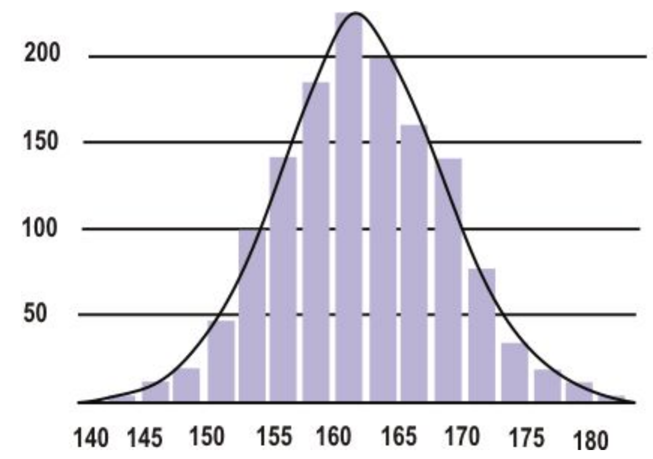
# Нормальное распределение:

## свойства и следствия из них

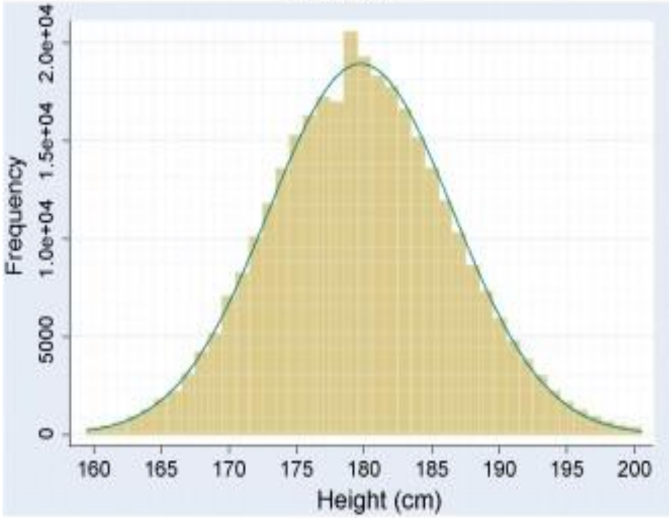
# Нормальное распределение

**Центральная предельная теорема в применении к Ψ**:

Если индивидуальная изменчивость некоторого свойства есть следствие действия множества причин, то распределение частот для всего многообразия проявлений этого свойства в генеральной совокупности соответствует кривой нормального распределения




