40. Проверка гипотез о совпадении средних двух нормальных совокупностей.

Status	Completed

Теорема Стьюдента

если средние и дисперсии совпадают n, m - размеры независимых нормальных выборок X, Y соответственно.

$$\frac{\overline{X}-\overline{Y}}{\sqrt{1/n+1/m}\sqrt{\frac{nS_X^2+mS_Y^2}{n+m-2}}} \sim T_{n+m-2}$$

☝️ Тут требуется, чтобы дисперсии совпадали! Иначе нельзя применить такой критерий!

Доказательство

$$\overline{X} \in \Phi_{\alpha_1, \sigma^2/n}, \quad \overline{Y} \in \Phi_{\alpha_2, \sigma^2/m},$$

имеем

$$\overline{X}-\overline{Y}\in\Phi_{\alpha_1-\alpha_2,\,\sigma^2(1/n+1/m)}$$

и после стандартизации

$$\frac{\overline{X}-\overline{Y}-(\alpha_1-\alpha_2)}{\sqrt{\sigma^2\left(\frac{1}{n}+\frac{1}{m}\right)}}\in\Phi_{0,1}.$$

Далее, по свойству распределения хи-квадрат

$$\frac{nS_X^2}{\sigma^2}+\frac{mS_Y^2}{\sigma^2}\in\chi^2_{n+m-2};$$

эта случайная величина не зависит от $\overline{X}-\overline{Y}$. Таким образом,

$$\frac{\overline{X}-\overline{Y}-(\alpha_1-\alpha_2)}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}}:\sqrt{\frac{1}{n+m-2}\frac{nS_X^2+mS_Y^2}{\sigma^2}}\in T_{n+m-2}.$$

Если верна гипотеза H_1, то $\alpha_1-\alpha_2=0$ и

$$\psi=\frac{\overline{X}-\overline{Y}}{\sqrt{\frac{1}{n}+\frac{1}{m}}\sqrt{\frac{nS_X^2+mS_Y^2}{n+m-2}}}\in T_{n+m-2}.$$

Критерий Стьюдента (о совпадении средних)

τ - квантиль данного распределения.

$$d=\frac{\overline{X}-\overline{Y}}{\sqrt{1/n+1/m}\sqrt{\frac{nS_X^2+mS_Y^2}{n+m-2}}}\sim T_{n+m-2}$$

$$\delta=\begin{cases}0, & |d|<\tau_{1-\epsilon/2}\\ 1, & |d|\ge\tau_{1-\epsilon/2}\end{cases}$$

☝️ Подумай про квантили! Тут распределение Стьюдента, оно похоже на нормальное распределение, оно симметричное, значит можно взять один квантиль и взять по модулю **ИЛИ** с таким же успехом поставить двойное неравенство где есть еще квантиль $\tau_{\epsilon/2}=-\tau_{1-\epsilon/2}$.

Свойства

1. Критерий состоятельный.

1. **Доказательство**. Если средние разные, то $\overline{X} - \overline{Y}$ стремится к некоторой константе, $\sqrt{\frac{nS_X^2 + mS_Y^2}{n+m-2}}$ стремится к константе, $\sqrt{1/n + 1/m}$ стремится к нулю, то есть дробь стремится к бесконечности. А квантили стремятся к квантилям нормального распределения, то есть не к бесконечности.

2. Критерий имеет размер $1 - \epsilon$.
3. Реально допустимый уровень значимости: $\epsilon^* = P(\eta \geq d) = 1 - T_{n+m-2}(d)$.