W MŠ 22.5.2017

Pr. 1 *a) Aké chyby poznáme pri testovaní štatistických hypotéz a aký je vzťah medzi nimi? Ako súvisí hladina významnosti testu s niektorou z chýb? (5b)*
b) Pri testovaní štatistických hypotéz pre rôzne hladiny významnosti $\alpha \in \{0,1,\ 0,05,\ 0,01\}$ boli vypočítané nasledujúce kritické oblasti (oblasti zamietania H_0): $(-\infty,\ 1,2)$, $(-\infty,\ 1,6)$, $(-\infty,\ 0,8)\}$ Priraďte oblastiam správne hladiny významnosti. (Bez zdôvodnenia rozhodnutia 1b.)

Riešenie 1 *a) hladina významnosti testu je α. Ak zmenšíme α zväčšíme β a naopak.*

výsledok testu/skutočnosť	H_0 platí	H_0 neplatí
H_0 zamietame	I. druh = α	$\times$
H_0 nezamietame	$\times$	II. druh= β

b) Ak K_α je oblasť zamietania H_0 na hladine významnosti α, tak platí $P(T \in K_\alpha) = \alpha$, kde T je štatistika, tak pre $K_{\alpha_1} \subseteq K_{\alpha_2}$ platí $\alpha_1 \leqslant \alpha_2$. Stačí obrázok.

$$\alpha = 0,1 \Rightarrow (-\infty,\ 1,6)$$
$$\alpha = 0,05 \Rightarrow (-\infty,\ 1,2)$$
$$\alpha = 0,01 \Rightarrow (-\infty,\ 0,8)$$

Pr. 2 *Nech $X_1, \cdots, X_n$ je náhodný výber, $E(X_i) = \mu$ a $D(X_i) = \sigma^2$ a $\overline{X} = \frac{1}{n}\sum_i X_i$.*
a) Nech $Y = \sum_i a_i X_i$, $a_i \in < 0, 1 >$. Pre aké a_i $E(Y) = \mu$ a $D(Y)$ je najmenšia? (Bez zdôvodnenia (1b), ináč (5b).)
b) Ak $X_i \sim N(\mu,\ \sigma^2)$. Aké rozdelenie má $\overline{X}$ a aké sú jeho parametre? (2b)
c) Ak $X_i \sim Bi(n,\ p)$, potom pomocou momentovej metódy odhadnite parameter p. (3b)

Riešenie 2 *a)*

$$\mu = E(Y) = \sum_i a_i E(X_i) = \mu \sum_i a_i \quad \Rightarrow \quad \sum_i a_i = 1$$
$$\sigma^2 = D(Y) = \sum_i a_i^2 D(X_i) = \sigma^2 \sum_i a_i^2 \quad \& \quad \sum_i a_i = 1$$

Riešime viazaný extrém funkcie viac premenných: hľadáme minimum funkcie

$$f(a_1, \cdots, a_n) = \sigma^2 \sum_i a_i^2 \quad \textit{za podmienky} \quad 1 - \sum_i a_i = 0$$

Mne by stačilo, keby vedeli naformulovať viazaný extrém, nepredpokladám, že by to vedelo mnoho študentov vyriešiť.

$$L(a_1, \cdots, a_n, \lambda) = f(a_1, \cdots, a_n) - \lambda(1 - \sum_i a_i)$$
$$\frac{\partial L(a_1, \cdots, a_n, \lambda)}{\partial a_i} = 2\sigma^2 a_i - \lambda = 0$$
$$\Rightarrow \quad a_i = \frac{\lambda}{2\sigma^2} \quad \forall i.$$

Po dosadení do podmienky dostaneme:

$$1 = \sum_i a_i = \sum_i \frac{\lambda}{2\sigma^2} = \frac{n\lambda}{2\sigma^2}$$
$$1 = \frac{n\lambda}{2\sigma^2} \quad \Rightarrow \quad \lambda = \frac{2\sigma^2}{n}$$

Teda

$$a_i = \frac{\frac{2\sigma^2}{n}}{2\sigma^2} = \frac{1}{n}$$

$a_i = 1/n$
b) $\overline{X} \sim N(\mu,\ \frac{\sigma^2}{n})$ (2b)
c) Ak $X_i \sim Bi(n,\ p)$, $E(X) = np = \mu \approx \overline{X}$.

$$np = \mu \quad \Rightarrow \quad p \approx \frac{\overline{X}}{n}$$

---

počet kliknutí	1	2	3	4	5	6	15
počet zákazníkov	1	26	24	10	15	20	4

Pr. 3 *Firma sledovala na webe počet kliknutí na svoju reklamu počas dňa u 100 náhodne vybraných zákazníkov. Nájdite dolný a horný kvartil, medián, modus a zostrojte krabicový (box) graf. Nájdite extremálne hodnoty (outlayers), ak existujú. Výsledky štatistického prieskumu sú v nasledujúcej tabuľke:*

Riešenie 3 *Modus* $Mo = 2$ *Median* $Me = 3$, $Q_L = 2$, $Q_U = 5$, $R = 5 - 3 = 2$ $1,5R = 3$, $Q_L - 1,5R = -1 < minx = 1$, $Q_U + 1,5R = 5 + 3 = 8 < 15 \quad \Rightarrow \quad 15$ *je vybočujća hodn.*

Pr. 4 *Generátor náhodných čísel vygeneroval* 100 *čísel z množiny* $\{0,1,2,3\}$. *Otestujte na hladine významnosti* $\alpha = 0,05$, *či ide o náhodný výber z* $Bi\left(3,\, \frac{1}{2}\right)$. *Výsledky sú v nasledujúcej tabuľke:*

