Procvičovací úkolu č.9 - Řešení příkladu č.3

Příklad č.3: Asymptotický interval spolehlivosti pro parametr θ alternativního rozdělení
Pro vybranou politickou stranu se v předvolebním průzkumu vyslovilo 60 z 1000 dotázaných osob. Stanovte 95 % asymptotický IS pro pravděpodobnost že tato politická strana ve volbách překročí 5 % hranici pro vstup do parlamentu.

Řešení:

V zadání se po nás jinými slovy chce, abychom stanovili 95 % intervalový odhad parametru θ, tedy 95 % asymptotický interval spolehlivosti. Protože nás zajímá, zda strana překročí 5 % hranici, sestrojíme 95 % levostranný IS pro parametr θ, který má obecný tvar

$$\langle dh; \infty)$$

Protože stanovujeme jednostranný IS, bude mít dolní hranice tvar:

$$dh = M - \sqrt{\frac{M(1-M)}{n}} u_{1-\alpha}$$

Před výpočtem musíme ověřit Haldovu podmínku dobré aproximace.

$$nc(1-c) > 9 \rightarrow n\theta(1-\theta) > 9$$

Protože přesnou hodnotu parametru θ neznáme, musíme ji v Haldově podmínce nahradit výběrovým průměrem M.

$$M = 60/1000 = 0.06$$

Hadlova podmínka má tedy tvar

$$nc(1-c) = nM(1-M) = 1000 * 0.06 * 0.94 = 56.4 > 9,$$

Tedy Haldova podmínka je splněna.

Dolní hranice IS má tedy tvar:

$$dh = M - \sqrt{\frac{M(1-M)}{n}} qnorm(0.95) = 0.06 - \sqrt{\frac{0.06 * 0.94}{1000}} * 1.6448 = 0.04765$$

95 % asymptotický IS pro parametr θ má tedy tvar:

$$\langle 0.04765; \infty) \tag{1}$$

Odpověď Protože 95 % asymptotický interval spolehlivosti pro parametr θ obsahuje i čísla menší než 0.05, může se stát, že politická strana ve volbách nepřekročí 5 % hranici pro vstup do parlamentu (Získá například jenom 4.81 % hlasů).

Poznámka: Pokud by IS obsahoval pouze čísla větší než 0.05, např. by měl tvar $\langle 0.0523; \infty)$, potom by závěr byl, že politická strana pravděpodobně překročí 5 % hranici pro vstup do parlamentu.

!!Upozornění!!: V zadání příkladu jsme měli pouze stanovit 95 % IS a na základě jeho tvaru určit, zda se politická strana do parlamentu dostane, nebo ne. V případě, že bychom chtěli otestovat hypotézu, že politická strana překročí pětiprocentní hranici a do parlamentu se dostane, bychom postupovali následovně.

Příklad č.3: Asymptotický interval spolehlivosti pro parametr θ alternativního rozdělení
Pro vybranou politickou stranu se v předvolebním průzkumu vyslovilo 60 z 1000 dotázaných osob. Na hladině významnosti $\alpha = 0.05$ otestujte, zda politická strana překročí 5 % hranici pro vstup do parlamentu.

$H_0 : \theta \leq 0.05$
$H_1 : \theta > 0.05$ (pravostranná alternativa)

Nejprve musíme zjistit, zda je splněna Haldova podmínka dobré aproximace

$$nc(1-c) > 9 \to n\theta(1-\theta) > 9.$$

O parametru θ nyní předpokládáme, že jeho hodnota je známá, a sice, že $\theta = 0.05$, jak je řečeno v nulové hypotéze. Haldova podmínka dobré aproximace má tedy v tomto případě tvar:

$$n\theta(1-\theta) = 1000 * 0.05 * (0.95) = 47.5 > 9.$$

Haldova podmínka dobré aproximace je splněna a tedy můžeme přistoupit k testování.

(a) Testování pomocí kritického oboru: Testovací statistika T_0 má tvar

$$T_0 = \frac{M-\theta}{\sqrt{\frac{\theta(1-\theta)}{n}}}$$

$$t_0 = \frac{0.06-0.05}{\sqrt{\frac{0.05*0.95}{1000}}} = 1.450953.$$

Proti pravostranné alternativě postavíme pravostranný kritický obor

$$W = \langle u_{1-\alpha}; \infty).$$

Kritický obor má tedy tvar

$$W = \langle 1.644854; \infty).$$

Závěr testování: Protože $t_0 = 1.45 \notin W = \langle 1.644854; \infty)$, $H0$ nezamítáme na hladině významnosti $\alpha = 0.05$.

(b) Testování pomocí intervalu spolehlivosti
Proti pravostranné alternativě H_1 postavíme 95 % levostranný interval spolehlivosti.

$$\langle dh; \infty),$$

kde

$$\begin{aligned} dh &= M - \sqrt{\frac{M(1-M}{n}} u_{1-\alpha} \\ &= 0.06 - \sqrt{\frac{0.06*0.94}{1000}} qnorm(0.95) \\ &= 0.06 - \sqrt{\frac{0.06*0.94}{1000}} * 1.6448 \\ &= 0.047647 \end{aligned}$$

Závěr testování: Protože $0 \in IS = \langle 0.04764; \infty)$, $H0$ nezamítáme na hladině významnosti $\alpha = 0.05$.

(c) Testování pomocí p-hodnoty
V části a) jsme spočítali hodnotu testovací statistiky $t_0 = 1.450953$

Protože máme pravostrannou alternativní hypotézu H_1, vypočítáme p-hodnotu podle vzorce

$$\text{p-val} = P(T_0 \geq t_0) = 1 - P(T_0 \leq t_0) = 1 - pnorm(t_0) = 0.07339654$$

Závěr testování: Protože p-hodnota $= 0.07339654 > 0.05 = \alpha$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

Interpretace výsledků testování: Všechny tři testy vedly k nezamítnutí H_0; na hladině významnosti $\alpha = 0.05$ tedy nezamítáme nulovou hypotézu o tom, že $\theta \leq 0.05$. Nemáme tedy dostatek indicií k tomu, abychom s jistotou řekli, že se politická strana do parlamentu dostane. (resp. Na základě testování jsme zjistili, že politická strana *pravděpodobně* nepřekročí 5 % hranici a do parlamentu se tedy nedostane).

Poznámka: Slovo '*poravděpodobně*' používáme proto, že pokud H_0 nezamítáme, znamená to, že nemáme dostatek indicií, abychom H_0 zamítli. To ale ještě neznamená, že H_0 stoprocentně platí. Z toho důvodu při stanovování závěru o nulové hypotéze nikdy nepoužíváme slovní spojení 'H_0 *platí*', ale používáme spojení 'H_0 *nezamítáme*'.