Практическое использование статистических критериев

ПМИ ФКН ВШЭ, 28 сентября 2019 г.

Денис Деркач[1], Алексей Артёмов[1,2],

(Ряд слайдов заимствован у Евгения Рябенко (Фейсбук))

[1] ФКН ВШЭ [2] Сколтех

Небольшая классификация критериев

- Одновыборочные $\{X_i\}_{i=1}^n$
  - Критерий Неймана-Пирсона 1 (согласие с одной из двух альтернатив)
  - Лемма Неймана-Пирсона 2
  - Критерий Колмогорова-Смирнова 3 (согласие с распределением)
- Двухвыборочные $\{X_i\}_{i=1}^n, \{Y_j\}_{j=1}^m$
  - Критерий Стьюдента 4 (равенство средних)
  - Критерий перестановок 5 («перемешиваемость» (равенство) распределений)
  - Ранговый критерий MWW 6 (равенство распределений)
  - Знаковый критерий 7 (равенство распределений)

Варианты двухвыборочных гипотез

О положении:

$\mathbb{H}_0\colon \mathbb{E}X = \mathbb{E}Y,$ $\quad$ $\mathbb{H}_1\colon \mathbb{E}X \;{<}{\neq}{>}\; \mathbb{E}Y;$

$\mathbb{H}_0\colon \operatorname{med} X = \operatorname{med} Y,$ $\quad$ $\mathbb{H}_1\colon \operatorname{med} X \;{<}{\neq}{>}\; \operatorname{med} Y;$

$\mathbb{H}_0\colon \mathbf{P}(X > Y) = 0.5,$ $\quad$ $\mathbb{H}_1\colon \mathbf{P}(X > Y) \;{<}{\neq}{>}\; 0.5;$

$\mathbb{H}_0\colon F_X(x) = F_Y(x),$ $\quad$ $\mathbb{H}_1\colon F_X(x) = F_Y(x+\Delta), \Delta \;{<}{\neq}{>}\; 0$

$\mathbb{H}_0\colon F_X(x) = F_Y(x),$ $\quad$ $\mathbb{H}_1\colon F_X(x) \;{<}{\neq}{>}\; F_Y(x).$

О рассеянии:

$\mathbb{H}_0\colon \mathbb{D}X = \mathbb{D}Y,$ $\quad$ $\mathbb{H}_1\colon \mathbb{D}X \;{<}{\neq}{>}\; \mathbb{D}Y;$

$\mathbb{H}_0\colon F_X(x) = F_Y(x+\Delta),$ $\quad$ $\mathbb{H}_1\colon F_X(x) = F_Y(\sigma x+\Delta), \sigma \;{<}{\neq}{>}$

Критерий Неймана-Пирсона

1 Показательный пример

- Пусть $X_i \sim \mathcal{N}(\theta, \sigma^2), i = 1, \ldots, n$, а нулевая гипотеза заключается в том, что $\theta = \theta_0$.
- Назовите достаточную статистику в этой задаче.

Показательный пример

- Пусть $X_i \sim \mathcal{N}(\theta, \sigma^2), i = 1, \ldots, n$, а нулевая гипотеза заключается в том, что $\theta = \theta_0$.
- Назовите достаточную статистику в этой задаче.
- Достаточная статистика: $T(\mathbf{X}^\ell) = \overline{X}_n$.
- Какое распределение $T(\mathbf{X}^\ell)$, когда верна $\mathbb{H}_0$?