(a) Height



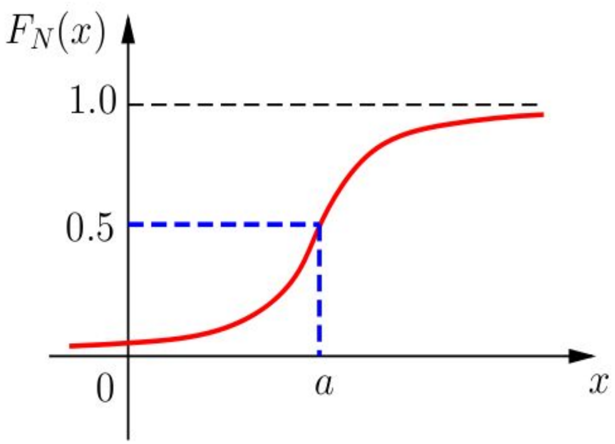
# Закон нормального распределения

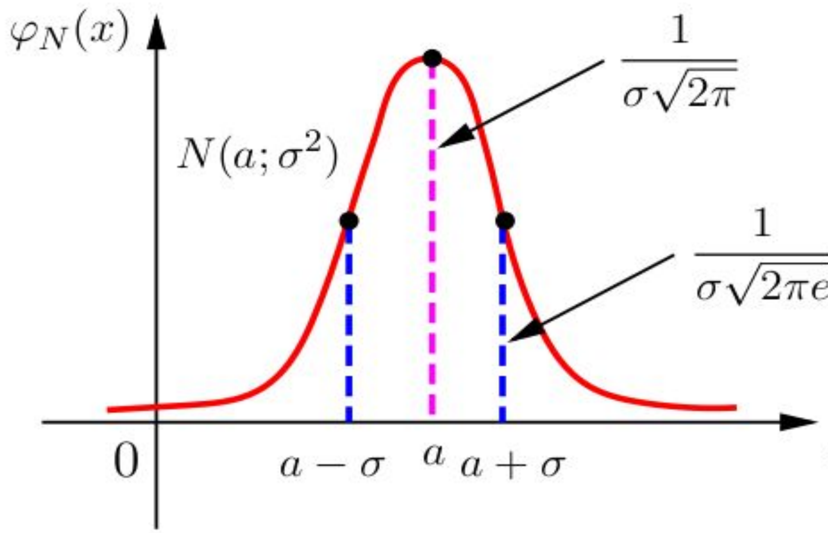
Непрерывная случайная величина X имеет нормальный закон распределения (закон Гаусса) с параметрами α и β, если ее плотность вероятности имеет вид:

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

Где:

