

Testování hypotéz:

Př.: Test správné míry:

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

hypotéza H_1 : Ne

$$X \sim N(\nu, \sigma^2) \quad \text{--- } \rightarrow \text{2 měř.}$$

$$H_0: \nu = \mu$$

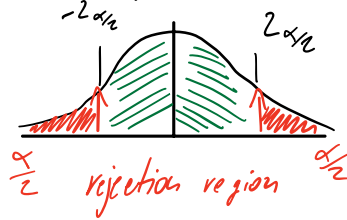
Máme X_1, \dots, X_n n. měření $\sim N(\nu, \sigma^2)$
tohle chceme zjistit

Stavování statistiky, kterou budeme používat:

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

Proč zvolim takhle?

Bloud $\nu = \mu$, pak $T \sim N(0, 1)$



Tohle je ten rozsah,
ve kterém přijmeme,
že vícepravou vlnu je
v folovanci.

podud $|T| > 2 \cdot \alpha/2$, zamítáme.

Př.: Test nadržování

$$X_1, \dots, X_n \sim N(\nu_1, \sigma_1^2)$$

$$Y_1, \dots, Y_m \sim N(\nu_2, \sigma_2^2)$$

$$H_0: \nu_1 = \nu_2 ?$$

$$H_1: \nu_1 \neq \nu_2 ?$$

Zmoudeme 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 ukázal, že jde o $N(0, 1)$

Test dobré shody

Test dobré shody

Pr. Házíme kostkou 600x

	1	2	3	4	5	6
# shod	92	120	88	98	95	107
# kousků	100	100	100	100	100	100
Mitt. hodnota	$\frac{92}{100}$	$\frac{120}{100}$	$\frac{88}{100}$	$\frac{98}{100}$	$\frac{95}{100}$	$\frac{107}{100}$

Obs. $E(\pi)$

Max. věrohodnost (max. likelihood)

Likelihood ratio

např. data x_1, \dots, x_k

$P(X_1=x_1, \dots, X_k=x_k) = \binom{n}{x_1, \dots, x_k} \vartheta_1^{x_1} \dots \vartheta_k^{x_k}$
 $H_0: \vartheta = (\frac{1}{6}, \dots, \frac{1}{6}) = \vartheta^*$
 $H_1: \vartheta \neq \vartheta^*$

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

MAX. VĚROHODNOST

POZOR. VĚROH. ZA PŘEDP. H_0

$(n, \vartheta_1, \vartheta_2, \dots, \vartheta_k)$
 ϑ

(X_1, \dots, X_k) s parametry

$L(x, \vartheta)$

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

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

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

Taylor. polynom $\approx \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} = \chi^2$ (χ^2 -krit. řešení)

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

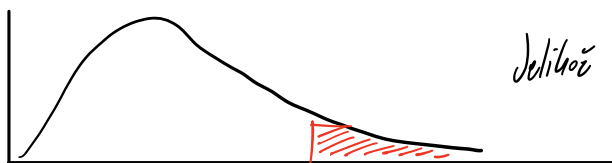
$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{8^2 + 20^2 + 12^2 + 2^2 + 5^2 + 7^2}{100} = 6,86$$

$\chi^2 = 0 \dots$ dokonalá shoda

Je potřeba vědět, jestli to je OK nebo NE

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

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



Jelikož $6,86 < 11,1$, tak přijmeme.

Tudíž kostka je férová.

\hookrightarrow zamítnutí region

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

Rejection - Sampling