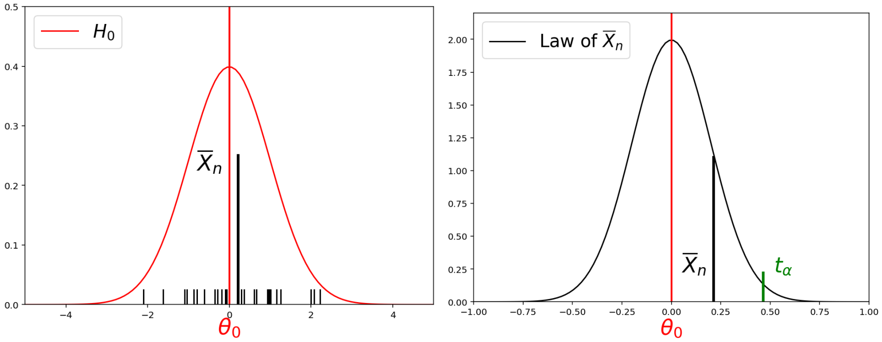
1 Показательный пример

- Пусть $X_i \sim \mathcal{N}(\theta, \sigma^2), i = 1, \ldots, n$, а нулевая гипотеза заключается в том, что $\theta = \theta_0$.
- Назовите достаточную статистику в этой задаче.
- Достаточная статистика: $T(\mathbf{X}^\ell) = \overline{X}_n$.
- Какое распределение $T(\mathbf{X}^\ell)$, когда верна $\mathbb{H}_0$?
- $T(\mathbf{X}^\ell) \sim \mathcal{N}(\theta_0, \sigma^2/n)$
- Какова критическая область? (Когда отклоняем $\mathbb{H}_0$?)

1 Показательный пример

- Пусть $X_i \sim \mathcal{N}(\theta, \sigma^2), i = 1, \ldots, n$, а нулевая гипотеза заключается в том, что $\theta = \theta_0$.
- Назовите достаточную статистику в этой задаче.
- Достаточная статистика: $T(\mathbf{X}^\ell) = \overline{X}_n$.
- Какое распределение $T(\mathbf{X}^\ell)$, когда верна $\mathbb{H}_0$?
- $T(\mathbf{X}^\ell) \sim \mathcal{N}(\theta_0, \sigma^2/n)$
- Какова критическая область? (Когда отклоняем $\mathbb{H}_0$?)
- Критическая область $\mathcal{R}_\alpha = [t_\alpha, \infty)$, т.е. $T(\mathbf{X}^\ell) \geqslant t_\alpha$.

Показательный пример



Показательный пример

- Подсчитайте вероятность ложной тревоги в этой задаче.

Показательный пример

- Подсчитайте вероятность ложной тревоги в этой задаче.

$$\begin{aligned} \alpha &= \mathbb{P}_{\theta_0}\left(\frac{\sqrt{n}(\overline{X}_n - \theta_0)}{\sigma} \geqslant \frac{\sqrt{n}(t_\alpha - \theta_0)}{\sigma}\right) = \\ &= 1 - \Phi\left(\frac{\sqrt{n}(t_\alpha - \theta_0)}{\sigma}\right). \end{aligned}$$

- Как выбрать t_α, чтобы $\alpha \leqslant \alpha_0$?

Показательный пример

- Подсчитайте вероятность ложной тревоги в этой задаче.

$$\begin{aligned} \alpha &= \mathbb{P}_{\theta_0}\Big(\frac{\sqrt{n}(\overline{X}_n - \theta_0)}{\sigma} \geqslant \frac{\sqrt{n}(t_\alpha - \theta_0)}{\sigma}\Big) = \\ &= 1 - \Phi\Big(\frac{\sqrt{n}(t_\alpha - \theta_0)}{\sigma}\Big). \end{aligned}$$

- Как выбрать t_α, чтобы $\alpha \leqslant \alpha_0$?

$$t_{\alpha_0} = \theta_0 + \sigma x_{1-\alpha_0}/\sqrt{n}$$

Показательный пример

- Подсчитайте вероятность ложной тревоги в этой задаче.

$$\begin{aligned}\alpha &= \mathbb{P}_{\theta_0}\left(\frac{\sqrt{n}(\overline{X}_n - \theta_0)}{\sigma} \geqslant \frac{\sqrt{n}(t_\alpha - \theta_0)}{\sigma}\right) = \\ &= 1 - \Phi\left(\frac{\sqrt{n}(t_\alpha - \theta_0)}{\sigma}\right).\end{aligned}$$

- Как выбрать t_α, чтобы $\alpha \leqslant \alpha_0$?

$$t_{\alpha_0} = \theta_0 + \sigma x_{1-\alpha_0}/\sqrt{n}$$

- Пусть на самом деле верна альтернатива $\mathbb{H}_1 : \theta = \theta_1$, причем $\theta_1 > \theta_0$. Какова вероятность ошибки 2-го рода?

