

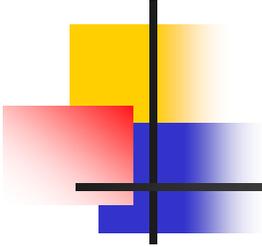
ОБРАБОТКА И АНАЛИЗ ЧИСЛОВОЙ ИНФОРМАЦИИ

ЦЕЛИ И ЗАДАЧИ

2018 © Левандовский С.А.,
профессор, канд. техн. наук

(на основе материалов профессора Румянцева М.И.)

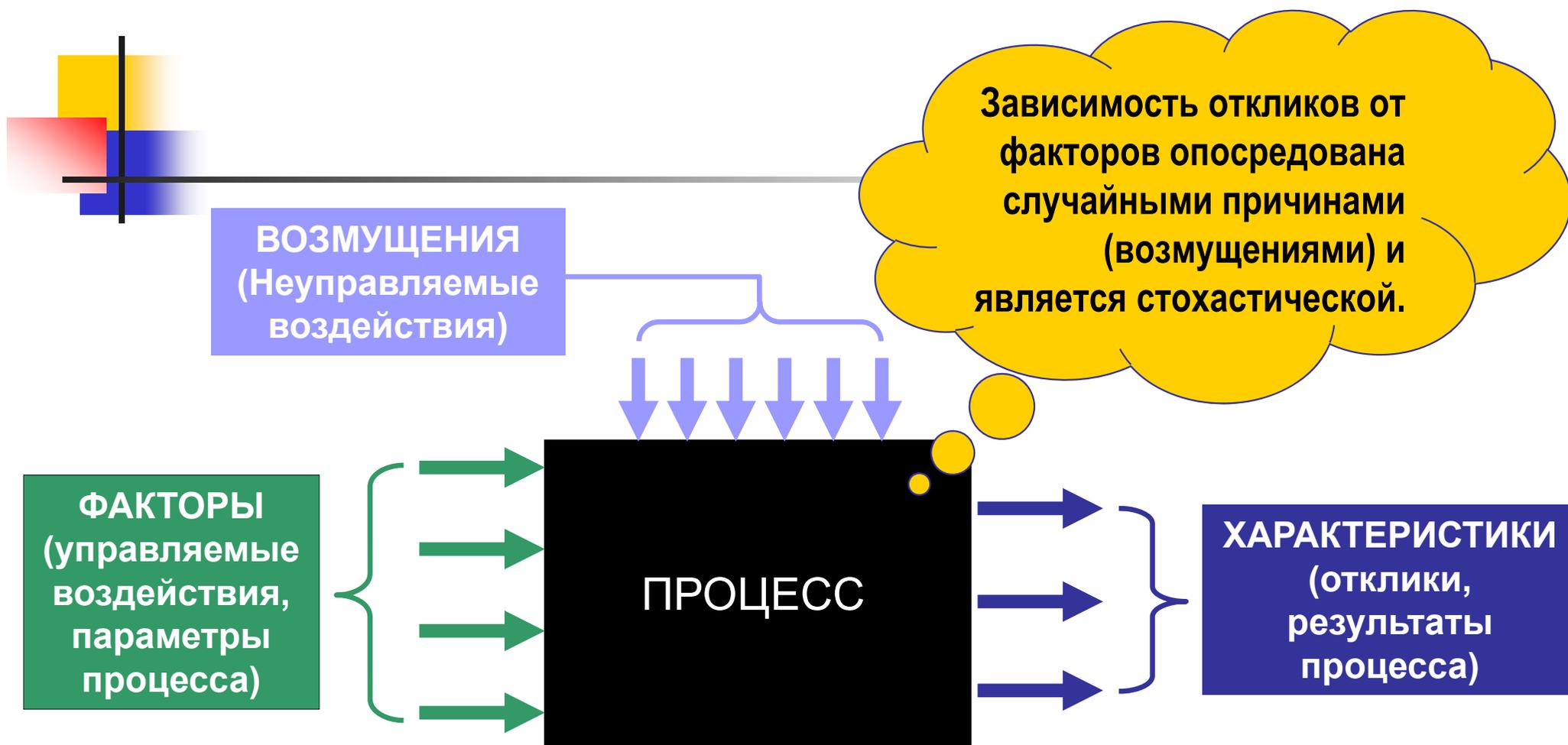
Кафедра технологий обработки материалов, ФГБОУ ВО МГТУ



ЦЕЛЬ ОБРАБОТКИ И АНАЛИЗА ЧИСЛОВОЙ ИНФОРМАЦИИ

Сделать надежные выводы о функционировании того или иного объекта на основании массива числовых данных, полученных в результате наблюдений за ним

СТОХАСТИЧНОСТЬ



Стохастичность зависимости проявляется, например, в том, что при одних и тех же значениях факторов в различные моменты времени будут обнаружены различные значения отклика

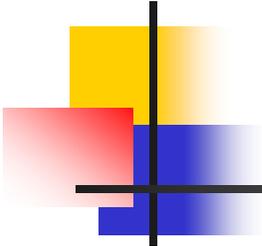
СЛУЧАЙНЫЙ ХАРАКТЕР ЧИСЛОВОЙ ИНФОРМАЦИИ



Из-за наличия возмущений, числа в массиве данных, на основании которых необходимо делать выводы об объекте, представляют собой значения случайных величин.

