TESTOVÁNÍ HYPOTÉZ II

H_0	pravdivá	nepravdivá
nezamítnuta	správně	chyba 2. druhu (β)
zamítnuta	chyba 1. druhu (α)	správně

$$P(X \in W \mid H_0) \leq \alpha$$

X – množina možných výsledků)

Příklad. Testování preventivního účinku séra.

$0{,}25$

$n = 30$ očkovaných zvířat

3 zvířata onemocní

$H_0: p = 0{,}25$

$H_1: p < 0{,}25$

$X = \{0, 1, \dots, 30\}$ mohla onemocnět

$$P(X \le 3 \mid H_0) = P(X=0 \mid H_0) + P(X=1 \mid H_0) + P(X=2 \mid H_0) + P(X=3 \mid H_0) = 0{,}75^{30} + \binom{30}{1} 0{,}25 \cdot 0{,}75^{29} + \binom{30}{2} 0{,}25^{2} \cdot 0{,}75^{28} + \binom{30}{3} 0{,}25^{3} \cdot 0{,}75^{27} \doteq$$

$$\doteq 0{,}0268 + 0{,}0086 + 0{,}0018 + 0{,}0002 < 0{,}05$$

$\Rightarrow H_0$ zamítáme na hladině významnosti: 0,05

TESTY O PARAMETRECH NORMÁLNÍHO ROZDĚLENÍ

$(X_1, X_2, \ldots, X_n)$ je náhodný výběr z $N(\mu, \sigma^2)$. Necht σ^2 je známý parametr.

$H_0: \mu = \mu_0 \qquad H_1: \mu \neq \mu_0$

$$W = \left\{ x \in R^n \,\middle|\, |\bar{x} - \mu_0| \geq \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}} \right\},$$

prostě

$$P\left(|\bar{X}-\mu|<\frac{\sigma}{\sqrt{n}}u_{1-\alpha/2}\right)=1-\alpha$$

Kdyby

$$H_0:\mu=\mu_0,\quad H_1:\mu>\mu_0$$

$$W=\left\{x\in R^n \,\middle|\, \bar{x}-\mu_0\geq\frac{u_{1-\alpha}}{\sqrt{n}}\cdot\sigma\right\}$$

Příklad: 336 kg

dodávka: 144 výrobků, $\bar{x} = 344$ kg

$s = 52$ kg (neznáme σ)

Práce: $H_0: \mu = 336$, $H_1: \mu > 336$

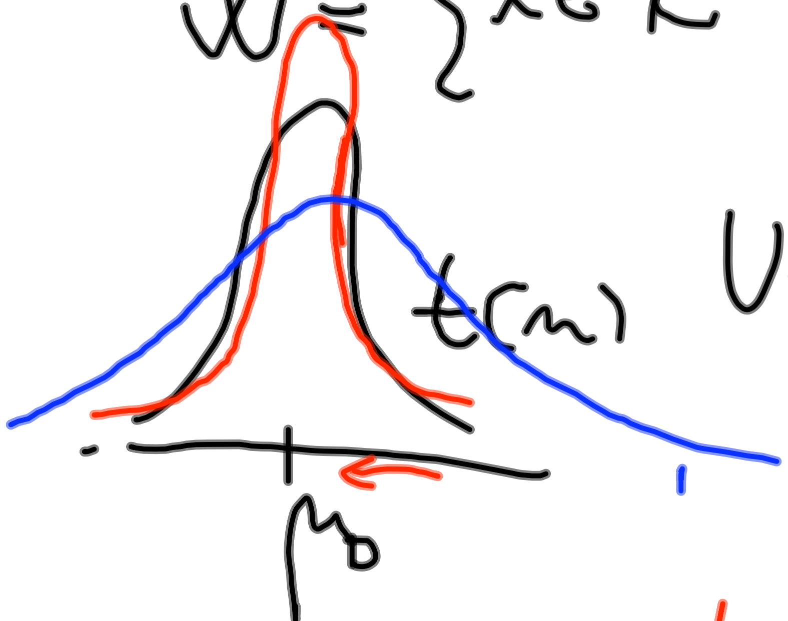
$$(\mu - \mu_0 <) \quad \frac{344 - 336}{\frac{52}{\sqrt{144}}} \doteq 1{,}85 > 1{,}64$$

$u_{0{,}095}$

$\alpha = 0{,}05$

Matematicky rigorózní postup.

$$W = \left\{ x \in \mathbb{R}^n \,\middle|\, \bar{x} - \mu_0 \geq \frac{s}{\sqrt{n}} \cdot t_{1-\alpha}(n-1) \right\}$$



