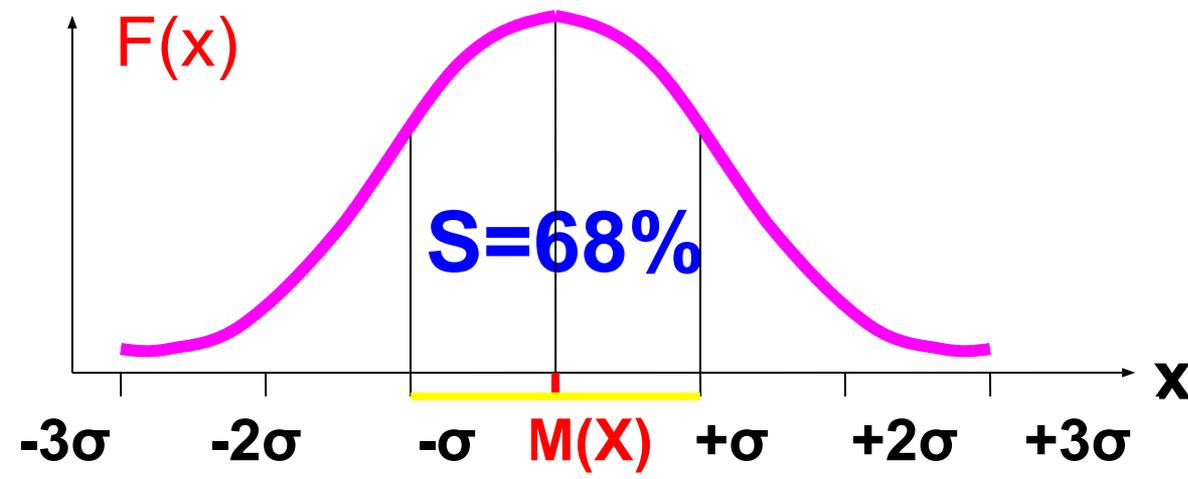
• Вероятность появления случайной величины в интервале значений  $M(X) \pm 3\sigma$  равна 99,97%. Это соответствует условию нормировки - **площадь под кривой равна 1**, Т.Е - практически все случайные величины нормального распределения находятся под кривой Гаусса.



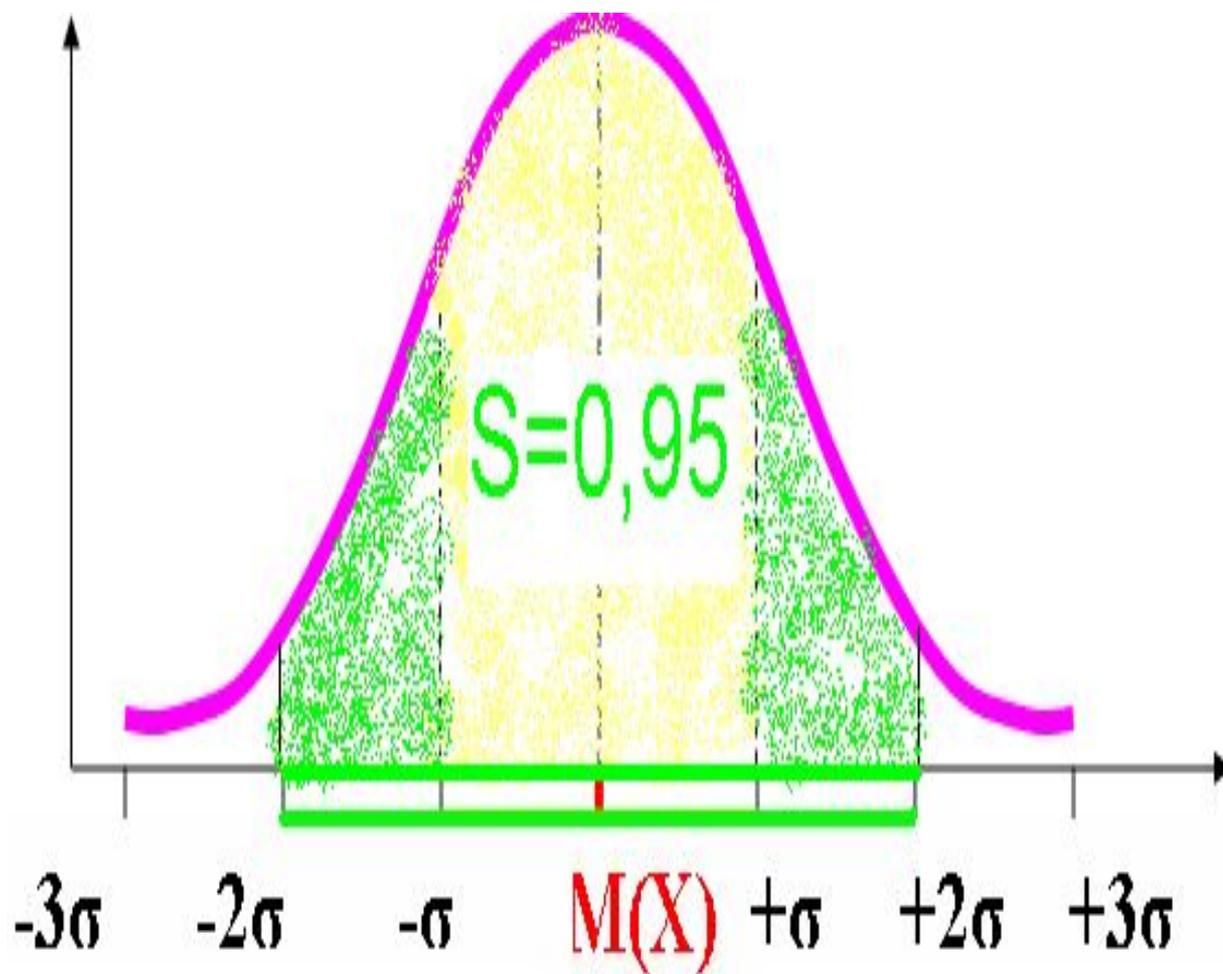
- Вероятность появления случайной величины в интервале значений  $M(x) \pm \sigma$  равна **68%**