- β — среднеквадратичное отклонение (σ);
- α — среднее (M);
- e, π - константы



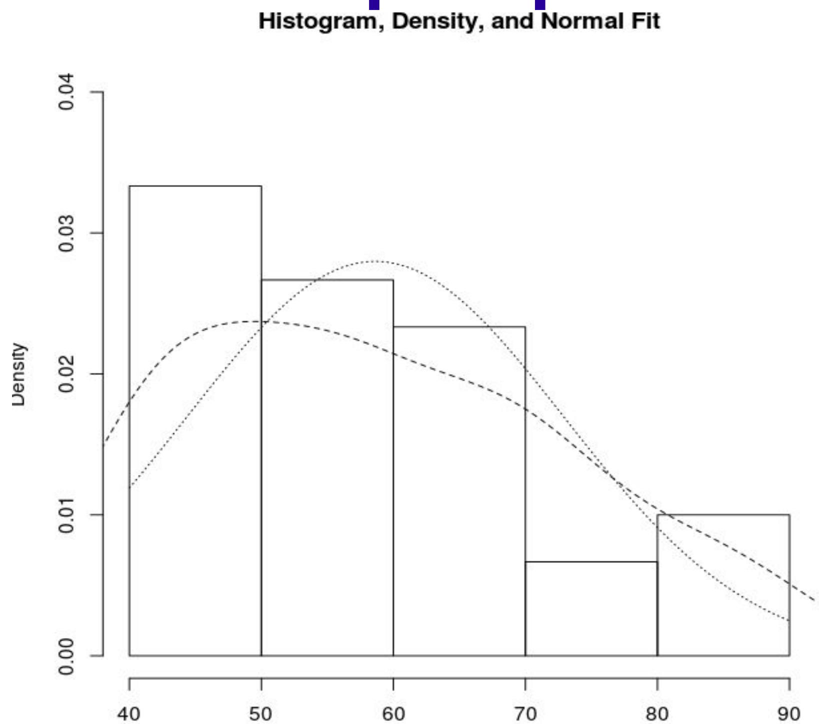
# Свойства нормального распределения

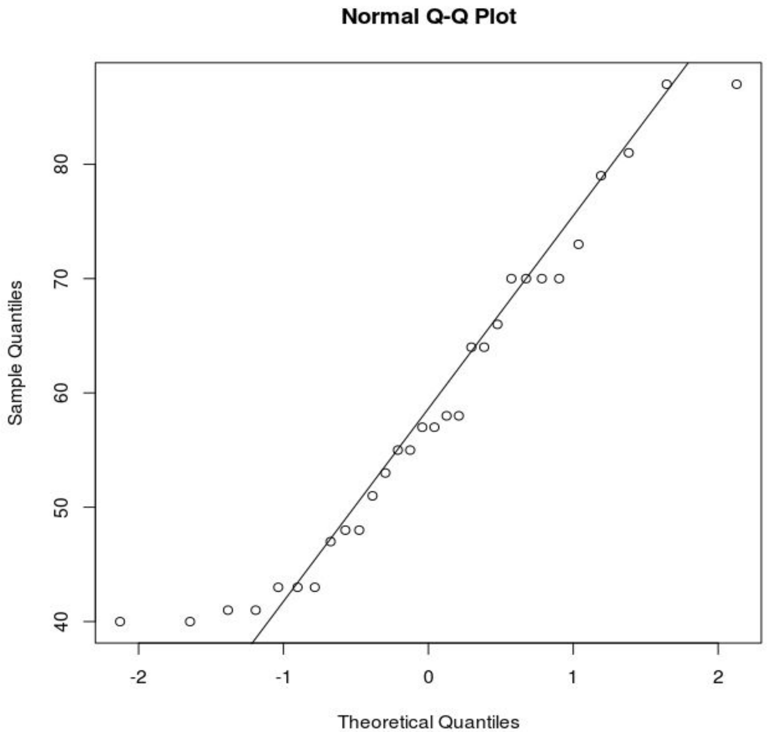
- Правило 3 сигм (99,72% значений лежат в рамках M+/-3σ)
- Распределение симметрично (А=0), эксцесс (мера остроты пика) Е = 0
- Мода, медиана и среднее совпадают
- Значения, лежащие на равном расстоянии от М (среднего), имеют равную частоту в выборке

# Проверка распределения на «нормальность»

- Графический способ (QQ-plot);
- Статистический критерий Колмогорова-Смирнова (N>50 человек) ;
- W-критерий Шапиро-Уилка (8<N<50 человек);
- Критерий асимметрии и эксцесса
- См. ГОСТ Р ИСО 5479—2002

# Графический способ




- Определить эмпирические процентили (5%, 10% ...);
- Посчитать теоретические процентили (через z-значения и оценки σ и X ген.совокупности)
- Разместить значения как точки с координатами (эмпирический процентиль; теоретический процентиль)
- Точки должны лежать на прямой

# Критерий асимметрии и эксцесса

1. Определить среднее арифметическое (М) и стандартное отклонение (σ).

2. Рассчитать показатели асимметрии и эксцесса.

- A= $\dfrac{\sum\left(\mathrm{X}_i - \mathrm{M}\right)^3}{\mathrm{N}\cdot\sigma^3}$ E= $\dfrac{\sum\left(\mathrm{X}_i - \mathrm{M}^4\right)}{\mathrm{N}\cdot\sigma^4}$ -3

3. Рассчитать критические значения А и Е

$$A_{кр} = 3\cdot\sqrt{\frac{6\cdot(n-1)}{(n+1)\cdot(n+3)}}$$

$$E_{кр} = 5\cdot\sqrt{\frac{24\cdot n\cdot(n-2)\cdot(n-3)}{(n+1)^2\cdot(n+3)\cdot(n+5)}}$$