Показательный пример

- Подсчитайте вероятность ложной тревоги в этой задаче.

$$\begin{aligned}\alpha &= \mathbb{P}_{\theta_0}\left(\frac{\sqrt{n}(\overline{X}_n - \theta_0)}{\sigma} \geqslant \frac{\sqrt{n}(t_\alpha - \theta_0)}{\sigma}\right) = \\ &= 1 - \Phi\left(\frac{\sqrt{n}(t_\alpha - \theta_0)}{\sigma}\right).\end{aligned}$$

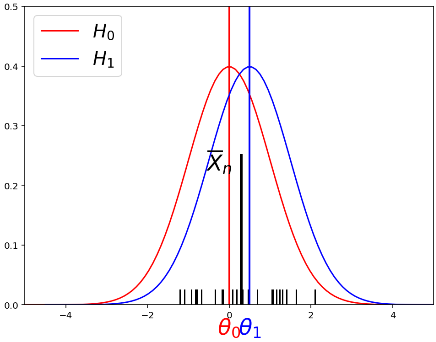
- Как выбрать t_α, чтобы $\alpha \leqslant \alpha_0$?

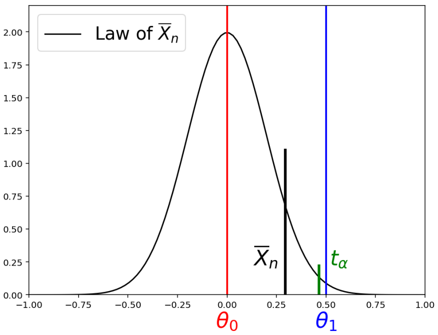
$$t_{\alpha_0} = \theta_0 + \sigma x_{1-\alpha_0}/\sqrt{n}$$

- Пусть на самом деле верна альтернатива $\mathbb{H}_1 : \theta = \theta_1$, причем $\theta_1 > \theta_0$. Какова вероятность ошибки 2-го рода?

$$\beta = \mathbb{P}_{\theta_1}\left(\overline{X}_n < t_{\alpha_0}\right) = \Phi\left(x_{1-\alpha_0} - \frac{\sqrt{n}(\theta_1 - \theta_0)}{\sigma}\right).$$

Показательный пример





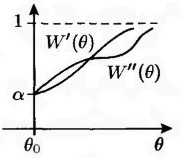
Лемма Неймана-Пирсона

Сравнение двух критериев

- Критерии на самом деле задаются критическими множествами
- Пусть имеется два критерия, заданных множествами $\mathcal{R}'_\alpha$ и $\mathcal{R}''_\alpha$. Какой выбрать?

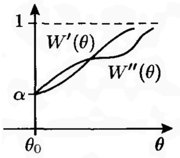
Сравнение двух критериев

- Критерии на самом деле задаются критическими множествами
- Пусть имеется два критерия, заданных множествами $\mathcal{R}'_\alpha$ и $\mathcal{R}''_\alpha$. **Какой выбрать?**
- **Сложная альтернатива:** необходимо сравнивать функции мощности $W'(\theta)$ и $W''(\theta)$



Сравнение двух критериев

- Критерии на самом деле задаются критическими множествами
- Пусть имеется два критерия, заданных множествами $\mathcal{R}'_\alpha$ и $\mathcal{R}''_\alpha$. **Какой выбрать?**
- **Сложная альтернатива:** необходимо сравнивать функции мощности $W'(\theta)$ и $W''(\theta)$



- **Простая альтернатива:** существует наиболее мощный критерий (Неймана-Пирсона).
- Идея: при заданной (достаточно малой) вероятности ошибки 1-го рода α постараться уменьшить вероятность ошибки 2-го рода β насколько возможно за счет подбора критического множества $\mathcal{R}_\alpha$.

Сравнение двух критериев

- Пусть задана выборка $\mathbf{X}^\ell$
- Гипотеза $\mathbb{H}_0$ и альтернатива $\mathbb{H}_1$ порождают в выборочном пространстве $\mathbb{R}^\ell$ меры $\mathbb{P}_0$ и $\mathbb{P}_1$
- Таким образом, необходимо найти множество G такое, что $\mathbb{P}_0(G) \leqslant \alpha$ и $\mathbb{P}_1(G) \to \sup\limits_{G:\mathbb{P}_0(G)\leqslant\alpha} \mathbb{P}_1(G)$
- Рассмотрим систему вложенных множеств $G_c = \{\boldsymbol{x} \in \mathbb{R}^\ell : \frac{p_1(\boldsymbol{x})}{p_0(\boldsymbol{x})} \geqslant c\}$
- Пусть $\varphi(c) = \mathbb{P}_0(G_c)$, тогда $\varphi(c)$ убывает с ростом c

