

Критерии для  
проверки  
данных на  
соответствие  
нормальному  
закону  
распределения

Кустова А.А.

- **Распределение случайной величины** - это закон, описывающий область значений случайной величины и соответствующие вероятности появления этих значений.
- **Нормальный закон распределения** – закон вида:

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

$\mu$       Математическое ожидание  
 $\sigma$       Стандартное отклонение

# Свойства:

1. Полностью определяется средним и дисперсией
2. Среднее и медиана равны
3. Кривая симметрична относительно среднего (рис.7.2. а)
4. Сдвигается вправо, если среднее увеличивается, влево – если среднее уменьшается (рис.7.2. б)
5. Сплющивается, если дисперсия увеличивается, становится более остроконечной – дисперсия уменьшается (рис.7.2. в)
6. Пр

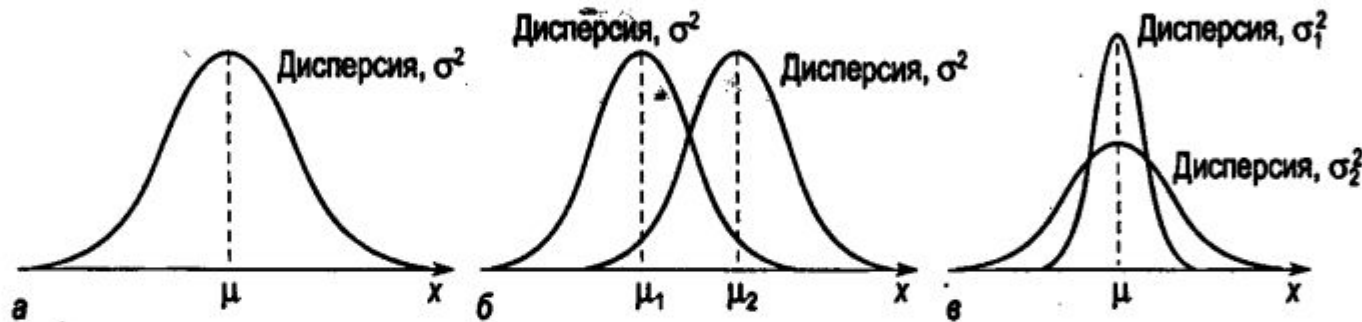
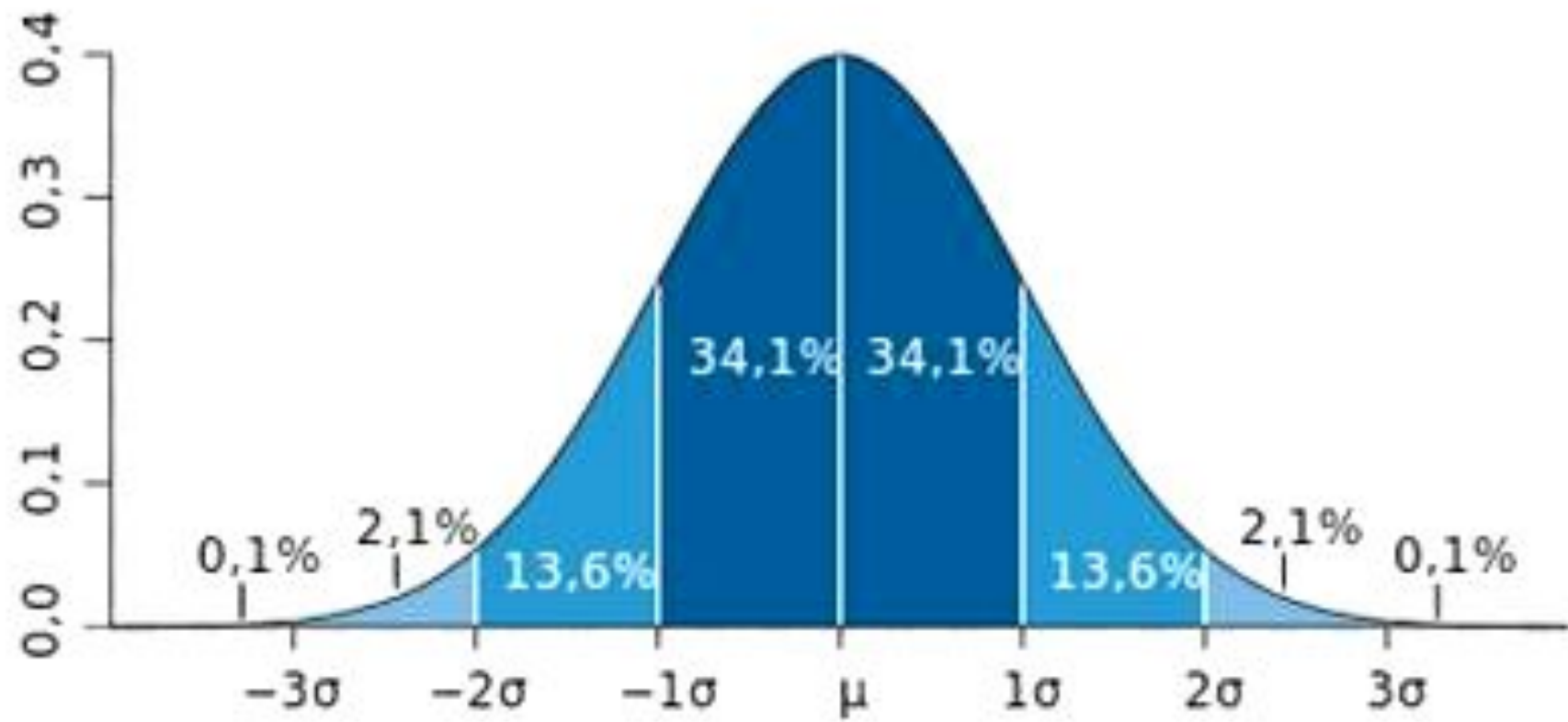