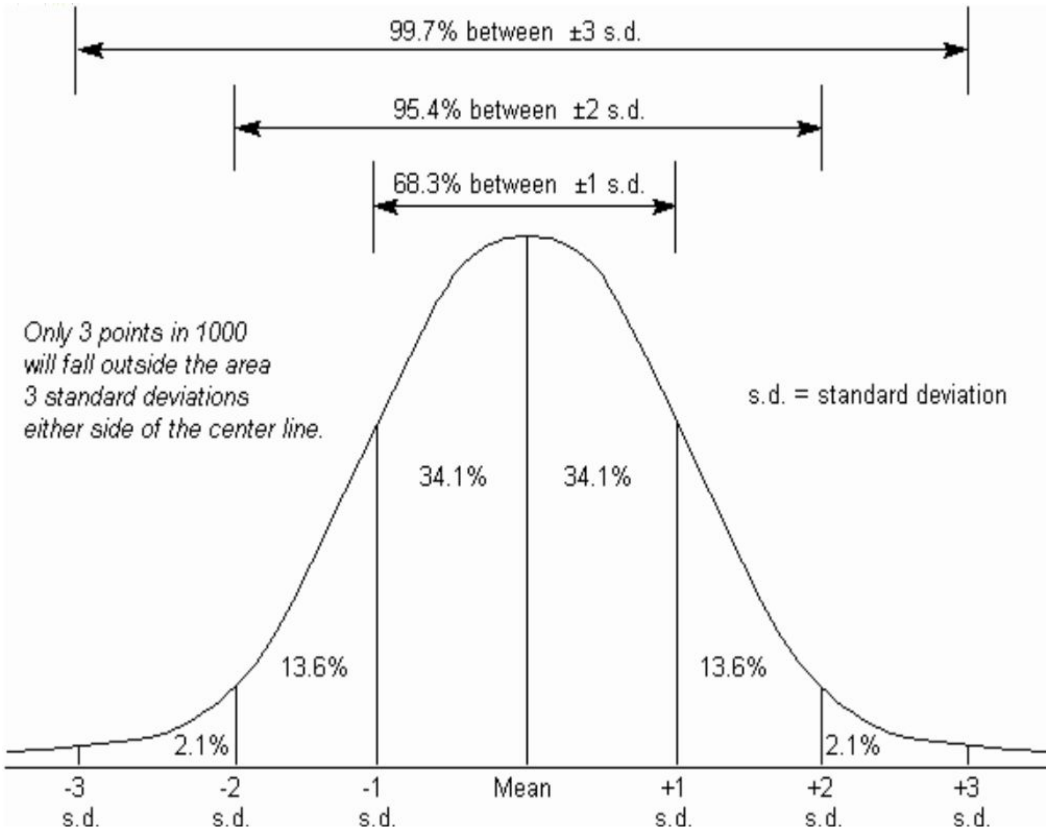
4. Если А<Акр и Е<Екр, распределение нормально

# Правило 3 сигм

При нормальном распределении:

- M(+/-)σ=68,26%
- M(+/-)2σ=95,44%
- M(+/-)3σ=99,72%,

M(+/-)3σ - интервал всех возможных значений




$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt$$

$$P(c < X < d) = \Phi\left(\frac{c-M}{s}\right) - \Phi\left(\frac{d-M}{s}\right)$$