Сравнение двух критериев

- На самом деле, $\varphi(c)$ убывает быстрее, чем $1/c$:

$$1 \geqslant \mathbb{P}_1(G_c) = \int\limits_{G_c} p_1(\boldsymbol{x})d\boldsymbol{x} \geqslant c \int\limits_{G_c} p_0(\boldsymbol{x})d\boldsymbol{x} = c\mathbb{P}_0(G_c) = c\varphi(c).$$

- Далее еще потребуем, чтобы плотности $p_1(\boldsymbol{x})$ и $p_0(\boldsymbol{x})$ были всюду положительны
- Дополнительно потребуем, чтобы $\forall \alpha \in (0, 1) \quad \exists c = c_\alpha : \quad \varphi(c_\alpha) = \alpha.$

Лемма Неймана-Пирсона

Лемма (Неймана-Пирсона)

Наиболее мощный критерий уровня α задается критическим множеством

$$G^* = G_{c_\alpha} = \left\{ \boldsymbol{x} \in \mathbb{R}^\ell : \frac{p_1(\boldsymbol{x})}{p_0(\boldsymbol{x})} \geqslant c_\alpha \right\}$$

- Пусть G — критическое множество уровня α.
- Тогда $\mathbb{P}_0(G_c) \geqslant \alpha = \mathbb{P}_0(G_{c_\alpha})$
- Пусть $I(\boldsymbol{x})$ — индикатор G_c, $I^*(\boldsymbol{x})$ — индикатор G_{c_α}
- Функция

$$f(\boldsymbol{x}) = (I^*(\boldsymbol{x}) - I(\boldsymbol{x}))(p_1(\boldsymbol{x}) - c_\alpha p_0(\boldsymbol{x}))$$

неотрицательна при всех $\boldsymbol{x} \in \mathbb{R}^\ell$

Лемма Неймана-Пирсона

- Функция

$$f(\boldsymbol{x}) = (I^*(\boldsymbol{x}) - I(\boldsymbol{x}))(p_1(\boldsymbol{x}) - c_\alpha p_0(\boldsymbol{x}))$$

неотрицательна при всех $\boldsymbol{x} \in \mathbb{R}^\ell$

- Поэтому

$$\begin{aligned} 0 \ &\leqslant \ \int\limits_{\boldsymbol{x}\in\mathbb{R}^\ell} f(\boldsymbol{x})d\boldsymbol{x} = \int\limits_{\boldsymbol{x}\in\mathbb{R}^\ell} I^*(\boldsymbol{x})p_1(\boldsymbol{x})d\boldsymbol{x} - \int\limits_{\boldsymbol{x}\in\mathbb{R}^\ell} I(\boldsymbol{x})p_1(\boldsymbol{x})d\boldsymbol{x} - \\ &- \ c_\alpha\Big[\int\limits_{\boldsymbol{x}\in\mathbb{R}^\ell} I^*(\boldsymbol{x})p_0(\boldsymbol{x})d\boldsymbol{x} - \int\limits_{\boldsymbol{x}\in\mathbb{R}^\ell} I(\boldsymbol{x})p_0(\boldsymbol{x})d\boldsymbol{x}\Big] = \\ &= \ \mathbb{P}_1(G^*) - \mathbb{P}_1(G) - c_\alpha \underbrace{\left[\mathbb{P}_0(G^*) - \mathbb{P}_0(G)\right]}_{\geqslant 0}. \end{aligned}$$

Критерий Неймана-Пирсона

- $\mathbb{H}_0 : \theta = \theta_0$ vs. $\mathbb{H}_1 : \theta = \theta_1$
- Статистика Неймана-Пирсона:
$$T = \frac{\mathcal{L}(\theta_1)}{\mathcal{L}(\theta_0)} = \frac{\prod_{i=1}^{n} f(X_i; \theta_1)}{\prod_{i=1}^{n} f(X_i; \theta_0)}. \tag{1}$$
- Допустим, что $\mathbb{H}_0$ отвергается при $T > k$. Выберем k так, что $\mathbb{P}_{\theta_0}(T > k) = \alpha$.
- Тогда, критерий Неймана-Пирсона (на основе статистики (2)) будет иметь наибольшую мощность $W(\theta_1)$ среди всех критериев размера α.

