Testování hypotéz o střední hodnotě ve výběrech z normálního rozdělení

Jednovýběrový t-test:

Data:

realizace náh. výběru $X_1, ..., X_n$, t. j. ***n* nezávislých pozorování náh. veličiny** $X \sim N(\mu, \sigma^2)$.

Hypotéza: H_0: $\mu = \mu_0$.

Oboustranná alternativa: H_1: $\mu \neq \mu_0$.

Bodový odhad pro μ: $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$.

Idea testu: $\bar{X}$ daleko od μ_0 => zamítáme H_0 ve prospěch H_1.

Intervalový odhad pro μ: I_μ $= [\bar{X} - \frac{S_n}{\sqrt{n}}\, t_{n-1}(1-\frac{\alpha}{2}); \bar{X} + \frac{S_n}{\sqrt{n}}\, t_{n-1}(1-\frac{\alpha}{2})]$,

kde $S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$ je bodový odhad rozptylu σ^2,

$t_{n-1}(1-\frac{\alpha}{2})$ je $(1-\frac{\alpha}{2})$ 100% - ní kvantil Studentova rozdělení t_{n-1}

a platí $P(I_\mu$ pokryje $\mu) = 1-\alpha$.

Nejčastěji volíme hladinu testu $\alpha = 0.05 = P$(zamítneme H_0, když platí).

Za platnosti H_0 máme: $P(\mu_0 \in I_\mu) = 1-\alpha$, a tedy $P(\mu_0$ neleží v $I_\mu) = \alpha$, odtud dostaneme

kritický obor testu: **μ_0 neleží v I_μ => zamítáme H_0** ve prospěch H_1 na hladině α.

Lze upravit na tvar: $|T_n| > t_{n-1}(1-\frac{\alpha}{2})$ => zamítáme H_0 ve prospěch H_1 na hladině α,

kde $T_n = \sqrt{n}\, \frac{1}{S_n} (\bar{X} - \mu_0)$ je testová statistika mající Studentovo rozdělení t_{n-1}.

<u>Ověření předpokladů</u>: nezávislost: intuitivně podle povahy dat,

normalita: histogram dat, test na normalitu dat (např. Shapiro-Wilk).

Nesplnění normality: 1-výběrový Wilcoxonův test.

Párový t-test:

Data:

nezávislá párová pozorování $(X_1,Y_1), \ldots, (X_n,Y_n)$,

taková, že $Z_i = X_i - Y_i$, $i = 1,\ldots, n$ tvoří výběr z $N(\mu,\sigma^2)$, kde $\mu = \mu_X - \mu_Y$.

Hypotéza: $H_0 : \mu_X = \mu_Y$, t.zn. $\mu = 0$.

Oboustranná alternativa: H_1: $\mu_X \neq \mu_Y$, t.zn. $\mu \neq 0$.

Provedení testu: aplikujeme jednovýběrový t-test na $Z_1, \ldots, Z_n$.

Dvouvýběrový t-test:

Data:

realizace 2 nezávislých náhodných výběrů $X_1, \ldots, X_m$ z rozdělení $N(\mu_X, \sigma^2)$

a $Y_1, \ldots, Y_n$ z rozdělení $N(\mu_Y, \sigma^2)$.

Hypotéza: H_0: $\mu_X = \mu_Y$, t.zn. $\mu_X - \mu_Y = 0$.

Oboustranná alternativa: H_1: $\mu_X \neq \mu_Y$.

Idea testu: $\bar{X}$ daleko od $\bar{Y}$ => zamítáme H_0 ve prospěch H_1.

Intervalový odhad pro $\mu_X - \mu_Y$: I = $\left[\bar{X} - \bar{Y} - \frac{S}{\sqrt{\frac{mn}{m+n}}}\, t_{m+n-2}(1 - \frac{\alpha}{2});\ \bar{X} - \bar{Y} + \frac{S}{\sqrt{\frac{mn}{m+n}}}\, t_{m+n-2}(1 - \frac{\alpha}{2})\right]$,

kde $S^2 = \frac{1}{m+n-2}\,[(m-1)S_X^2 + (n-1)S_Y^2]$ je vážený průměr odhadů rozptylu v 1. a v 2. výběru.

Kritický obor testu: **0 neleží v I => zamítáme H_0** ve prospěch H_1 na hladině α.

Lze upravit na tvar: $|T| > t_{m+n-2}(1 - \frac{\alpha}{2})$ => zamítáme H_0 ve prospěch H_1 na hladině α,

kde $T = \sqrt{\frac{mn}{m+n}}\,\frac{1}{S}\,(\bar{X} - \bar{Y})$ je testová statistika mající Studentovo rozdělení t_{m+n-2}.

Ověření předpokladů: nezávislost mezi výběry a uvnitř výběrů: intuitivně podle povahy dat,

shoda rozptylů: F-test (zamítá H_0: $\sigma_X^2 = \sigma_Y^2$, je li $\frac{S_X^2}{S_Y^2}$ daleko od 1),

normalita: histogramy nebo testy na normalitu v obou výběrech.

Nesplnění předpokladů:

normalita ANO → shoda rozptylů ANO → dvouvýběrový t- test,

normalita ANO → shoda rozptylů NE → Welchův test,

normalita NE → dvouvýběrový Wilcoxonův test.

Poznámky k normalitě:

1) Centrovaný a normovaný průměr $\sqrt{n}\ \frac{1}{\sigma}\ (\overline{X} - \mu)$ veličin s rozdělením $N(\mu,\sigma^2)$ má $N(0,1)$,

nahradíme-li směrodatnou odchylku σ jejím odhadem S_n , dostaneme $T_n \sim t_{n-1}$.

Proto se v kritických oborech t-testů používají kvantily rozdělení t_{n-1}.

Pro $n \to \infty$ je Studentovo rozdělení o n-1 stupních volnosti t_{n-1} přibližně stejné jako $N(0,1)$.

2) CLV => centrovaný a normovaný průměr $\sqrt{n}\ \frac{1}{\sigma}\ (\overline{X} - \mu)$ má přibližně rozdělení $N(0,1)$.

Proto **při velkém *n* nemá porušení normality zásadní vliv na fungování testu**,

testujeme pak na hladině přibližně rovné α . To platí pro jednovýběrový i dvouvýběrový t-test.

Podklady ze statistického softwaru k testování hypotéz :

Statistické softwarové produkty (R, Statistica, NCSS, SPSS ...) umožňují testovat

hypotézy o střední hodnotě i jiné - např. hypotézu, že

- data jsou realizací náhodného výběru z normálního rozdělení,
- dvě spojité veličiny jsou nekorelované,
- dvě diskrétní veličiny jsou nezávislé,
- střední hodnoty v $k > 2$ výběrech jsou stejné,
- chování náhodné veličiny ovlivňuje jeden nebo více faktorů.

Na výstupu ke každému testu je softwarem uvedena tzv.

***p*-hodnota**, což je **nejmenší hladina, na které zamítáme testovanou hypotézu** .

Provedení testu na hladině α pomocí p-hodnoty:

$\boldsymbol{p \le \alpha}$ => zamítáme $\mathbf{H_0}$ na hladině p a každé vyšší => **zamítáme $\mathbf{H_0}$ na hladině $\boldsymbol{\alpha}$**,

$\boldsymbol{p > \alpha}$ => zamítáme $\mathbf{H_0}$ až od hladiny p => **nezamítáme $\mathbf{H_0}$ na hladině $\boldsymbol{\alpha}$.**