Другие причины случайного характера числовой информации

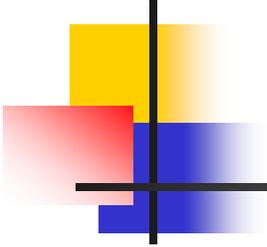
-  Погрешности измерения анализируемых параметров
-  Ограниченная точность работы исполнительных механизмов, с помощью которых задаются и поддерживаются требуемые значения факторов.



СЛУЧАЙНАЯ ВЕЛИЧИНА

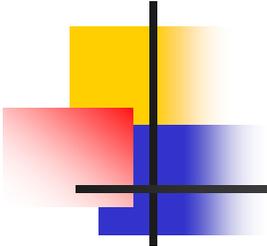
- ✓ Величина, которая в результате опыта может принять то или иное заранее неизвестное значение
- ✓ Величины, которые могут принимать в результате опыта различные значения, причем до опыта невозможно предвидеть, какими именно они будут
- ✓ Величина, измеряемая в исследуемых экспериментах, исходы которых заранее не известны и зависят от случайных причин
- ✓ Величина, принимающая в зависимости от случая те или иные значения с определёнными вероятностями.

Всякая случайная величина X обладает тем свойством, что при производстве в неизменных условиях n наблюдений будут зарегистрированы неодинаковые значения x_1, x_2, \dots, x_n среди которых, однако, могут встречаться и повторяющиеся.



ТИПЫ СЛУЧАЙНЫХ ВЕЛИЧИН

- ДИСКРЕТНАЯ
- НЕПРЕРЫВНАЯ

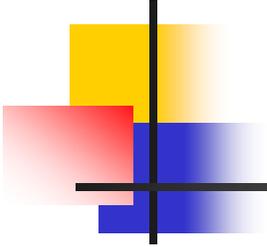


ДИСКРЕТНАЯ СЛУЧАЙНАЯ ВЕЛИЧИНА

Принимает лишь отдельные, изолированные одно от другого значения.

Таким свойством обладают атрибутивные признаки (цвет предмета, сорт продукции, годное или бракованное изделие и т. д.).

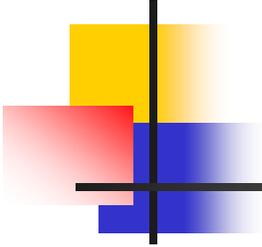
Отдельные значения подобных параметров определяются путем счета и поэтому значениями дискретной случайной величины могут быть только натуральные числа.



НЕПРЕРЫВНАЯ СЛУЧАЙНАЯ ВЕЛИЧИНА

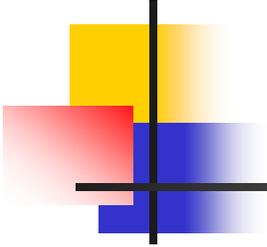
Принимает любые значения из некоторого свойственного ей числового интервала.

Таким свойством обладают количественные признаки (механические свойства материала, фактические размеры продукции, производительность агрегата при обработке конкретного профилеразмера и т. п.) Отдельные значения таких параметров определяются путем измерений или расчетов и поэтому могут быть любыми действительными числами.



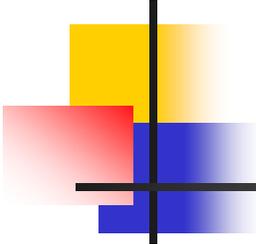
СВОЙСТВА СЛУЧАЙНОЙ ВЕЛИЧИНЫ

- **ИСТИННОЕ ЗНАЧЕНИЕ.**
Случайная величина имеет такое единственное значение, к которому тяготеют все остальные.
Характеристика – *математическое ожидание*
- **ВАРИАЦИЯ.**
Наблюдаемые значения случайной величины рассеяны относительно математического ожидания (отклоняются от него) в соответствии с особенностями, объективно присущими данной случайной величине.
Характеристики:
 - *дисперсия*
 - *стандартное отклонение (стандарт)*



ЗАКОНОМЕРНОСТИ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

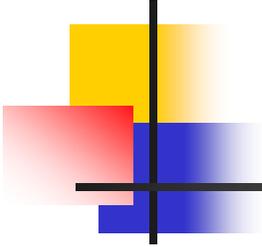
- ОГРАНИЧЕННЫЙ ИНТЕРВАЛ
ВАРЬИРОВАНИЯ.
- ЗАКОН РАСПРЕДЕЛЕНИЯ.



ОГРАНИЧЕННОСТЬ ИНТЕРВАЛА ВАРЬИРОВАНИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Случайная величина может принимать значения только из интервала, величина и положение которого на числовой оси обусловлены ее физической природой.

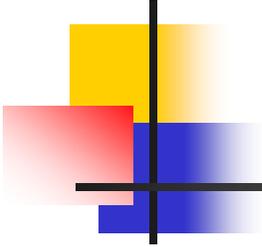
Случайность проявляется лишь в том, что значение, которое будет обнаружено в некоторый момент наблюдения, заранее не известно.



ЗАКОН РАСПРЕДЕЛЕНИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Каждое возможное значение случайной величины проявляется внутри интервала ее варьирования с совершенно определенной вероятностью.