**$M(x) \pm \sigma$**



Вероятность появления случайной величины в интервале значений  $M(x) \pm 2\sigma$  равна **95%**

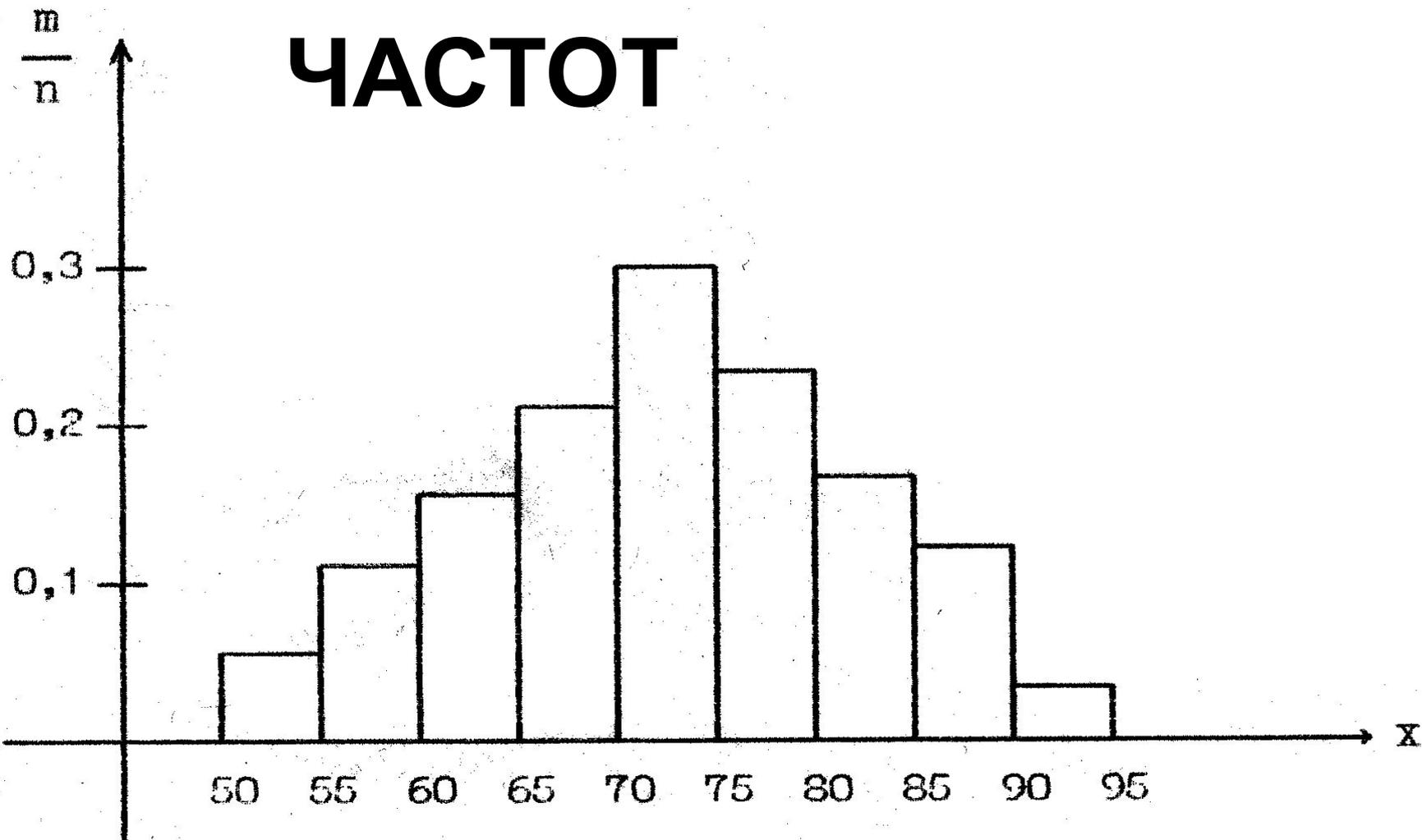


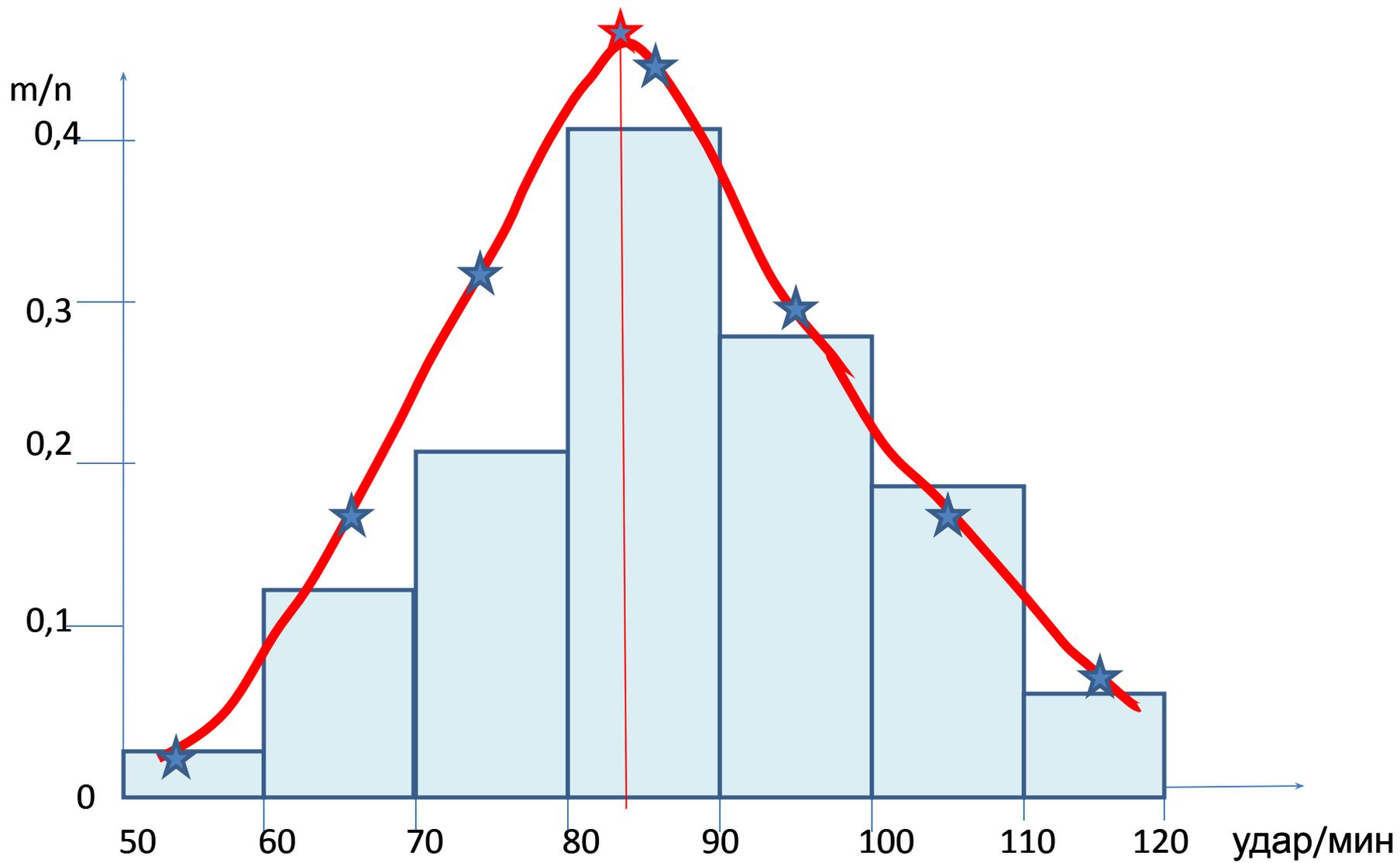
$M(x) \pm 2\sigma$

Для нормального закона распределения характерен **симметричный** вид гистограммы

- **Гистограмма частот** - совокупность смежных прямоугольников, построенных на одной прямой линии. Основания прямоугольников одинаковы. Высоты прямоугольников равны относительной частоте  $m/n$  (вероятности).

# ГИСТОГРАММА ЧАСТОТ





# Гистограмма частот и кривая Гаусса

