

# Testování hypotéz:

**Př:** Test správné míry:

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

hypotéza  $H_1$ : Ne

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

$$H_0: \nu = \mu$$

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

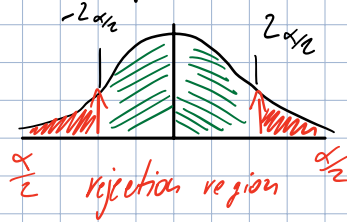
Stavujeme statistiku, kterou budeme používat:

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

Proč zvolila tahle?

Tohle je ten rozsah,  
ve kterém přijímáme,  
že náhodnou velič. je  
v tabulce.

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



podud  $|T| > 2 \cdot \frac{\alpha}{n}$ , zamítáme.

**Př:** Test nůdroování

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

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

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

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

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 ukázal, že jde o  $N(0, 1)$

# Test dobré shody

## Test dobré shody

Př. 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.  $E(x)$   
 $(n, \vartheta_1, \vartheta_2, \dots, \vartheta_k)$   
 $\vartheta$

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

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

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

Max. věrohodnost (max. likelihood)

Likelihood ratio

naměřením data  $x_1, \dots, x_k$

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

MAX. VĚROHODNOST

POZOR. VĚROHODNOST

ZA PŘEDP.  $H_0$

$$L(x_i, \vartheta)$$

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

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

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

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

Taylor. polynom

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

$\chi^2$ -krit. řešení

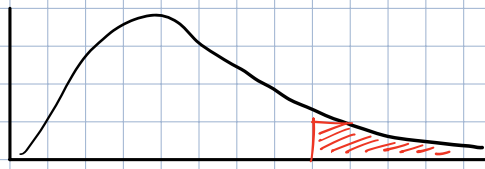
$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{8^2 + 20^2 + 12^2 + 2^2 + 5^2 + 7^2}{100} = 6,86$$

$\chi^2 = 0 \dots$  dokonalá shoda

Je potřeba určit, jestli to je AN nebo NE

To nám určí  $\chi^2$ -krit. s F dist. fci a Q kumulac.

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



$\leftarrow$  zamítací region

Jelikož  $6,86 < 11,1$ , tak přijímáme.

Tudíž kostka je férová!

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



# Rejection - Sampling