Совокупность возможных значений случайной величины и соответствующих им вероятностей появления называются законом распределения вероятности (законом распределения, распределением) данной случайной величины



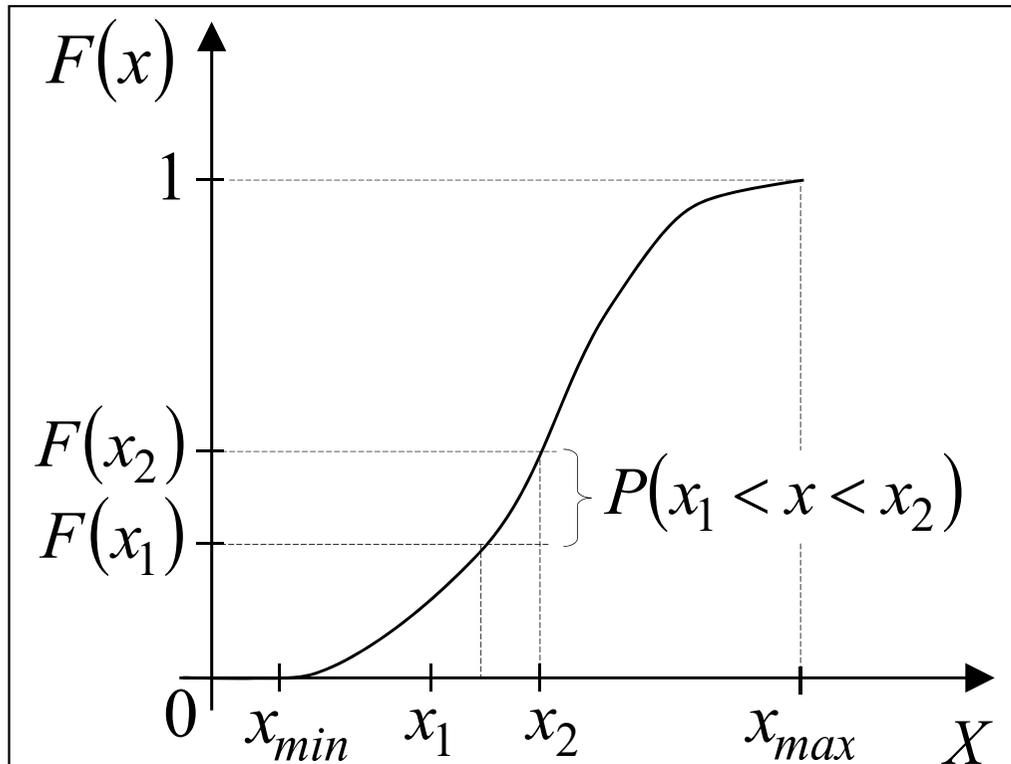
ФОРМЫ ОТОБРАЖЕНИЯ РАСПРЕДЕЛЕНИЯ

- ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ
(Density Function)
- ПЛОТНОСТЬ РАСПРЕДЕЛЕНИЯ
(Distribution Function)

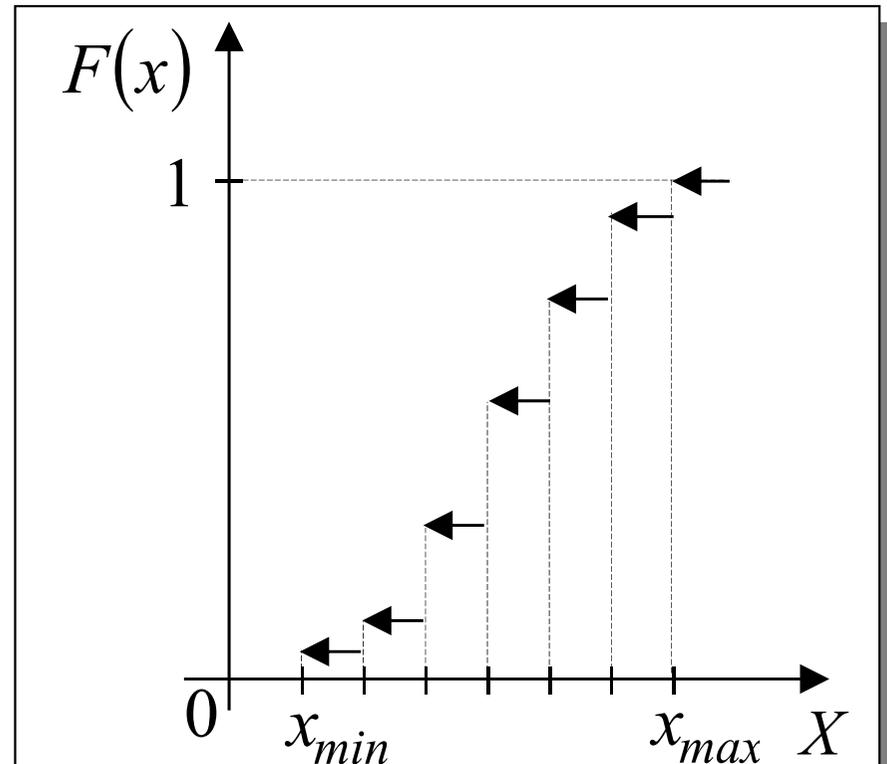
ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ

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

Отображает вероятность события, заключающегося в том, что случайная величина (например X) примет значение меньше, чем произвольное действительное число x