Рис. 7.2. Функция плотности нормального распределения вероятности,  $x$ . Симметричное относительно среднего  $\mu$ , дисперсия  $\sigma^2$  (а). Результат изменения среднего ( $\mu_2 > \mu_1$ ) (б). Результат изменения дисперсии ( $\sigma_1^2 < \sigma_2^2$ ) (в).



## ***Стандартное нормальное распределение***

- Нормальное распределение с нулевым средним и стандартным отклонением, равным 1.
- В литературе используют обозначение  $N(0,1)$ .

- Критерии согласия:

1. Колмогорова-Смирнова (Лиллиефорса)
2. Шапиро-Уилка
3. Пирсона

## Одновыборочный критерий нормальности Колмогорова-Смирнова

- Основан на максимуме разности между эмпирическим распределением выборки и предполагаемым распределением
- $D = \max |F_{\text{эмп}}(x) - F_{\text{гип}}(x)|$
- Для выборок большого объема
- Применим для непрерывнораспределенных случайных величин
- Можно использовать и для проверки согласия с любым другим законом

## Критерий Лиллиефорса

- статистический критерий, являющийся модификацией критерия Колмогорова–Смирнова.
- Используется для проверки нулевой гипотезы о том, что выборка распределена по нормальному закону для случая, когда параметры нормального распределения (математическое ожидание и дисперсия) априори неизвестны.