# Стандартная шкала

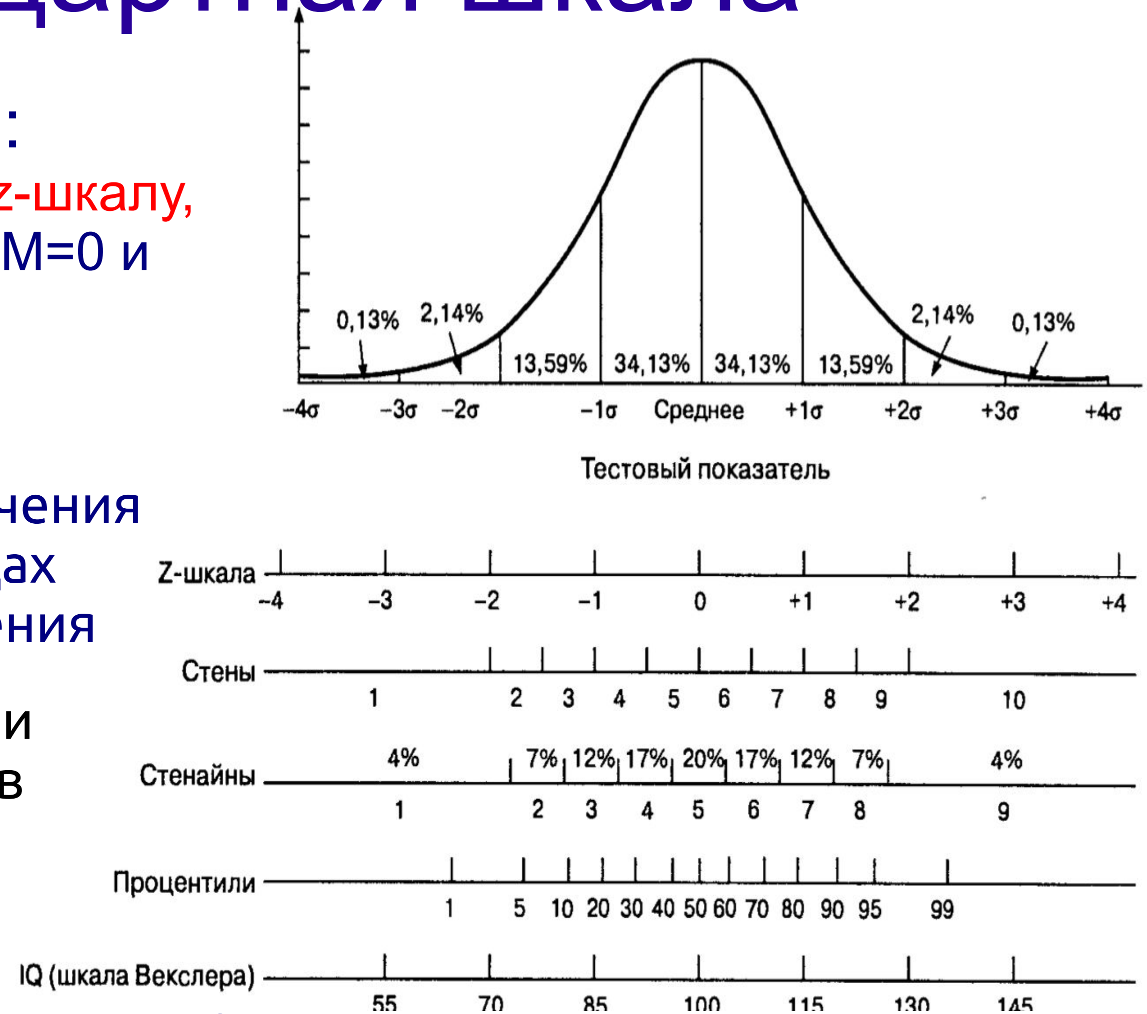
- **Стандартизация**:
  перевод измерений в z-шкалу, т.е. шкалу со средним M=0 и σ=1

$$z_i=(x_i-M)/\sigma$$

- Все полученные z-значения выражаются в единицах стандартного отклонения

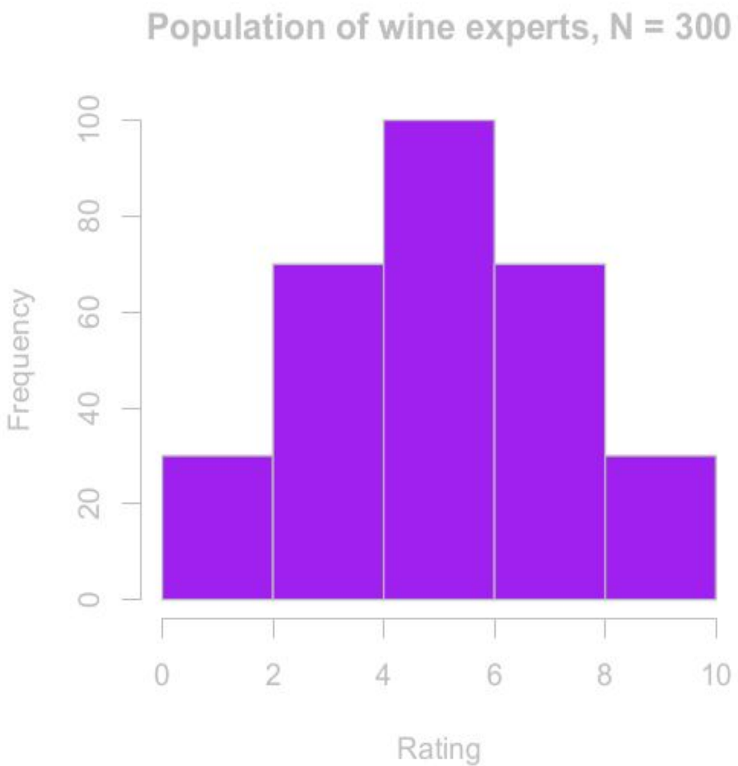
Z-шкала используется при стандартизации тестов

