Přednáška 6 – Testy hypotéz obecně. Testy jedné veličiny

Interval spolehlivosti (IS) – interval, ve kterém se nachází hodnota odhadovaného parametru θ s pravděpodobností $1-\alpha$

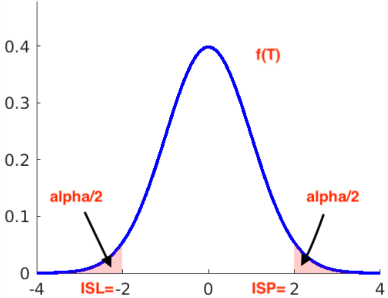
- Hranice – kvantil a kritická hodnota pro normované T a $f(T)$

α – hladina významnosti

$$\alpha = 0.05 \quad \text{nebo} \quad 0.01 \quad \text{nebo} \quad 0.1$$
$$1-\alpha = 0.95 \quad \text{nebo} \quad 0.99 \quad \text{nebo} \quad 0.90$$

$P(\hat{\theta} \in IS) = 1-\alpha$ (bílá plocha)
$P(\hat{\theta} < ISL) = \frac{\alpha}{2}, \quad P(\hat{\theta} > ISP) = \frac{\alpha}{2}$
(růžová plocha)

- oboustranný $IS = (ISL, ISP)$



- Hledáme μ, ale známe σ^2 – normální rozdělení $f(T)$
- Hledáme μ, neznáme σ^2 – Studentovo rozdělení $f(T)$
- Hledáme σ^2 – χ^2-rozdělení $f(T)$
- Hledáme podíl p – přibližně normální rozdělení $f(T)$

Příklad IS

- intervalový odhad ceny olejového filtru pro vozidlo Škoda Octavia na hladině významnosti 0.05
- z dlouhodobých měření víme, že rozptyl cen $\sigma^2 = 6$
- výběr cen z eshopů:

$$X = [100\ 154\ 110\ 137\ 119\ 109\ 138\ 164\ 136\ 85\ 138\ 105\ 121\ 149]$$

Konstrukce IS v software:

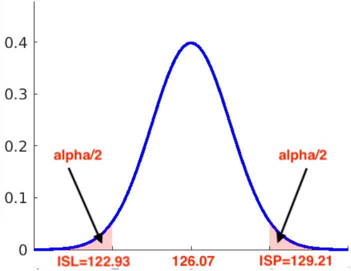
- α (pokud není, použije se 0.05)
- rozdělení souboru

předpoklad: cena $\sim N(\mu, \sigma^2)$

- $T = \bar{X} = 126.07$Kč
- $f(T)$ – normální rozdělení

(protože známe rozptyl souboru σ^2)

$$IS = (122.93\text{Kč},\ \ 129.21\text{Kč})$$



$$\bar{X} \in IS$$

Jednostranné IS – levostranný a pravostranný

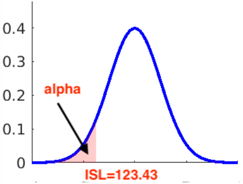
Odhad minimální ceny

- levostranný IS

(pouze levá hranice)

$IS = (ISL, \infty) = (123.43, \infty)$

$P(\hat{\theta} < ISL) = \alpha$



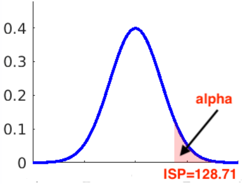
Odhad maximální ceny

- pravostranný IS

(pouze pravá hranice)

$IS = (-\infty, ISP) = (-\infty, 128.71)$

$P(\hat{\theta} > ISP) = \alpha$



Poznámka:

- některé software – pouze oboustranný a jednostranný IS
- jiná interpretace směru IS
- použití IS v rámci testování hypotéz

Testy hypotéz – obecně. Dvě hypotézy H_0 a H_A

Nulová hypotéza H_0

tvrzení podle úlohy

Příklad:

$$\theta = \hat{\theta}_0$$

P-hodnota – p_h, p_v, p_{value}

Pravděpodobnost toho, že platí nulová hypotéza H_0

Alternativní hypotéza H_A

popírá nulovou hypotézu H_0

Příklad:

$\theta \neq \hat{\theta}_0$ – oboustranný test

nebo

$\theta < \hat{\theta}_0$ – levostranný test

nebo

$\theta > \hat{\theta}_0$ – pravostranný test

Ptáme se, zda můžeme nulovou hypotézu H_0 na zvolené hladině významnosti α zamítnout (tzv. testujeme hypotézu H_0)