## Критерий Шапиро-Уилка

- Основан на отношении оптимальной линейной несмещенной оценки дисперсии к ее обычной оценке
- Обладает высокой мощностью
- Используется для объёма выборок  $8 \leq n \leq 50$
- Имеет вид:

$$W = S^2 / \sum_{i=1}^n (X_i - \bar{X})^2,$$

$$S = \sum_k a_k [X_{(n+1-k)} - X_{(k)}],$$

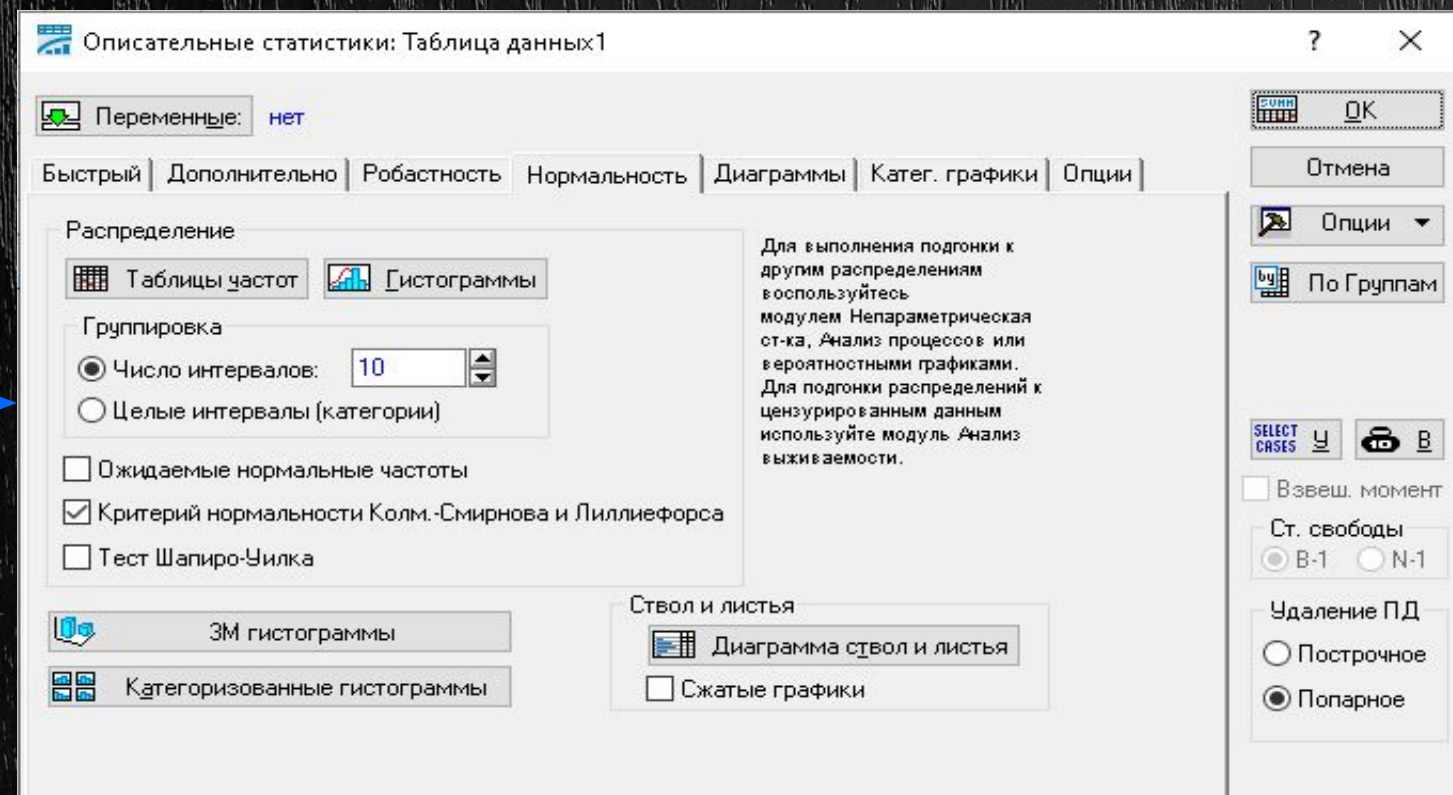
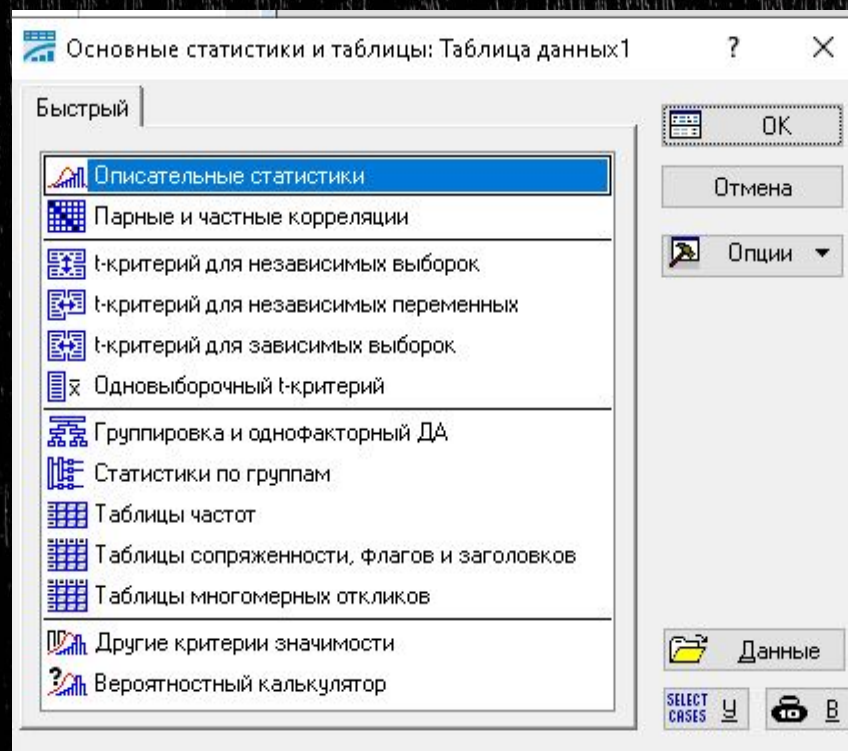
