

Testování hypotéz:

Př: Test správné míry:

hypotéza H_0 : Ano, dostali jsme $\mu = 0,5\%$

hypotéza H_1 : Ne

$$X \sim N(\mu, \sigma^2) \quad \rightarrow \text{známe}$$

$$H_0: \mu = \mu_0$$

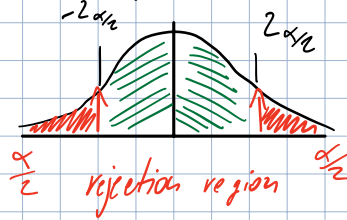
Máme $X_1, \dots, X_n$ n.h. měření $\sim N(\mu, \sigma^2)$
tohle chceme zjistit

Stavíme statistiku, kterou budeme používat:

$$T := \frac{\bar{X}_n - \mu_0}{\sigma / \sqrt{n}}$$

Proč zvolila takhle?

Blud $\mu = \mu_0$, pak $T \sim N(0, 1)$



podud $|T| > 2\alpha/2$, zamítáme.

Tohle je ten rozdíl,
ve kterém přijímáme,
že náhodná veličina je
v tabulce.

Př: Test nůžování

$$X_1, \dots, X_n \sim N(\mu_1, \sigma_1^2)$$

$$Y_1, \dots, Y_m \sim N(\mu_2, \sigma_2^2)$$

$$H_0: \mu_1 = \mu_2 \quad ?$$

$$H_1: \mu_1 \neq \mu_2 \quad ?$$

Zavedeme statistiku:

$$T = \frac{\bar{X}_n - \bar{Y}_m}{\sigma \cdot \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim N(0, 1)$$

$$ET = \frac{E\bar{X}_n - E\bar{Y}_m}{\dots} = 0$$

$$\text{var } T = \frac{1}{\sigma^2 \left(\frac{1}{n} + \frac{1}{m} \right)} \cdot \left(\underbrace{\text{var } \bar{X}_n}_{\frac{\sigma^2}{n}} + \underbrace{\text{var } \bar{Y}_m}_{\frac{\sigma^2}{m}} \right) = 1$$

Tím pádem potom provádím
stejný test jako v minulém případě,
protože jsem uvažoval, že jde o $N(0, 1)$

Test dobré shody

Test dobré shody

Pr. Házíme kostkou 600x

	1	2	3	4	5	6
# shůt	92	120	88	98	95	107
# teorie	100	100	100	100	100	100

Max. věrohodnost (max. likelihood)
Likelihood ratio
např. data $x_1, \dots, x_k$
 $(n, \vartheta_1, \vartheta_2, \dots, \vartheta_k)$
 ϑ

$$P(X_1=x_1, \dots, X_k=x_k) = \binom{n}{x_1, \dots, x_k} \vartheta_1^{x_1} \dots \vartheta_k^{x_k}$$

$$H_0: \vartheta = (\frac{1}{6}, \dots, \frac{1}{6}) = \vartheta^*$$

$$H_1: \vartheta \neq \vartheta^*$$

Max. věrohodnost (max. likelihood)

Likelihood ratio

např. data $x_1, \dots, x_k$

$$\frac{L(x, \hat{\vartheta})}{L(x, \vartheta^*)}$$

MAX. VĚROHODNOST

POZOR. VĚROHODNOST

ZA PŘEDP. H_0

$$L(x, \vartheta)$$

$$O_i = x_i, E_i = n\vartheta_i^*$$

$\hat{\vartheta}$ takové ϑ , že $L(x, \vartheta)$ je maximální

$$\hat{\vartheta}_i = \frac{x_i}{n}$$

$$G = 2 \log \frac{L(x, \hat{\vartheta})}{L(x, \vartheta^*)} = 2 \log \prod_{i=1}^k \frac{\left(\frac{x_i}{n}\right)^{x_i}}{\left(\frac{1}{6}\right)^{x_i}} = 2 \sum_{i=1}^k x_i \log \frac{x_i}{n\vartheta_i^*} = 2 \sum O_i \log \frac{O_i}{E_i}$$

Taylor. polynom

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

χ^2 -kritérium

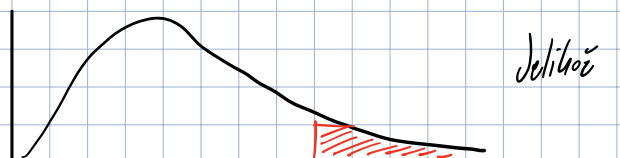
$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{8^2 + 20^2 + 12^2 + 2^2 + 5^2 + 7^2}{100} = 6,86$$

$\chi^2 = 0 \dots$ dokonalá shoda

Je potřeba uvažovat, jestli to je ale nebo NE

To nám určí χ^2 -kritérium s F dist. fci a Q kvantilem.

Chci přesnost 95% $\rightarrow Q(0,95) = 11,1$



$\hookrightarrow$ zamítnutí region

Jelikož $6,86 < 11,1$, tak přijmeme.

Tudíž kostka je šetrná.

$$1 - F(6,86) = 0,23$$

Rejection - Sampling