$T \notin IS$

$$p_h < \alpha$$

Zamítáme H_0 na hladině α
(nepotvrzujeme tímto H_A)

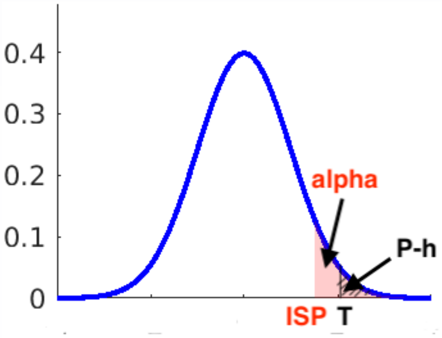
signifikantní výsledek

$T \in IS$

$$p_h > \alpha$$

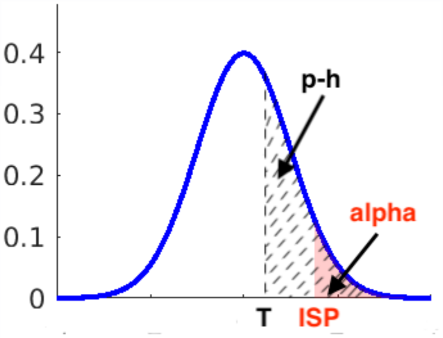
Nezamítáme nulovou hypotézu H_0
(nepotvrzujeme ji – pouze nevylučujeme)

Zamítáme H_0 vs. nezamítáme H_0



$p_h < \alpha$

Zamítáme H_0



$p_h > \alpha$

Nezamítáme H_0

- Závěr, zda zamítáme nulovou hypotézu – na základě p-hodnoty
- P-hodnota – statistický software

Postup při testování hypotéz

- správně zvolit test
- H_0 – co testujeme
- směr testu – určujeme podle H_A (neplatí pro všechny testy, některé mohou být pouze obou- nebo pouze jednostranné)
- hladina významnosti α (pokud není, použije se 0.05)

Dále – podle pravidla:

$$\text{Pokud} \quad p_h < \alpha, \quad \text{zamítáme} \quad H_0$$

$$\text{Pokud} \quad p_h > \alpha, \quad \text{nezamítáme} \quad H_0$$

Příklad TH – obecně

Sledujeme rychlost vozidel v tunelu Blanka na úseku s maximální povolenou rychlostí 70 km/h se směrodatnou odchylkou 1. Policie tvrdí, že řidiči systematicky porušují pravidla a jezdí na úseku vyšší rychlostí. Chceme otestovat toto tvrzení na hladině významnosti 0.05 za předpokladu normality dat.

Nulová hypotéza:

$H_0: \quad \mu = 70$ (nebo $\mu > 70$) – podle tvrzení

Alternativní hypotéza – opačné tvrzení:

$H_A: \quad \mu < 70$ – určuje, že test je levostranný

P-hodnota: $p_h = 0.0178644 < 0.05$ – zamítáme H_0.
Závěr: není pravda, že řidiči systematicky převyšují rychlost 70 km/h

Jaké testy budeme používat:

- Testy pro jeden výběr
- Testy pro dva výběry
- Testy více výběrů
- Testy nezávislosti
- Testy na vhodnost k regresní analýze
- Testy pro validaci regrese
- Testy pro diskrétní data

Testy hypotéz pro jeden výběr

- parametrické testy, kde pracujeme s daty, která pochází z normálního rozdělení
- neparametrické testy, kde data nepochází z normálního rozdělení (nesplňují předpoklad normality)
- testy rozdělení dat

Tabulka testů hypotéz:

http://staff.utia.cas.cz/uglickich/pdfka/volbaTH.pdf

Testy rozdělení

Nulová a alternativní hypotézy pro všechny testy rozdělení:

H_0 : data pochází z testovaného rozdělení

H_A : data nepochází z testovaného rozdělení

Pro ověření normality dat:

H_0 : data mají normální rozdělení

H_A : data nemají normální rozdělení

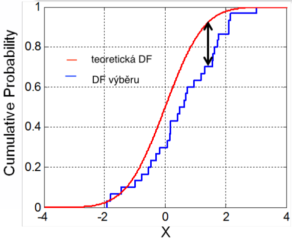
w/s test normality

$$T = \frac{w}{s} = \frac{\max(X) - \min(X)}{s}$$

Kolmogorov-Smirnovův test

- teoretické rozdělení výběru
- testování normality
- stejné rozdělení dvou výběrů
- pouze pravostranný

$$T = \max(F_{\text{výběr}}(X_i) - F_{\text{teor.}}(X_i))$$



Zdroj: Wikipedia

Test porovnává vzdálenosti mezi distribuční funkcí na hodnotách výběru (z četností) a distribuční funkcí teoretického rozdělení