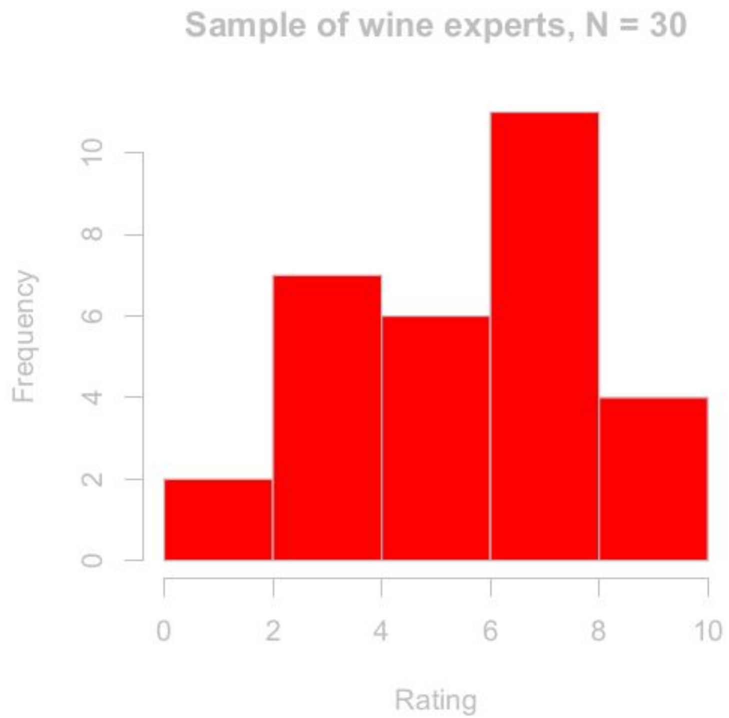
$$S_i=\sigma_s z_i+M_s$$




- Для стенов (st.ten) $M_s=5{,}5$ ; $\sigma_s=2$
- Для T-баллов $M_s=50$ ; $\sigma_s=10$
- Для IQ-баллов $M_s=100$ ; $\sigma_s=15$