Пример

- $X_i \sim \mathcal{N}(\mu, \sigma^2)$, причем дисперсия σ^2 известна
- $\mathbb{H}_0 : \mu = \mu_0$ vs. $\mathbb{H}_1 : \mu = \mu_1$
- Статистика Неймана-Пирсона:

$$T = \frac{\mathcal{L}(\theta_1)}{\mathcal{L}(\theta_0)} = \frac{\prod_{i=1}^{n} \mathcal{N}(X_i; \mu_1, \sigma^2)}{\prod_{i=1}^{n} \mathcal{N}(X_i; \mu_0, \sigma^2)}. \tag{2}$$

- Подсчитайте статистику критерия (упростите)

Пример

- $X_i \sim \mathcal{N}(\mu, \sigma^2)$, причем дисперсия σ^2 известна
- $\mathbb{H}_0 : \mu = \mu_0$ vs. $\mathbb{H}_1 : \mu = \mu_1$
- Статистика Неймана-Пирсона:

$$T = \frac{\mathcal{L}(\theta_1)}{\mathcal{L}(\theta_0)} = \frac{\prod_{i=1}^{n} \mathcal{N}(X_i; \mu_1, \sigma^2)}{\prod_{i=1}^{n} \mathcal{N}(X_i; \mu_0, \sigma^2)}. \tag{2}$$

- Подсчитайте статистику критерия (упростите)
- Получается

$$\begin{aligned} T &= \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{\ell} \left[(X_i - \mu_1)^2 - (X_i - \mu_0)^2 \right] \right] = \\ &= \exp\left[\frac{n}{\sigma^2} (\mu_1 - \mu_0) \underbrace{\left[\overline{X}_\ell - \frac{\mu_1 + \mu_0}{2} \right]}_{\text{важен знак!}} \right] \end{aligned}$$

Критерий согласия Колмогорова

3 Выборочная функция распределения

- Пусть дана выборка $\{X_i\}_{i=1}^n$ из распределения $F(x)$.
- Выборочной функцией распределения называется функция $\widehat{F}_n(x)$:

$$\widehat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{X_i \leqslant x}.$$

  где $I_{X_i \leqslant x}$ указывает, попало ли наблюдение X_i в область $(-\infty, x]$:

$$I_{X_i \leqslant x} = \begin{cases} 1, & X_i \leqslant x; 0, \\ X_i > x. \end{cases}$$

- Легко проверить, что $\widehat{F}_n(x)$ — состоятельная оценка $F(x)$, при этом $\widehat{F}_n(x) \sim \mathrm{Bin}(n, F(x))$.

Выборочная функция распределения

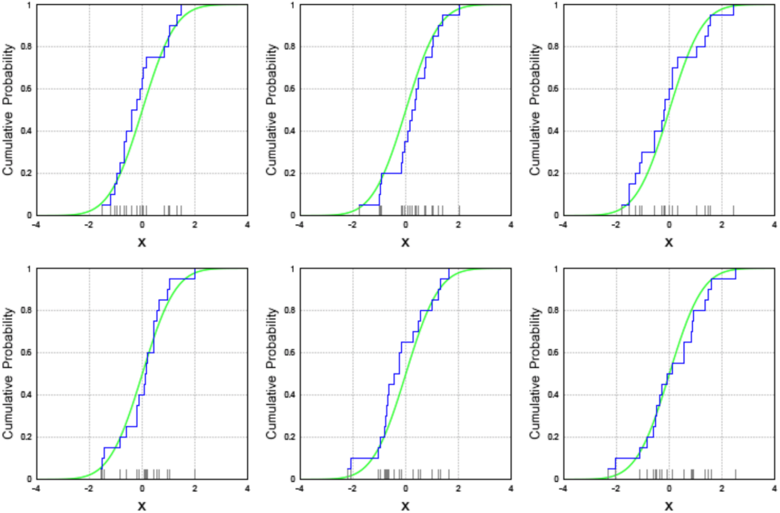


Рис.: https://en.wikipedia.org/wiki/Empirical_distribution_function

Критерий согласия Колмогорова

- Пусть выборочная функция распределения $\widehat{F}_n$, построенная по выборке $\{X_i\}_{i=1}^n$, имеет вид:

$$\widehat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{X_i \leqslant x},$$

- Задача:

$$\mathbb{H}_0 : \widehat{F}_n(X) = F(X) \quad \text{vs.} \quad \mathbb{H}_1 : \widehat{F}_n(X) \neq F(X)$$

