Statistika

(D360P03Z, D360P03U)

Karel Zvára

15. listopadu 2004

připomeňme příklad s hrací kostkou

- odhadujeme pravděpodobnost šestky
- kostka A: $n = 100, n_A = 17, f_A = 0{,}17, \Rightarrow 95\%$ int. spol.$(0{,}10; 0{,}24)$
- kostka B: $n = 100, n_B = 41, f_B = 0{,}41 \Rightarrow 95\%$ int. spol.$(0{,}31; 0{,}51)$
- důležitý rozdíl: u kostky A patří $1/6 = 0{,}167$ do intervalu spolehlivosti; u kostky B nikoliv
- co se dá z tohoto zjištění usoudit?
- použijeme **testování hypotéz**

testování hypotéz (1)

- **(nulová) hypotéza** H_0: – zjednodušuje situaci, zpravidla se jí snažíme vyvrátit, abychom věcně něco prokázali
- **alternativa** H_1: **(alternativní hypotéza)** – opak nulové hypotézy, zpravidla to, co chceme dokázat
- možná rozhodnutí
  - **zamítnout** H_0 pokud naše data svědčí proti H_0
  - **nezamítnout** H_0 (přijmout H_0) pokud *není dost důvodů* H_0 zamítnout
- nelze zaručit bezchybnost rozhodnutí

testování hypotéz (2)

- protože nelze zaručit bezchybnost rozhodnutí, mohou nastat chyby:
  - **chyba 1. druhu**, když zamítneme platnou hypotézu
  - **chyba 2. druhu**, když nepoznáme, že hypotéza neplatí a nezamítneme ji
- nechceme často *chybně* zamítat H_0 (falešně něco věcně prokazovat), proto zvolíme nízkou hladinu testu α (nejčastěji $\alpha = 5\%$)
- **hladina testu** α = maximální přípustná pravděpodobnost chyby 1. druhu
- **síla testu** = pravděpodobnost správného zamítnutí neplatné hypotézy

schéma testování hypotéz

rozhodnutí	H_0 platí	H_0 neplatí
H_0 zamítnout	chyba 1. druhu (pst $\leq \alpha$) hladina testu	správné rozhodnutí (pst $1-\beta$) síla testu
H_0 nezamítnout (přijmout)	správné rozhodnutí (pst $\geq 1-\alpha$)	chyba 2. druhu (pst β)

postup při rozhodování

- zvolit hypotézu H_0, alternativu H_1
- zvolit hladinu testu α
- zvolit metodu rozhodování (test)
- z dat spočítat testovou statistiku T a porovnat ji s tabelovanou kritickou hodnotou
- když padne T do **kritického oboru**, pak H_0 zamítnout (zpravidla, když $T \geq t_0$, t_0 – kritická hodnota)
- **kritický obor** – množina těch výsledků pokusu (např. hodnot T), kdy budeme hypotézu zamítat

příklad: padá na kostce šestka příliš často?

- chceme na 5% hladině prokázat, že pravděpodobnost šestky je velká (tj. větší než 1/6)
- H_0 : P(padne šestka) $= 1/6 \qquad (= \pi_0)$
- H_1 : P(padne šestka) $> 1/6 \qquad (\neq \pi_0)$
- co svědčí pro neplatnost hypotézy?
„šestka padá mnohem častěji, než by měla“
- provedeme $n = 100$ pokusů, Y počet šestek
- hypotézu budeme zamítat, když $Y > y_0$
- za platnosti H_0 má počet šestek Y rozdělení bi$(n, 1/6)$
- y_0 zvolit tak, aby za hypotézy bylo P$(Y > y_0) \leq \alpha$

příklad přesné volby kritického oboru

y_0	19	20	21	22	23
P$(Y > y_0)$	0,220	0,152	0,100	0,063	0,038

- podmínku P$(Y > y_0) \leq 0{,}05$ splňuje $y_0 = 23$
- padne-li ve 100 nezávislých hodech kostkou více než 23 šestek, budeme na **5% hladině zamítat hypotézu**, že pst šestky je 1/6 **ve prospěch alternativy**, že pst šestky je větší než 1/6 (dáno zvolenou alternativou)
- padlo nám $Y = Y_0 = 17$ šestek, hypotézu nezamítáme, což neznamená, že bychom hypotézu prokázali
- pro $\alpha = 10$ % bychom zvolili $y_0 = 21$

příklad: volba kritického oboru (přibližně)

- použijme přibližné tvrzení za H_0 $Y \sim N(n\pi_0, n\pi_0(1-\pi_0))$, potom

$$
\begin{aligned}
P(Y > y_0) &= 1 - P(Y < y_0) \\
&= 1 - P\left(\frac{Y - n\pi_0}{\sqrt{n\pi_0(1-\pi_0)}} < \frac{y_0 - n\pi_0}{\sqrt{n\pi_0(1-\pi_0)}}\right) \\
&\doteq 1 - \Phi\left(\frac{y_0 - n\pi_0}{\sqrt{n\pi_0(1-\pi_0)}}\right) = \alpha (= 0{,}05)
\end{aligned}
$$