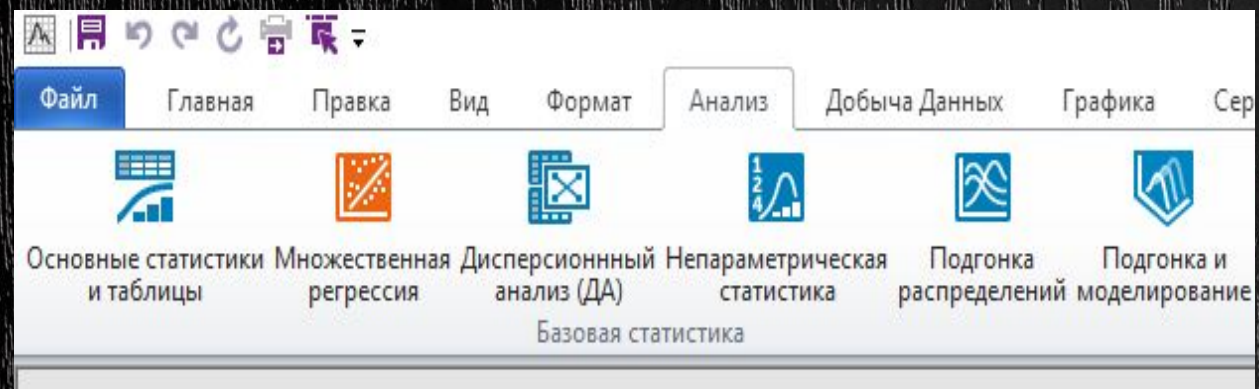
## Критерий Пирсона

- Отвечает на вопрос, с одинаковой ли частотой встречаются разные значения признака в эмпирическом и теоретическом распределении

- Имеет вид:

$$\chi^2_{\text{набл}} = \sum \frac{(n_i - n_i')^2}{n_i'}$$

- Объем выборки должен быть достаточно большим (не менее 30-50)
- Является общим критерием согласия





Спасибо за  
внимание!