Критерий согласия Колмогорова

- Статистика критерия для $\widehat{F}_n(x)$:

$$D_n = \sup_x |\widehat{F}_n(x) - F(x)|,$$

- Распределение статистики Колмогорова: если $F(X) \in C^1(\mathbb{X})$, то для введённой статистики справедливо:

$$\forall t > 0\colon \lim_{n \to \infty} P(\sqrt{n} D_n \leqslant t) = K(t) = \sum_{j=-\infty}^{+\infty} (-1)^j e^{-2j^2 t^2}.$$

- Если $\sqrt{n} D_n$ превышает критическое значение K_α уровня α, то $\mathbb{H}_0$ отвергается. Иначе не отвергается на уровне α.
- Если α достаточно близко к 1, то: $K_\alpha \approx \sqrt{-\frac{1}{2} \ln \frac{1-\alpha}{2}}$.
- Асимптотическая мощность критерия равна 1.

Критерий Стьюдента

4 Критерий Стьюдента (t-test)

- Пусть $X, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$, где (μ, σ^2) неизвестны.
- Задача
$$\mathbb{H}_0 : \mu = \mu_0 \quad vs. \quad \mathbb{H}_1 : \mu \neq \mu_0$$
- Обозначим через S_n^2 выборочную дисперсию. Статистика критерия:
$$T = \frac{\sqrt{n}(\overline{X_n} - \mu_0)}{S_n}$$
- Основная гипотеза отвергается, если $|T| > t_{n-1,\alpha/2}$, где $t_{n-1,\alpha/2}$ – квантиль распределения Стьюдента с $n-1$ степенями свободы.
- При больших n выполняется $T \sim \mathcal{N}(0, 1)$, то есть при больших n t-критерий эквивалентен критерию Вальда.

Распределение Стьюдента (t-test)

Определение

Случайная величина имеет распределение Стьюдента (t-распределение) с k степенями свободы, если:

$$f(t) = \frac{\Gamma(\frac{k+1}{2})}{\sqrt{k\pi}\Gamma(\frac{k}{2})(1 + \frac{t^2}{k})^{\frac{k+1}{2}}}$$

При $k \to \infty$ t-распределение стремится к стандартному нормальному распределению. При $k = 1$ t-распределение совпадает с распределением Коши.

- t-критерий используют, когда распределение данных близко к нормальному, а размер выборки невелик

Критерий Стьюдента (t-test)

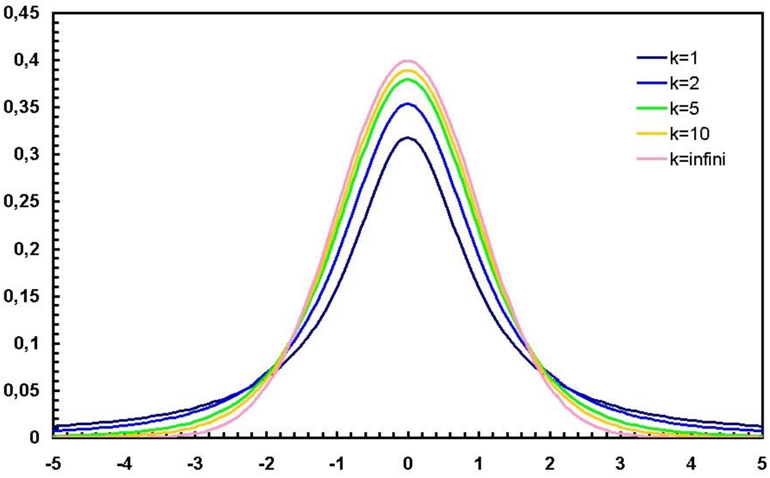


Рис.: https://ru.wikipedia.org/wiki/Распределение_Стьюдента

Критерий Стьюдента (t-test)