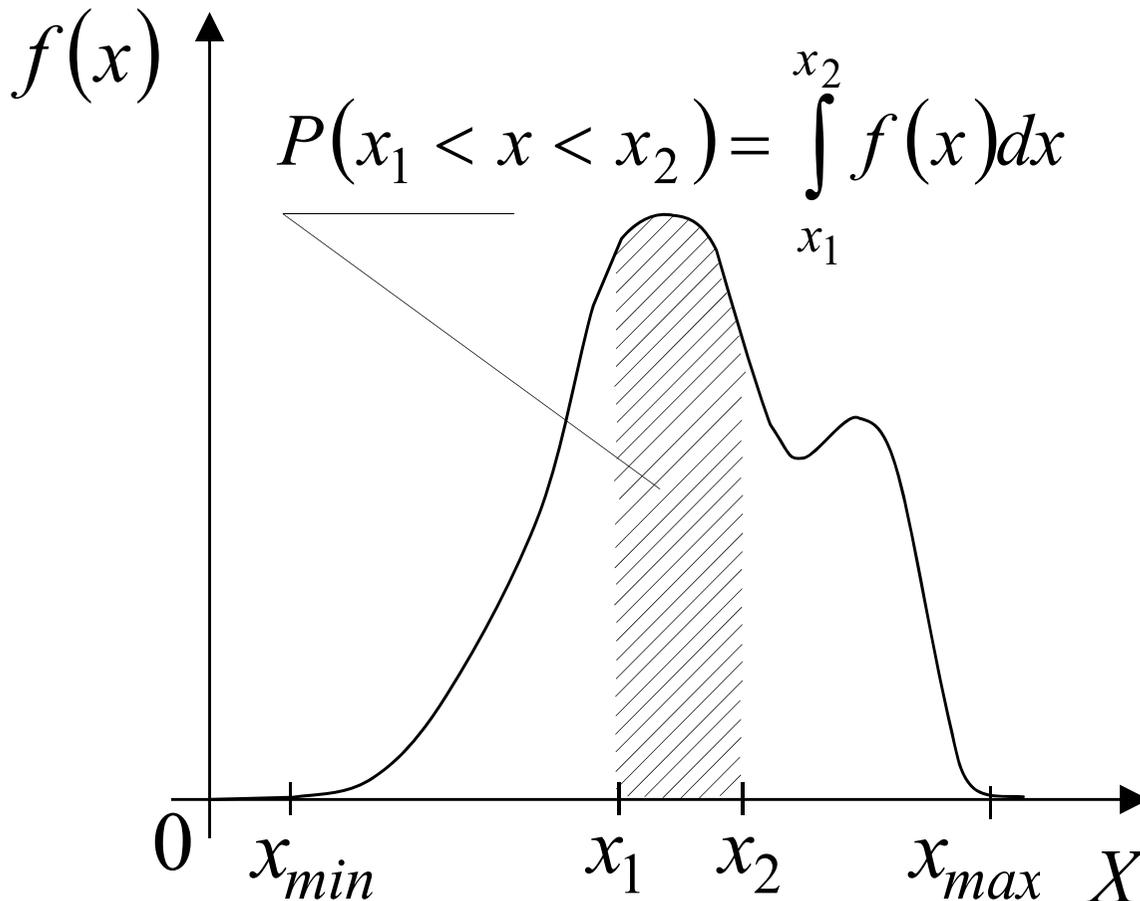
Для непрерывной величины



Для дискретной величины

ПЛОТНОСТЬ РАСПРЕДЕЛЕНИЯ

$$f(x) = P(x_1 < X < x_2)$$



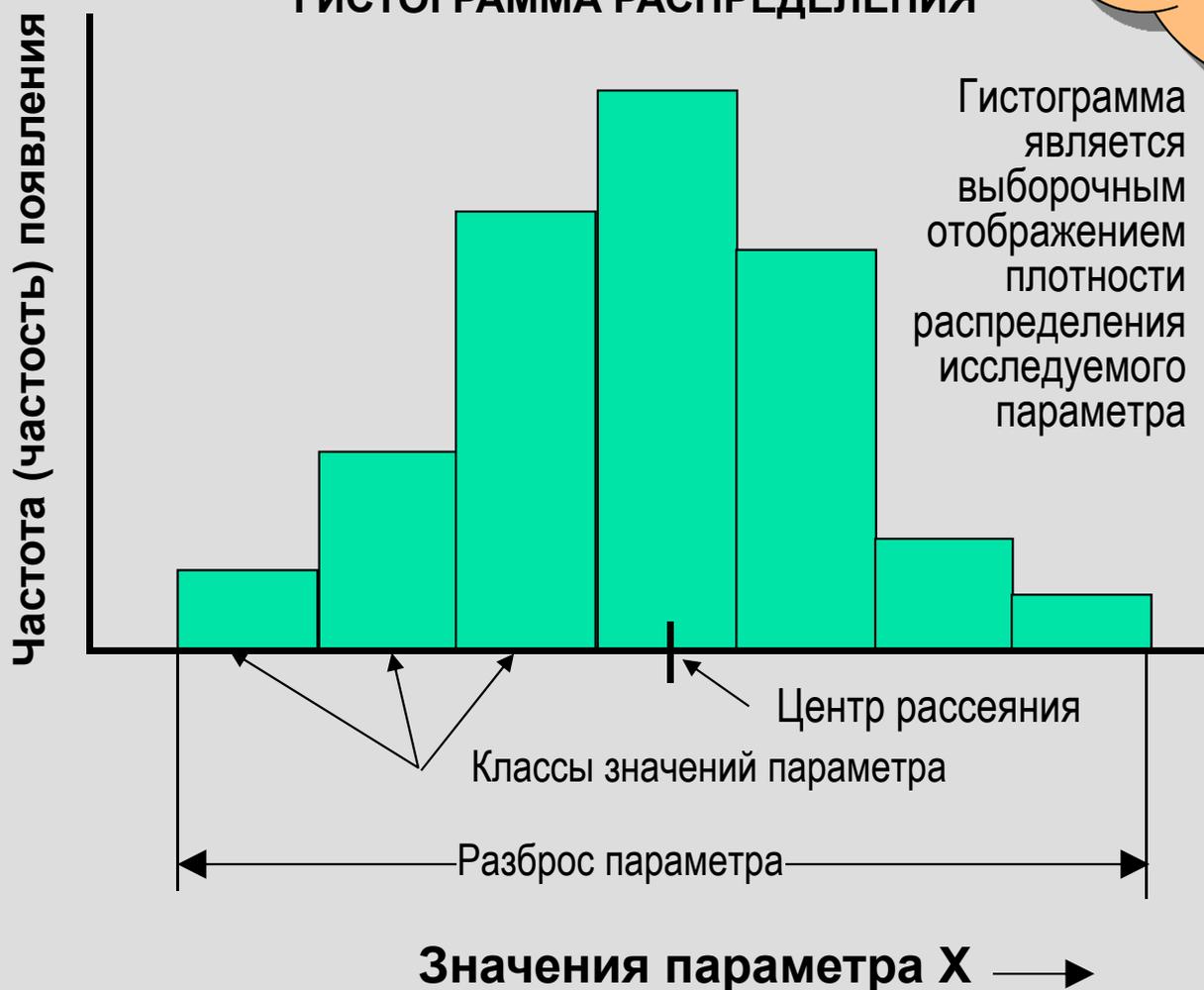
Отображает вероятность события, состоящего в том, что произвольное значение x случайной величины X находится в некотором наперед заданном интервале $\{x_1; x_2\}$.