číslo	0	1	2	3
početnosť	20	20	50	10

číslo	0	1	2	3
počet.ť f_i	20	20	50	10
očakávaná np_i	$100\frac{1}{8} = 12,5$	$100\frac{3}{8} = 37,5$	$100\frac{3}{8} = 37,5$	$100\frac{1}{8} = 12,5$
$(f_i - np_i)^2$	$7,5^2$	$(-17,5)^2$	$(-12,5)^2$	$(-2,5)^2$
$\frac{(f_i-np_i)^2}{np_i}$	$7,5^2/12.5$	$17,5^2/37,5$	$12,5^2/37,5$	$2,5^2/12,5$
$\frac{(f_i-np_i)^2}{np_i}$	$4,5$	$8,167$	$4,167$	$0,5$

Riešenie 4 H_0 *dáta majú* $Bi(3,\,1/2)$ H_1 *dáta nemajú* $Bi(3,\,1/2)$
dosadíme do vzorca:

$$\chi^2 = \sum_i \frac{(f_i - np_i)^2}{np_i} = 17,334$$

tab. hod. $\chi^2(0,05, \quad ,3) = 7,815$ *Pretože* $17,334 > 7,815$ H_0 *zamietame.*

Pr. 5 *V meste LM sa vyrába pivo Uf. V náhodnom výbere 100 domácností mesta LM sa sledovala spotreba Uf v priebehu roka. Z náhodného výberu sa zistil výberový priemer* $\overline{x} = 18l$ *a smerodajná odchýlka* $S = 8,5\ l$. *Priemerná ročná spotreba Uf na jednu domácnosť v SR je 17,8 l a smerodajná odchýlka* $\sigma = 6l$. *Na hladine vznamnosti* $\alpha = 0,05$*:*
a) otestujte či rozptyl je v LM je taký ako v SR alebo väčší.
b) otestujte či priemerná spotreba Uf v LM je taká istá ako v SR, alebo rôzna.
V oboch prípadoch naformulujte H_0 *a* H_1.

Riešenie 5 *a)* H_0 $\sigma^2 = 6^2$ *oproti* H_1 $\sigma^2 > 6^2$
Dosadíme do vzorca

$$U = \frac{(n-1).S^2}{\sigma_0^2} \sim \chi^2(n-1)$$

Pretože $U = \frac{99.8,5^2}{36} = \frac{7152,75}{36} = 198,688$ *Tab.hodn.* $113,145$. *Pretože* $198,688 > 113,145$, H_0 *zamietame.*
b) H_0 $\mu = 17,8$ *oproti* H_1 $mu \neq 17,8$
Štatistika

$$T = \frac{\sqrt{100}(\overline{x} - 17,8)}{8,5} \sim t(99).$$

$$T = \frac{10(18 - 17,8)}{8,5} = 0,235.$$

tab.hod. $T(99,\, 0,975) = 1,984$. *Pretože* $|0,235| = 0,235 < 1,984$, H_0 *nezamietame.*

Pr. 6 *Dvaja hráči hrajú Človeče nezlob se. Hodia naraz dvomi kockami. Pre každého z nich platia iné pravidlá. Ak obe čísla sú párne X zostáva stáť a Y sa vráti späť o jedno políčko, ak jedno je párne a druhé nepárne X ide dopredu o jedno polícko a Y o dve políčka a ak obe sú nepárne, tak X ide dopredu o dve a Y o jedno. Nájdite združené a marginálne rozdelenia pravdepodobnosti náh. vek.* (X, Y) *a vypočítajte* $\rho(X, Y)$. *Majú obaja rovnakú šancu vyhrať?*

Riešenie 6

$$\begin{aligned} E(X) &= 1/2 + 1/2 = 1, \quad E(Y) = -1/4 + 1 + 1/4 = 1, \\ E(X^2) &= 1/2 + 1 = 3/2, \quad E(Y^2) = 1/4 + 2 + 1/4 = 5/2, \\ E(XY) &= 1 + 1/2 = 3/2. \end{aligned}$$

Ω	PP	NP, PN	NN
pravd.	1/4	1/2	1/4
X	0	1	2
Y	-1	2	1
X^2	0	1	4
Y^2	1	4	1
XY	0	2	2

$$\begin{aligned} D(X) =& 3/2-1 = 1/2, \quad D(Y) = 5/2-1 = 3/2, \quad cov(X,Y) = 3/2-1 = 1/2, \\ \rho(X,Y) =& \frac{\frac{1}{2}}{\sqrt{\frac{1}{2}\frac{3}{2}}} = \frac{1}{\sqrt{3}}. \end{aligned}$$

Združené rozd.

X/Y	0	1	2
-1	1/4	0	0
2	0	1/2	0
1	0	1	1/4

Pr. 7 *N.p. X má funkciu hustoty $f_X(t)$*

$$f_X(t) = \begin{cases} \frac{t}{2} & 0 \leqslant t \leqslant a \\ \frac{1}{6}(4-t) & a < t \leqslant 4 \\ 0 & t \notin < 0,\, 4 > . \end{cases}$$

Vypočítajte čísla a, ak $F_X(a) = 0,25$. Načrtnite graf funkcie hustoty n.p. X, vypočítajte $P(X \geqslant 1,5)$ a vyznačte na grafe funkcie hustoty.

Riešenie 7

$$0,25 = F_X(a) = \int_0^a \frac{t}{2}\, dt = \left[\frac{t^2}{4}\right]_0^a = \frac{a^2}{4} \qquad \Rightarrow \qquad a = 1$$

$$P(X \geqslant 1,5) = \frac{1}{6}\int_{1,5}^4 (4-t)\, dt = \frac{1}{6}\left[4t - \frac{t^2}{2}\right]_{1,5}^4 = \frac{1}{6}\,(8-3,75) = \frac{4,25}{6} = 0,708$$