Shapiro-Wilkův test normality

$$W = \frac{b^2}{(n-1)s^2}$$

$$b = a_1(r_n - r_1) + a_2(r_{n-1} - r_2) + \ldots$$

- a_i – ze speciální tabulky
- r_i – pořadí hodnot
- jeden z nejčastěji používaných testů normality

Anderson-Darlingův test

- teoretické rozdělení výběru + testování normality

$$A^2 = -n - S$$

$$S = \sum_{i=1}^{n} \frac{2i-1}{n} \left[\ln(F(X_i)) + \ln\left(1 - F(X_{n+1-i})\right)\right]$$

- jeden z nejsilnějších statistických testů rozdělení

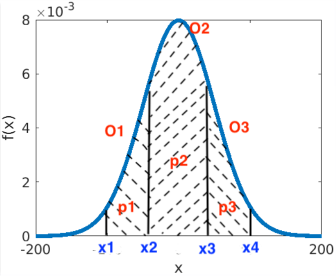
χ^2-test dobré shody

- teoretické rozdělení výběru
- testování normality

$$T = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

- O_i – pozorované četnosti (observed)
- E_i – očekávané četnosti (expected)

Ověření normality



Výpočet očekávaných četností E_i:

- pravděpodobnosti p_i

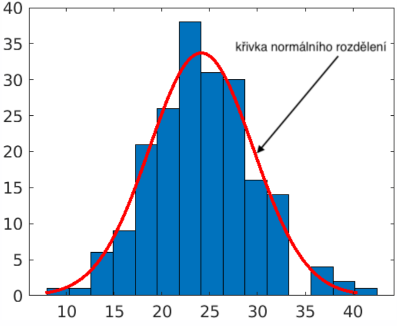
$$\begin{aligned} p_1 &= F(X_2) - F(X_1) \\ p_2 &= F(X_3) - F(X_2) \\ p_3 &= F(X_4) - F(X_3) \end{aligned}$$

$$E = [\underbrace{np_1}_{E_1} \quad \underbrace{np_2}_{E_2} \quad \underbrace{np_3}_{E_3}]$$

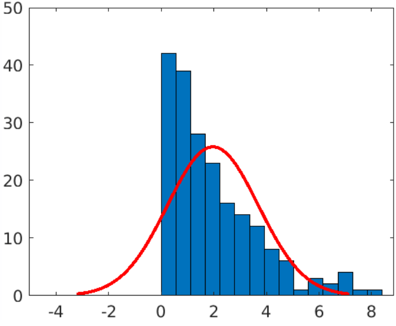
Grafické metody ověření normality

Histogram

Data **mají** normální rozdělení



Data **nemají** normální rozdělení

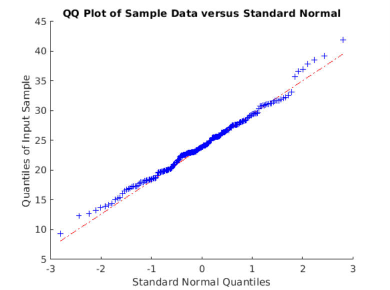


- předběžná analýza, není vhodný pro malý počet dat

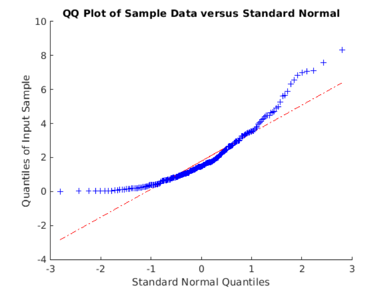
Q-Q graf

- porovnání rozdělení dvou výběrů – kvantil-kvantil graf

Data mají normální rozdělení



Data nemají normální rozdělení



Poznámka:

- jakýkoliv počet výběrů – testujeme každý výběr zvlášť
- neexistuje jeden univerzálně spolehlivý test normality
- software – několik druhů testů rozdělení
- uživatel – volba vhodného testu pro data
    - malé počty dat, velké hodnoty výběru, menší rozptyl atd.
- cvičení v Matlabu – Anderson-Darlingův test