Если функция распределения данной случайной величины непрерывна и дифференцируема, то

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

ВЫБОРОЧНЫЙ МЕТОД ИЗУЧЕНИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

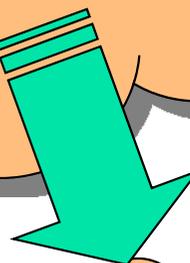
ГИСТОГРАММА РАСПРЕДЕЛЕНИЯ



$\{\chi_i\}_N$

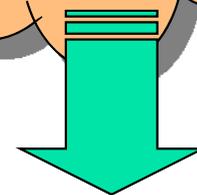
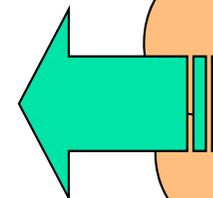
ГЕНЕРАЛЬНАЯ СОВОКУПНОСТЬ

полное множество возможных значений параметра в соответствии с его физической природой и особенностями процесса



ВЫБОРКА

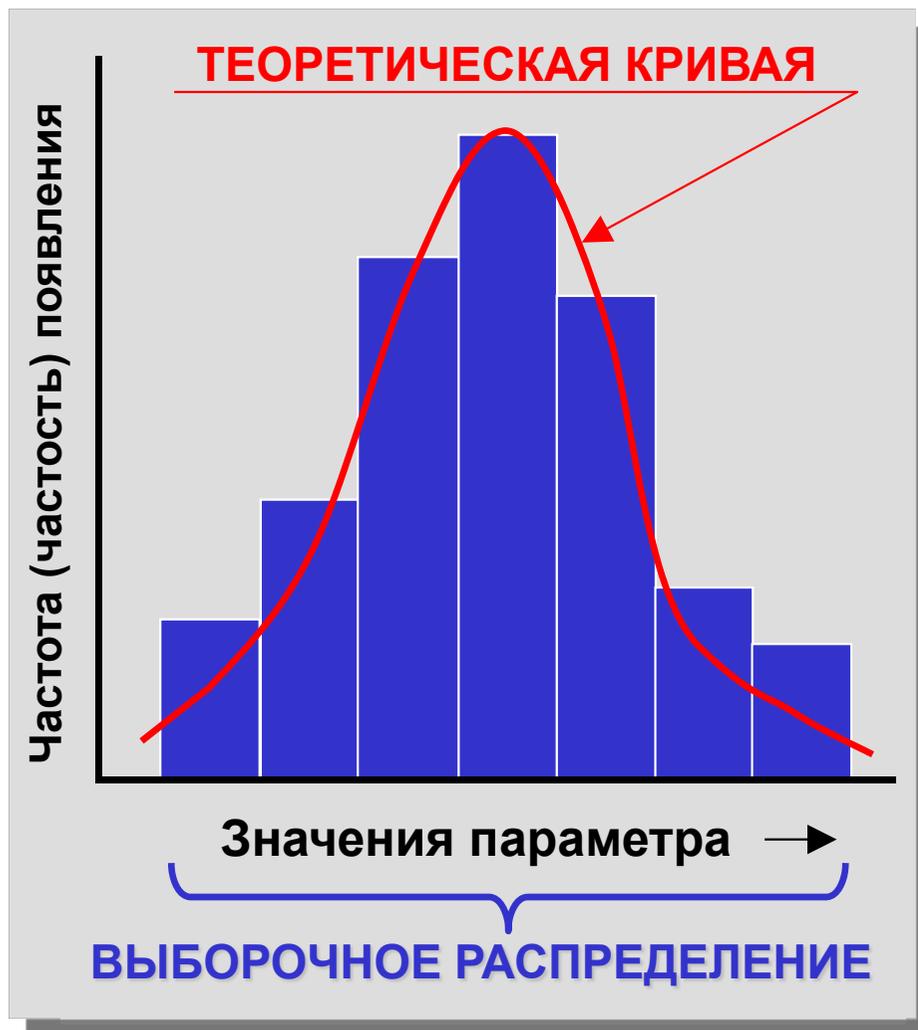
ограниченное множество значений, $\{x_i\}_n$ обнаруженных за время наблюдения