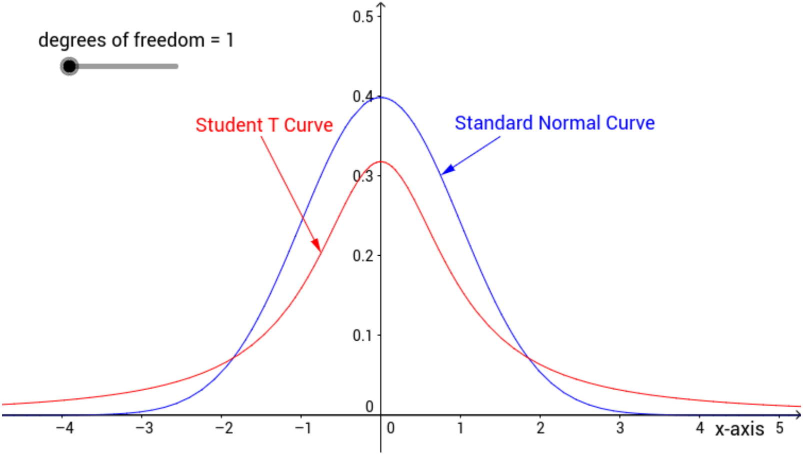


Рис.: http://tananyag.geomatech.hu/m/53882

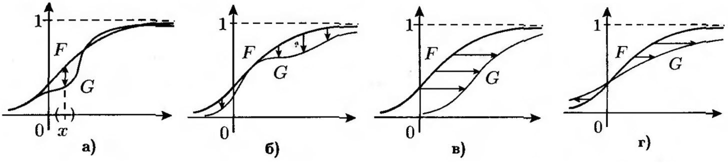
Двухвыборочный t-критерий

- $\{X_i\}_{i=1}^n, \{Y_j\}_{j=1}^m$ — две выборки из нормальных распределений $\mathcal{N}(\mu_X, \sigma_X^2), \mathcal{N}(\mu_Y, \sigma_Y^2)$
- Задача
$$\mathbb{H}_0 : \mu_X = \mu_Y \quad \text{vs.} \quad \mathbb{H}_1 : \mu_X \neq \mu_Y$$
- Механика проверки гипотезы та же, что и раньше.

Ранговые непараметрические критерии

Альтернативы однородности

- Имеем две выборки $\mathbf{X}^n \sim F(x)$ и $\mathbf{Y}^m \sim G(x)$
- Гипотеза однородности $\mathbb{H}_0 : F(x) = G(x), x \in \mathbb{R}$



- Бывает важно уловить отклонения от $\mathbb{H}_0$ только определенного типа (наличие прироста Y_j по сравнению с X_i)
- Сужение альтернативы $\implies$ более эффективные критерии
- Перестановки: гипотеза однородности против альтернативы неоднородности $\mathbb{H}_1$ (вариант а))
- Манн-Уитни: гипотеза однородности против альтернативы доминирования $\mathbb{H}_1$ (варианты б) и в))

Перестановочные непараметрические критерии

5 Критерий перестановок

- Критерий перестановок применяется для проверки того, отличаются ли распределения.
- Пусть $X_1, ..., X_m \sim F_X$ и $Y_1, ..., Y_n \sim F_Y$ - две независимые выборки. Требуется решить:

$$\mathbb{H}_0 : F_X = F_Y \quad vs. \quad \mathbb{H}_1 : F_X \neq F_Y$$

- Критерий перестановок — «точный» в том смысле, что он не использует предположения об асимптотической сходимости к нормальному распределению.

Критерий перестановок:

1. Обозначим через $T(x_1, ..., x_m, y_1, ..., y_n)$ некоторую тестовую статистику, например, $T(X_1, ..., X_m, Y_1, ..., Y_n) = |\overline{X}_m - \overline{Y}_n|$.
2. Положим $N = m + n$ и рассмотрим все $N!$ перестановок объединенной выборки $X_1, ..., X_m, Y_1, ..., Y_n$.
3. Для каждой из перестановок подсчитаем значение статистики T.
4. Обозначим эти значения $T_1, ..., T_{N!}$.

Теорема (Критерий перестановок)

Если $\mathbb{H}_0$ верна, то при фиксированных упорядоченных значениях $\{X_1, ..., X_m, Y_1, ..., Y_n\}$ значение статистики T распределены равномерно на множестве $T_1, ..., T_{N!}$.

Критерий перестановок

Теорема

Обозначим как перестановочное распределение статистики T такое, согласно которому:

$$\mathbb{P}_0(T = T_i) = \frac{1}{N!}, \quad i = 1, ..., N!$$

Пусть t_{obs} — значение статистики, которое было получено в опыте. Тогда:

$$p\text{-value} = \mathbb{P}(T > t_{obs}|f) = \frac{1}{N!}\sum_{j=1}^{N!}\mathbb{I}(T_j > t_{obs}), \quad f \in \mathcal{F}_0$$

Критерий перестановок: пример

- Пусть $(X_1, X_2, Y_1) = (1, 9, 3)$.
- Пусть $T(X_1, X_2, Y_1) = |\overline{X} - \overline{Y}| = 2$, тогда

Перестановка	Значение T	Вероятность
(1,9,3)	2	$1/6$
(9,1,3)	2	$1/6$
(1,3,9)	7	$1/6$
(3,1,9)	7	$1/6$
(3,9,1)	5	$1/6$
(9,3,1)	5	$1/6$

- p-value $= \mathbb{P}(T > 2) = 4/6$.

