`NMSA202` | **Cvičení 11**
22. 05. 2017

Intervalové odhady a testování hypotéz

1. Nechť $X_1, \ldots, X_n$ značí hodnoty IQ náhodně vybraných žáků osmé třídy. Předpokládejme, že $X_1, \ldots, X_n$ tvoří náhodný výběr z normálního rozdělení s neznámou střední hodnotou a rozptylem 9.
   (a) Odhadněte bodově střední hodnotu IQ žáků osmé třídy. Jaké jsou vlastnosti a rozdělení tohoto odhadu?
   (b) Sestrojte intervalový odhad o spolehlivosti $1 - \alpha$ pro střední hodnotu IQ žáků osmé třídy. Interpretujte.
   (c) Jak se bude tento interval měnit, budeme-li zvyšovat spolehlivost $1 - \alpha$?
   Jak se bude tento interval měnit, budeme-li zvyšovat počet dětí zahrnutých do experimentu (a ostatní hodnoty ve vzorcích zůstanou stejné)?
   (d) Po provedení měření jsme obdrželi následující hodnoty:

   111, 116, 105, 111, 110, 114, 108, 106, 112, 108, 112, 111, 105, 111, 108, 110.

   Jak byste na základě těchto konkrétních hodnot odhadli střední hodnotu IQ žáků osmé třídy? Uveďte bodový i 95%-ní intervalový odhad.
   (e) Napište 95%-ní dolní intervalový odhad pro střední hodnotu IQ žáků osmé třídy a interpretujte jej.
   (f) Odborníci A, B a C by rádi na základě těchto naměřených údajů potvrdili/vyvrátili svoje domněnky.
      (A) Odborník A tvrdí, že je střední hodnota IQ žáků 8. třídy rovno 110.
      (B) Odborník B zastává názor, že je střední hodnota IQ žáků 8.třídy jistě menší než 115.
      (C) Odborník C soudí, že je střední hodnota IQ žáků 8.třídy určitě vyšší než 110.
   Zformulujte nulovou a alternativní hypotézu, které chce testovat odborník A (resp. odborníci B a C). Proveďte vhodný test na hladině $\alpha = 0{,}05$.

2. Provádíme průzkum, zda v nejmenované hospodě okrádají své hosty. Zakoupíme proto 10 piv a změříme jejich objem. Obdrželi jsme následující hodnoty (v litrech):

   0,510, 0,462, 0,491, 0,466, 0,461, 0,503, 0,495, 0,488, 0,512, 0,505.

   Předpokládejme, že datům odpovídají nezávislé náhodné veličiny s normálním rozdělením.
   (a) Odhadněte střední hodnotu a rozptyl objemu jednoho natočeného piva. Připomeňte, jaké jsou teoretické vlastnosti a rozdělení těchto odhadů.
   (b) Zkonstruujte intervalový odhad pro střední hodnotu objemu jednoho piva o spolehlivosti 95 %. Leží předepsaná hodnota 0,5 l v tomto intervalu?
   (c) Napište dále oba 95%-ní jednostranné (dolní a horní) intervaly spolehlivosti pro střední hodnotu objemu jednoho piva a interpretujte je. Který z nich Vám připadá „zajímavější“ pro tuto situaci?
   (d) Zkonstruujte 90%-ní intervalový odhad pro rozptyl objemu jednoho piva.
   (e) Otestujte, zda je hostinský skutečně nepoctivý: Zformulujte nulovou a alternativní hypotézu, vyberte vhodný test, napište jeho kritický obor a proveďte jej na hladině $\alpha = 0{,}05$.
   (f) Jaká je souvislost mezi výsledkem tohoto testu a intervalovým odhadem, který jsme konstruovali?

3. *Centrum pro výzkum veřejného mínění* v dubnu 2013 zveřejnilo výsledky aktuálního průzkumu mezi občany ČR staršími 15 let. Můžeme se dočíst, že z 1 059 respondentů jich 297 uvedlo, že „do korupčního jednání se zapojují téměř všichni veřejní činitelé“.
   (a) Odhadněte bodově podíl občanů ČR, kteří si myslí, že do korupčního jednání se zapojují téměř všichni veřejní činitelé. Jaké je rozdělení a asymptotické rozdělení tohoto bodového odhadu?
   (b) Určete asymptotický 95%-ní interval spolehlivosti pro podíl občanů, kteří si myslí, že do korupčního jednání se zapojují téměř všichni veřejní činitelé.
   (c) Jaký by musel být rozsah výběru n, aby intervalový odhad s asymptotickou spolehlivostí 95% pro tento podíl měl šířku nejvýše 0,03 (uvažujeme-li, že podíl občanů zahrnutých ve studii, kteří vyjadřují uvedený názor, se nezmění).