**ОПИСАТЕЛЬНЫЕ
СТАТИСТИКИ**

ОПИСАТЕЛЬНЫЕ СТАТИСТИКИ

Количественные
оценки характеристик
исследуемого
параметра с учетом его
стохастичности



- Характеристики положения:
 - Среднее выборочное
 - Мода
 - Медиана
 - ...
- Характеристики рассеяния (вариации):
 - Размах (интервал)
 - Дисперсия
 - Стандартное отклонение
 - ...
- Закон распределения (теоретическая кривая)

НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Распределение
непрерывной
случайной
величины, для
которого характерна
плотность
распределения вида

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

M_x – математическое ожидание (характеристика положения истинного значения случайной величины). Выборочная оценка – *среднее выборочное*:

$$M_x \approx \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

σ – стандартное отклонение (характеристика вариации случайной величины). Выборочная оценка – *выборочное стандартное отклонение*:

$$\sigma \approx s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

ОСОБЕННОСТИ ПЛОТНОСТИ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ

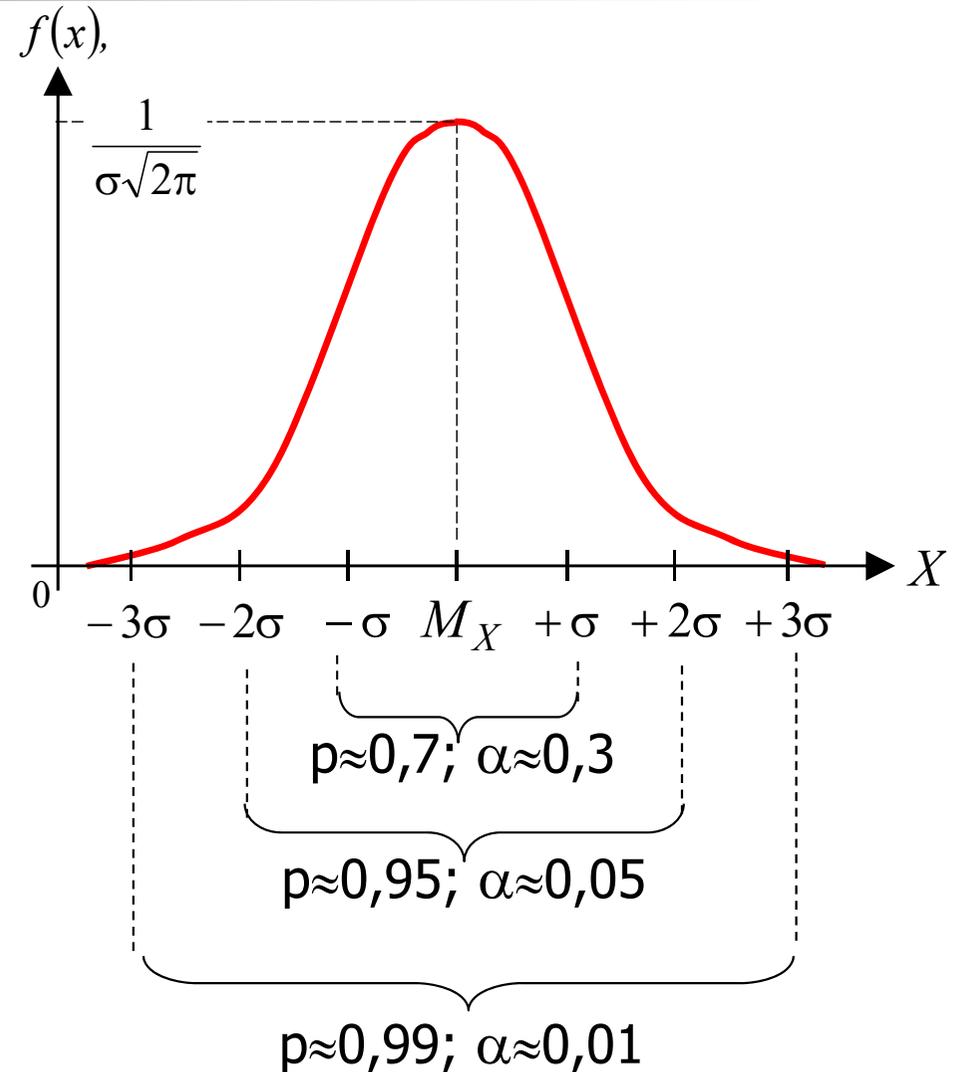
1. Кривая плотности распределения симметрична относительно M_X и при этой абсциссе достигает максимума, который равен

$$f_{max} = \frac{1}{\sigma\sqrt{2\pi}} = \frac{0,3989}{\sigma}$$

2. Если *независимые* случайные величины X_1 и X_2 имеют нормальные распределения, то для произвольных чисел α и β величина $Y = \alpha X_1 + \beta X_2$ также имеет нормальное распределение, причем

$$M_Y = \alpha M_{X_1} + \beta M_{X_2}$$

$$\sigma_Y = \sqrt{\alpha^2 \sigma_{X_1}^2 + \beta^2 \sigma_{X_2}^2}$$



СТАНДАРТНОЕ НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Распределение *непрерывной* случайной величины, для которого характерна плотность распределения вида