# Ошибки выборки



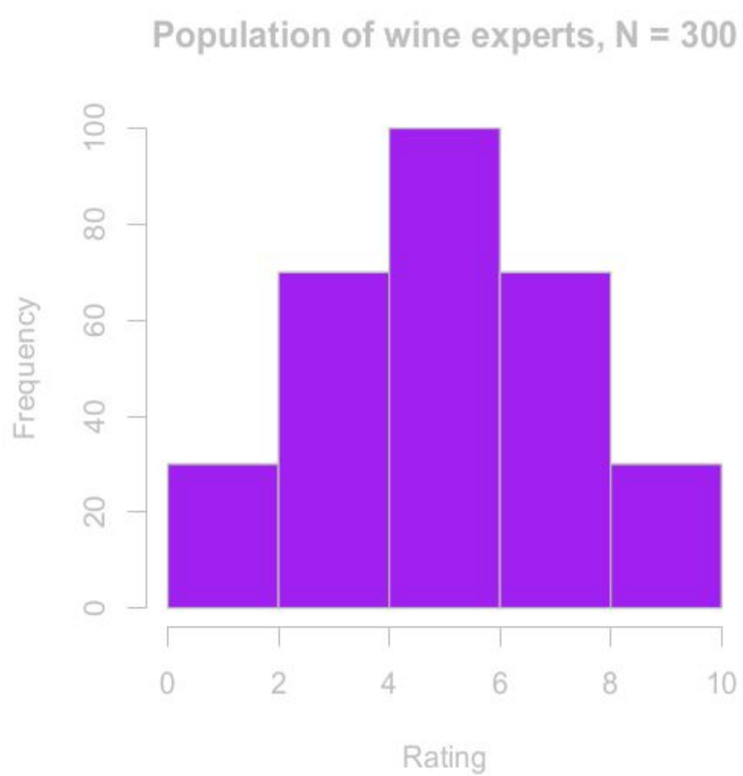

X = 5.5

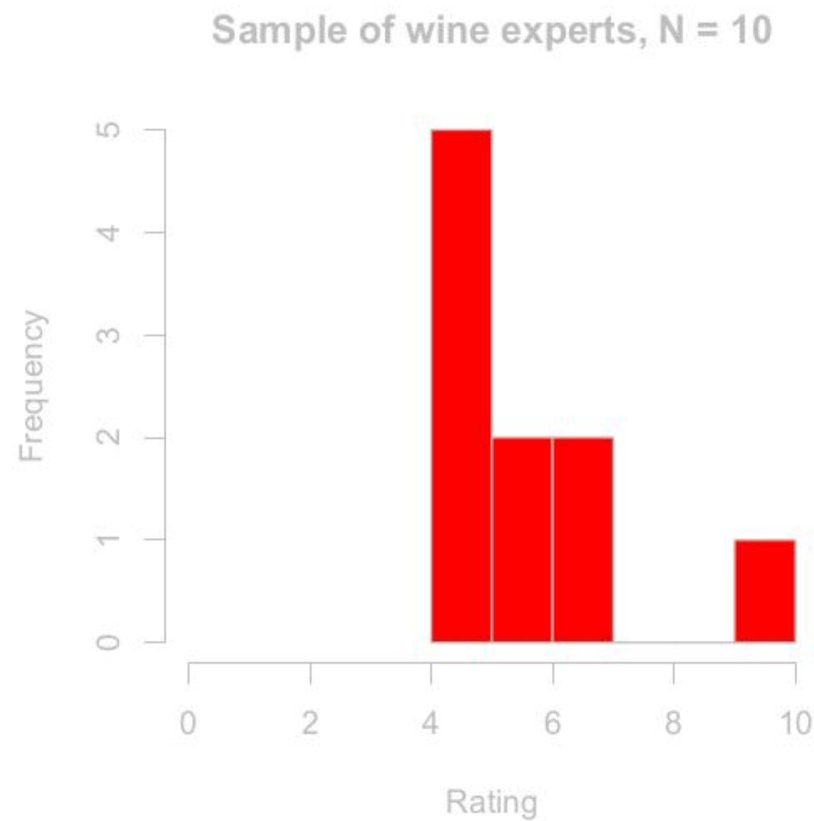
σ = 2.22

M = 5.93

s = 2.45

# Ошибки выборки




X = 5.5

σ = 2.22

M = 6.00

s = 1.70

# Чтобы не ошибиться

- Точечная оценка параметра=оценка одним числом

- Интервальная оценка параметра:

$$X_{min} < X < X_{max}$$

Интервал (Xmin,Xmax) = доверительный интервал

- Оценки (параметры) в генеральной совокупности при многократном измерении остаются в пределах точности измерения
- Статистические оценки в выборке (статистики) подвержены ошибкам и являются случайными величинами
- Мы можем только приблизительно оценивать параметры генеральной совокупности с помощью точечного или интервального оценивания