- tabulka kritických hodnot dá $z(\alpha)$, musí platit $z(\alpha) = \frac{y_0 - n\pi_0}{\sqrt{n\pi_0(1-\pi_0)}}$ tedy
  $y_0 = n\pi_0 + z(\alpha)\sqrt{n\pi_0(1-\pi_0)}$, v našem příkladu
  $y_0 = 100/6 + 1{,}645 \cdot \sqrt{500/36} \doteq 23$

p-hodnota

- spočítat k T_0 odpovídající p-hodnotu a porovnat ji s α
- p-**hodnota** p je nejmenší α, při kterém H_0 z daných dat ještě zamítáme
- p-hodnota p je za platnosti H_0 spočítaná *pravděpodobnost* výsledků stejně nebo *méně příznivých* pro H_0
- zamítnout H_0, když je $p \leq \alpha$
- p-hodnotu počítají moderní počítačové programy
- existují úlohy, kdy se rozhoduje pouze podle p-hodnoty (např. Fisherův exaktní test ve čtyřpolní tabulce)

příklad: rozhodování pomocí p-hodnoty

- snažíme se prokázat, že šestka padá příliš často
- padlo nám $Y_0 = 17$, proto (vzorec pro psti binomického rozdělení)

$$p = P(Y \geq 17) = \sum_{k=17}^{100} \binom{100}{k} \left(\frac{1}{6}\right)^k \left(1 - \frac{1}{6}\right)^{100-k} = 0{,}506$$

- protože 50,6 % $>$ 5 %, hypotézu nemůžeme na 5% hladině zamítnout, netvrdíme, že pst šestky je větší než 1/6
- neprokázali jsme však, že by hypotéza platila

příklad: kostka a oboustranná alternativa

- chceme ověřit, zda je kostka v pořádku
- pokusíme se prokázat, že šestka padá příliš často nebo příliš zřídka
- H_0 : P(padne šestka) $= 1/6$
- H_1 : P(padne šestka) $\neq 1/6$
- je to **oboustranná alternativa** (na rozdíl od jednostranné)
- *proti* hypotéze svědčí malé *nebo* velké hodnoty Y
- pst chyby 1. druhu α rozdělíme na dvě poloviny: pro příliš nalé a příliš velké Y

příklad: kostka, oboustranná alternativa

y_0	9	10	11	...	23	24	25
$\mathsf{P}(Y < y_0)$	$0{,}010$	$0{,}021$	$0{,}042$	...	$0{,}937$	$0{,}962$	$0{,}978$
$\mathsf{P}(Y > y_0)$	$0{,}979$	$0{,}957$	$0{,}922$	...	$0{,}038$	$0{,}022$	$0{,}012$

- H_0 zamítneme, když bude $Y < 10$ *nebo* když bude $Y > 24$
- skutečná pst chyby 1. druhu bude 0,021 + 0,022 = 0,043
- hodnoty v rozmezí 10 až 24 (včetně obou mezí) nesvědčí proti H_0

oboustranná alternativa přibližně

- H_0 : P(padne šestka) $= 1/6 \qquad (= \pi_0)$
- H_1 : P(padne šestka) $\neq 1/6 \qquad (\neq \pi_0)$
- proti alternativě svědčí Y hodně daleko od $\mathsf{E}\,Y$, tj. rel. četnost $f = Y/n$ daleko od π_0:

$$\mathsf{P}\left(\left|\frac{Y - n\pi_0}{\sqrt{n\pi_0(1-\pi_0)}}\right| > z(\alpha/2)\right) = \alpha$$

- zamítáme tedy, je-li

$$Y < n\pi_0 - z(\alpha/2)\sqrt{n\pi_0(1-\pi_0)} \doteq 9{,}36$$

nebo

$$Y < n\pi_0 + z(\alpha/2)\sqrt{n\pi_0(1-\pi_0)} \doteq 23{,}97$$

znovu hodnocení četností

23	11	34
30	17	47
17	1	18
70	29	99

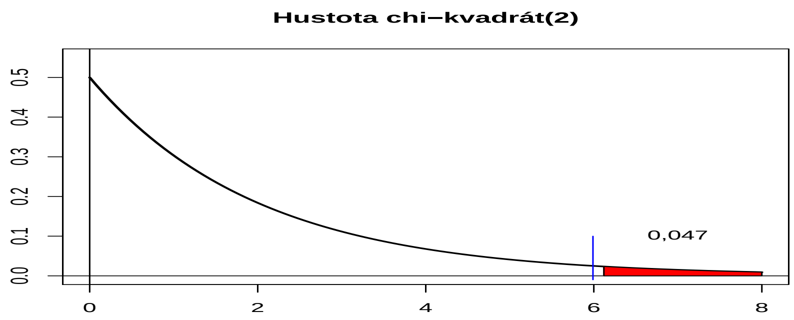
24,0	9,9	34
33,3	13,8	47
12,7	5,3	18
70	29	99

- statistika χ^2 porovnává skutečné četnosti (vlevo) s očekávanými (vpravo) za hypotézy

$$\chi^2 = \frac{(23-24{,}0)^2}{24{,}0} + \frac{(11-9{,}9)^2}{9{,}9} + \ldots + \frac{(1-5{,}3)^2}{5{,}3} = 6{,}12 > 5{,}99 = \chi^2_2(0{,}05)$$

- statistika χ^2 je větší, než kritická hodnota $\chi^2_2(0{,}05)$, závislost je na 5% hladině prokázána ($p = 0{,}047$)

grafická představa



- červená plocha = p-hodnota, modrá čára = hodnota statistiky