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

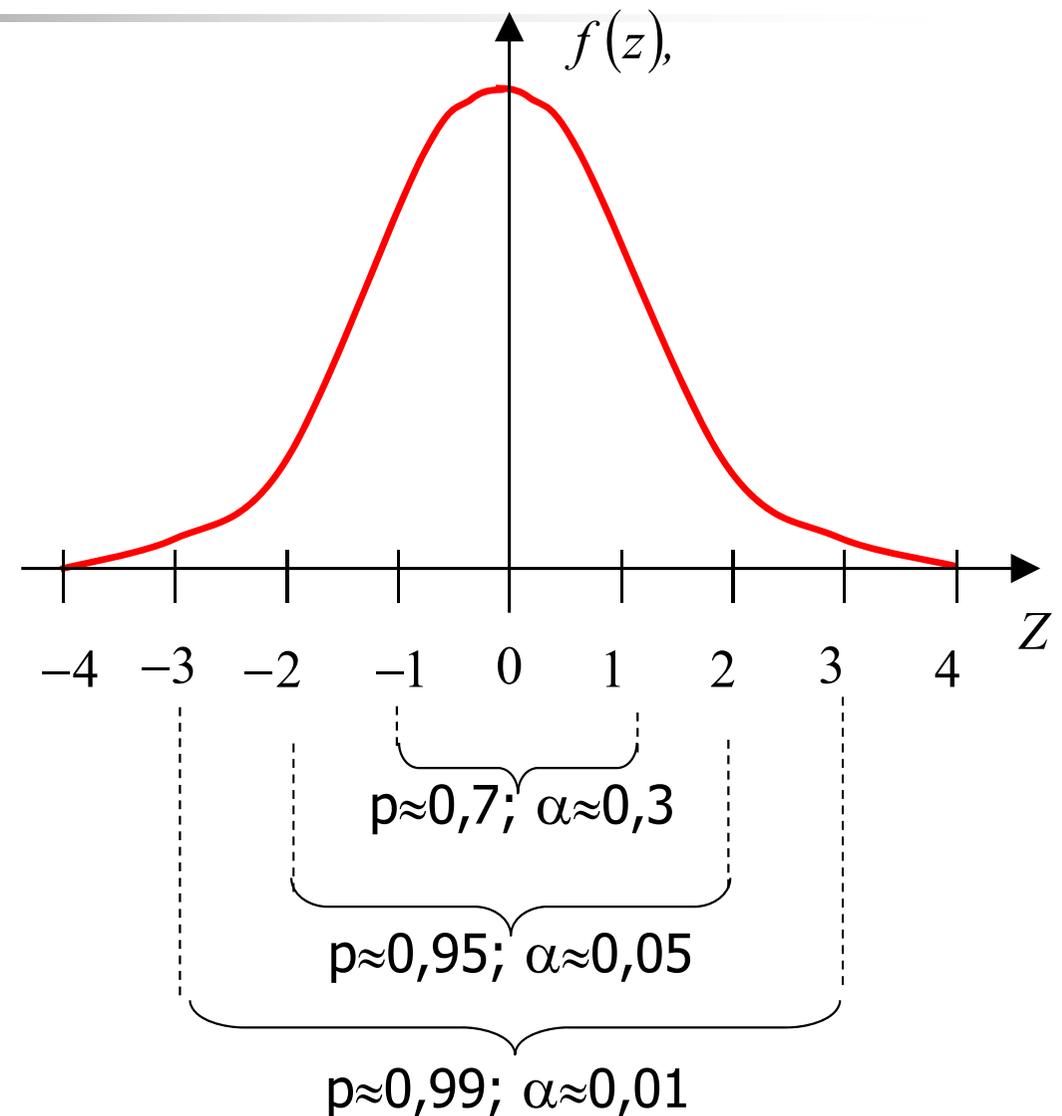
Кривая Гаусса

z – нормированное значение случайной величины:

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

ОСОБЕННОСТИ СТАНДАРТНОГО НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ

1. Плотность стандартного нормального распределения симметрична относительно оси ординат и имеет максимум, равный 0,3989.
2. Плотность стандартного нормального распределения является четной функцией, т.е. $f(z) = f(-z)$.
3. При $|z| = 4$ плотность стандартного нормального распределения равна нулю: $f(4) = f(-4) = 0$.
4. Значение функции ненормированного нормального распределения равно значению функции стандартного нормального распределения.