Вариационный ряд, ранги, связки

- Вариационный ряд:

$$X_1, \ldots, X_n \;\Rightarrow\; X_{(1)} \leqslant \ldots < \underbrace{X_{(k_1)} = \ldots = X_{(k_2)}}_{\text{связка размера } k_2 - k_1 + 1} < \ldots \leqslant X_{(n)}$$

- **Ранг** наблюдения X_i:
- если X_i не в связке, то $\operatorname{rank}(X_i) = r\colon X_i = X_{(r)}$,
- если X_i в связке $X_{(k_1)}, \ldots, X_{(k_2)}$, то $\operatorname{rank}(X_i) = \frac{k_1 + k_2}{2}$.

6 Критерий ранговых сумм MWW

- Построим вариационный ряд из объединенной выборки $(X_1, \ldots, X_n, Y_1, \ldots, X_m)$
  - Верна $\mathbb{H}_0 \implies$ значения Y_j рассеяны по всему ряду
  - Иначе средний ранг значений Y_j относительно большой
- Обозначим S_j ранг порядковой статистики $Y_{(j)}$ в этом ряду
- Положим $V = S_1 + \cdots + S_m$
- Критическая область: $V \geqslant c$, где $c = \mathrm{const}$
- **Большие выборки:** (Mann-Whitney-Wilcoxon, MWW)

$$U = \sum_{i=1}^{n} \sum_{j=1}^{m} I_{X_i < Y_j} \to \mathcal{N}\left(\frac{nm}{2}, \frac{nm(n+m+1)}{12}\right)$$

Знаковые непараметрические критерии

7 Двухвыборочный критерий знаков

выборки: $\mathbf{X}^\ell = (X_1, \ldots, X_n)$

$\mathbf{Y}^\ell = (Y_1, \ldots, Y_n)\,, X_i \neq Y_i$

выборки связанные

нулевая гипотеза: $\mathbb{H}_0\colon\ \mathbf{P}(X > Y) = \frac{1}{2}$

альтернатива: $\mathbb{H}_1\colon\ \mathbf{P}(X > Y) <\neq> \frac{1}{2}$

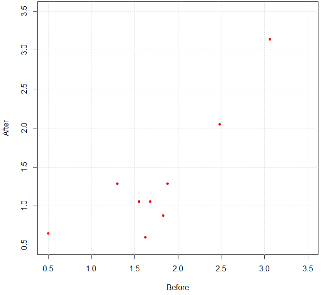
статистика: $T\left(\mathbf{X}^\ell, \mathbf{Y}^\ell\right) = \sum\limits_{i=1}^{n} \mathbb{I}_{\{X_i > Y_i\}}$

нулевое распределение: $\mathrm{Bin}(n, \frac{1}{2})$

Двухвыборочный критерий знаков: Примеры

- (Hollander & Wolfie, 29f): депрессивность 9 пациентов была измерена по шкале Гамильтона до и после первого приёма транквилизатора. Подействовал ли транквилизатор?

- H_0: уровень депрессивности не изменился.
  H_1: уровень депрессивности снизился.
- Критерий знаков: $p = 0.09$, 95% нижний доверительный предел для медианы изменения -0.041.



Двухвыборочный критерий знаков: Примеры

- (Laureysens et al., 2004): для 13 разновидностей тополей, растущих в зоне интенсивного загрязнения, в августе и ноябре измерялась средняя концентрация алюминия в микрограммах на грамм древесины.
- $\mathbb{H}_0:$ концентрация алюминия не менялась.
  $\mathbb{H}_1:$ концентрация алюминия изменилась.
- Для тополей 10 из 13 разновидностей концентрация алюминия увеличилась.
- Критерий знаков: $p = 0.0923$, 95% доверительный интервал для медианы изменения $[-0.687, 10.107]$.

Причины использовать критерий знаков

- Точные разности $X_i - Y_i$ неизвестны, известны только их знаки.
- Разности $X_i - Y_i$ при $\mathbb{H}_1$ могут быть небольшими по модулю, но иметь систематический характер по знаку.
- Разности $X_i - Y_i$ при $\mathbb{H}_0$ могут быть большими по модулю, но случайными но знаку.