Výpočet: $\dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} \doteq 1{,}85 > 1{,}65$

$t_{0,95}(143) = ?$

$t_{0,95}(100) \doteq 1{,}660$

$n \to \infty \quad 1{,}64$

ROZPTYL

$H_0: \sigma \leq \sigma_0^2$; $H_1: \sigma > \sigma_0^2$

$$W = \left\{ x \in R^n \,\middle|\, \frac{\sum_1^n (x_i - \mu)^2}{\sigma_0^2} \geq \chi^2_{1-\alpha}(n) \right\},$$

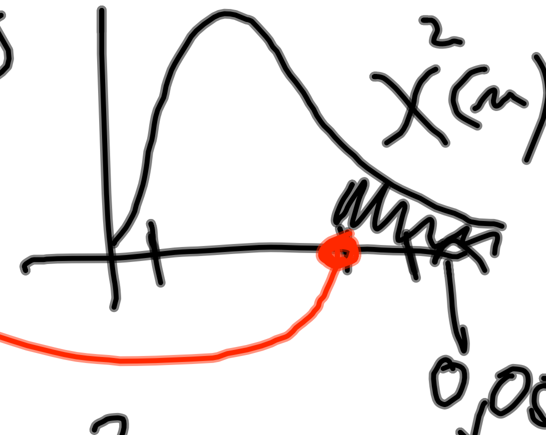
je-li známé μ.

Není-li známo μ, pak

$$W = \left\{ x \in R^n \,\middle|\, \frac{\sum_1^n (x_i - \bar{x})^2}{\sigma_0^2} \geq \chi^2_{1-\alpha}(n-1) \right\}.$$

Příklad: 25 výrobků, $\bar{x} = 95$, $s^2 = 88$

$$H_0: \sigma^2 \leq 80, \quad H_1: \sigma^2 \overset{\neq}{>} 80$$



$$\frac{24 \cdot 88}{80} \doteq 26{,}4 \;\begin{matrix}\leq \\ >\end{matrix}$$

$$26{,}4 \in \left(\chi^2_{0{,}025}(24), \chi^2_{0{,}975}(24)\right)$$

$$26{,}4 \in (12{,}4;\ 39{,}4)$$

TESTY SHODY

1.) χ^2 – test

$$\chi^2_n = \sum_{j=1}^{k} \frac{(n_j - n \cdot p_j)^2}{n p_j}$$

$k \leq n$ počet tříd

p_j – pravd. výskytu v j-té třídě

n_j – absolutní přečetnost j-té třídy

Pleade!: $\lim_{n\to\infty} P(\chi^2_n < r) =$

$= P(\chi^2(k-1) < r)$

Příklad.

typ půdy	počet sídel	% plochy	očekávaný počet sídel
I	26	32	17,0
II	9	25	13,2
III	18	43	22,8
	53	100	53,0

$$\chi^2_2 = \sum_{j=1}^{3} \frac{(n_j - np_j)^2}{np_j} =$$

$$= \frac{(26-17)^2}{17{,}0} + \frac{(9-13{,}2)^2}{13{,}2} +$$

$$+ \frac{(18-22{,}8)^2}{22{,}8} \doteq 7{,}11 > 5{,}99$$

$n=2;\ \alpha=0{,}05$

2.) Kolmogorov – Smirnov

$\alpha = 0{,}05$

$$1{,}36\sqrt{\frac{n_1+n_2}{n_1\cdot n_2}} \geq$$ největší absolutní hodnota „rozdílu dat"

n_1 – četnost 1. vzorku

n_2 – četnost 2. vzorku

$\alpha = 0{,}1$

$$1{,}63\sqrt{\frac{n_1+n_2}{n_1 n_2}}$$

H_0: není rozdíl mezi bohatými a chudými

Příklad.

věková kategorie	bohatí	kum. čet.	chudí	kum. fr.
dětský věk I	6	0,079	23	0,169
– 〃 – II	8	0,105	21	0,154
mládí	11	0,145	25	0,184
střední věk	29	0,382	36	0,265
starší střední	19	0,250	27	0,199
starý	3	0,039	4	0,029
	76		136	

Věk	bohatí	chudí	rozdíl
1.	0,079	0,169	0,090
2.	0,184	0,323	0,139
3.	0,329	0,507	0,178
4.	0,711	0,772	0,061
5.	0,961	0,971	0,010
6.	1,000	1,000	0,000

srovnání s kvantilem:

$$1{,}36 \cdot \sqrt{\frac{76+136}{78 \cdot 136}} \doteq 0{,}195 > 0{,}178$$

STEPHAN SHENNAN:
QUATIFYING ARCHEOLOGY

Edinburgh University Press
1988