КАК ПОЛУЧИТЬ НАДЕЖНЫЕ ВЫВОДЫ ПРИ НАЛИЧИИ СТОХАСТИЧНОСТИ

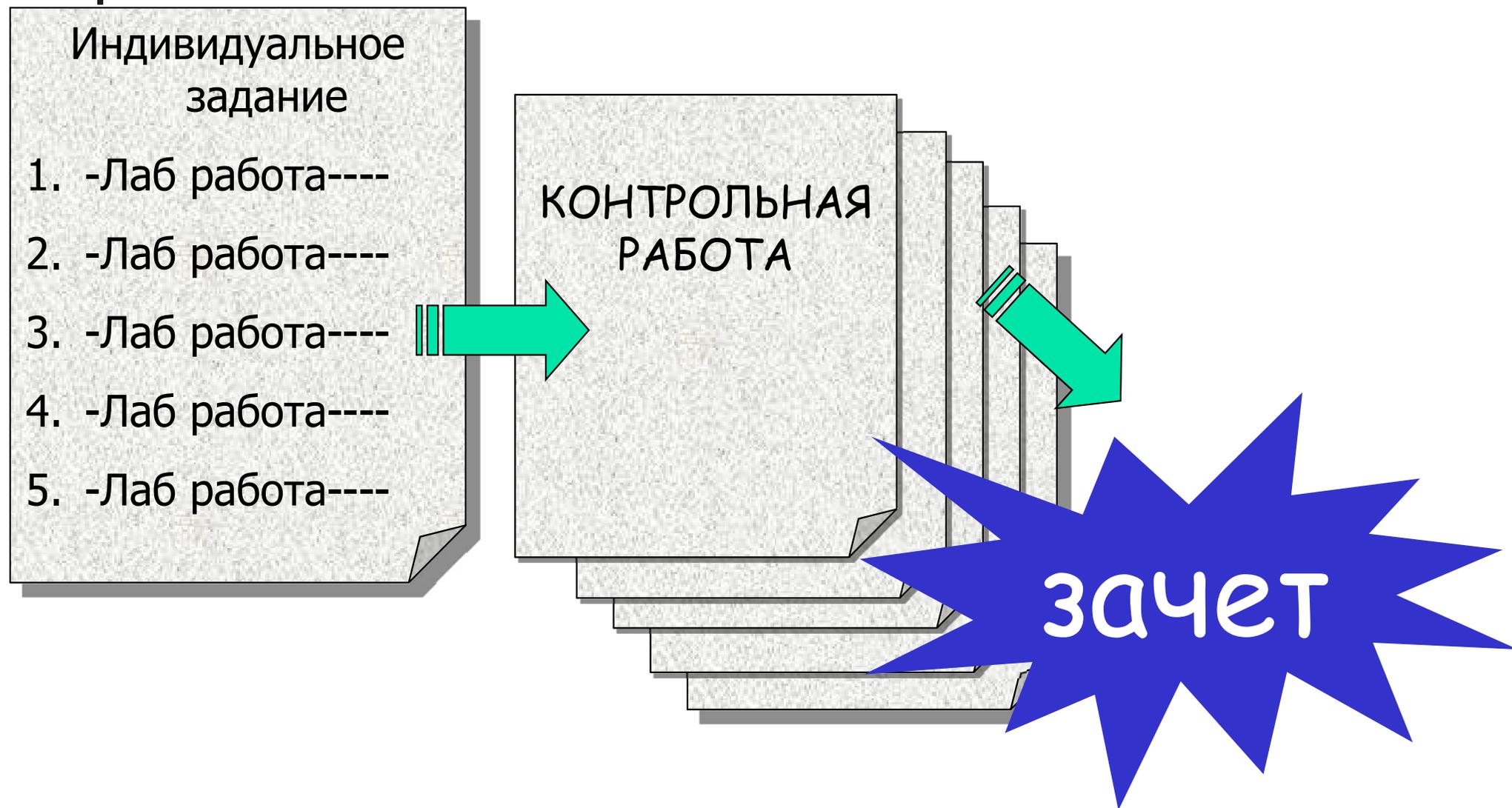
Для получения
надежных выводов
на основании числовой
информации,
которая содержит
случайную
составляющую,
необходимо применять
статистические методы

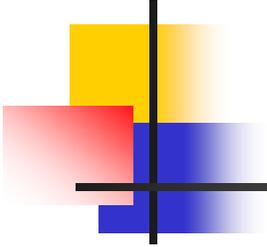


ОСНОВНЫЕ ЗАДАЧИ И МЕТОДЫ ОБРАБОТКИ И АНАЛИЗА ЧИСЛОВОЙ ИНФОРМАЦИИ



ОБЪЕМ САМОСТОЯТЕЛЬНОЙ РАБОТЫ И АТТЕСТАЦИЯ





СОДЕРЖАНИЕ КОНТРОЛЬНОЙ РАБОТЫ

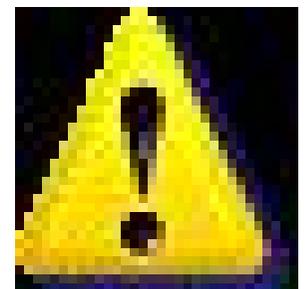
1. Обработка и анализ выборки
2. Построение выборочного распределения
3. Корреляционный анализ
4. Парный регрессионный анализ
5. Множественный регрессионный анализ

РЕКОМЕНДУЕМОЕ СОДЕРЖАНИЕ ОТЧЕТА

По каждой из задач приводятся:

- Название задачи
- Цель решения
- Теоретическое введение
- Результаты решения в виде распечатки
- Выводы

ОТ
РУКИ



РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

1. Румянцев М.И., Левандовский С.А., Ручинская Н.А., Черкасов К.Е., Логинов А.В. Статистические методы обработки и анализа числовой информации, контроля и управления качеством проката: Учебное пособие, 2-е изд., перераб. и доп. Магнитогорск, МГТУ, 2014. 262 с.
2. Львовский Е.Н. Статистические методы построения эмпирических формул: Учебн. пособие для вузов, 2-е изд., перераб. и доп. М., Высш. шк. , 1988. 239 с.
3. Минько А.А. Статистический анализ в MS Excel. М., Изд. дом «Вильямс», 2004. 448 с.
4. Румянцев М.И. Обработка и анализ выборки. Магнитогорск, МГТУ, 2003.
5. Румянцев М.И. Построение выборочного распределения. Магнитогорск, МГТУ, 2003.
6. Румянцев М.И. Корреляционный анализ. Магнитогорск, МГТУ, 2003.
7. Румянцев М.И. Парный регрессионный анализ. Магнитогорск, МГТУ, 2003.
8. Румянцев М.И. Множественный регрессионный анализ. Магнитогорск, МГТУ, 2003.