4. Chceme porovnat průměrnou výšku jedenáctiletých chlapců a dívek. Studie se zúčastnilo 25 chlapců (veličiny $X_1, \dots, X_{25}$) a 20 dívek (veličiny $Y_1, \dots, Y_{20}$). Obdrželi jsme následující výsledky:

$$\overline{X} = 147{,}1, \quad \overline{Y} = 147{,}8, \quad S_X^2 = 10{,}9, \quad S_Y^2 = 14{,}7.$$

   (a) Lze předpokládat, že výška chlapců i dívek má normální rozdělení se stejným rozptylem σ^2. Sestrojte intervalový odhad pro rozdíl průměrné výšky o spolehlivosti 95 %.
   (b) Zjistěte, zda lze tvrdit, že je střední hodnota výšky chlapců o více než 10 cm větší než střední hodnota výšky dívek. Zformulujte nulovou a alternativní hypotézu, kterou chceme testovat. Vyberte a proveďte vhodný test a uveďte jeho předpoklady.

5. Zkonstruujte intervalový odhad parametru λ o asymptotické spolehlivosti 95 % pro data z příkladu 2 z minulého cvičení.

6. U několika leváků byla měřena síla stisku levé a pravé ruky.

Levá	140	90	125	130	95	121	85	97	131	110
Pravá	138	87	110	132	96	120	86	90	129	100

   Potvrzují data domněnku, že levá ruka je silnější? Vyberte vhodný test, popište jeho předpoklady. Formulujte hypotézy a napište kritický obor.

7. Chystáte se zavést na trh nový prostředek na hubnutí. Provedli jste studii na 250 náhodně vybraných dobrovolnících a zjistili jste, že snížení hmotnosti nastalo u 140 z nich. Budete moci na základě těchto dat připravit slogan „Opravdu efektivní hubnutí: úspěch ve více než 50 % případů“? Zformulujte nulovou a alternativní hypotézu a sestrojte test pomocí centrální limitní věty. Proveďte jej, přičemž vezměte v úvahu, že reklamní slogan může být „nepravdivý“ s pravděpodobností nejvýše 0,01.

Opakování z přednášky

Model: $X_1, \dots, X_n$ je náhodný výběr (nezávislé stejně rozdělené veličiny) z rozdělení F_θ, které závisí na neznámém parametru $\theta \in \Theta$.

Minulé cvičení: bodový odhad parametrické funkce $g(\theta)$ $\leftrightsquigarrow$ náhodná veličina
+ jeho „pěkné“ vlastnosti $\leftrightsquigarrow$ nestrannost, konzistence

Intervalový odhad parametrické funkce $g(\theta)$ $\leftrightsquigarrow$ interval s náhodnými mezemi, který pokrývá skutečnou (neznámou) hodnotu $g(\theta)$ s předepsanou pravděpodobností $1-\alpha$

- Intervalový odhad parametrické funkce $g(\theta)$ o spolehlivosti $1-\alpha$ je dvojice náhodných veličin (L_n, U_n) takových, že

$$\mathsf{P}_\theta\big(L_n < g(\theta) < U_n\big) = 1-\alpha \quad \text{pro všechna } \theta \in \Theta. \tag{11.1}$$

- Veličina $D_n = D_n(X_1, \dots, X_n)$ se nazývá dolní intervalový odhad $g(\theta)$ o spolehlivosti $1-\alpha$, jestliže

$$\mathsf{P}_\theta\big(D_n < g(\theta)\big) = 1-\alpha \quad \text{pro všechna } \theta \in \Theta. \tag{11.2}$$

- Veličina $H_n = H_n(X_1, \dots, X_n)$ se nazývá horní intervalový odhad $g(\theta)$ o spolehlivosti $1-\alpha$, jestliže

$$\mathsf{P}_\theta\big(g(\theta) < H_n\big) = 1-\alpha \quad \text{pro všechna } \theta \in \Theta. \tag{11.3}$$

Podobně jako jsme měli u bodového odhadu, $L_n = L_n(X_1, \dots, X_n)$, $U_n = U_n(X_1, \dots, X_n)$, $D_n = D_n(X_1, \dots, X_n)$ a $H_n = H_n(X_1, \dots, X_n)$ jsou borelovské funkce náhodného výběru $X_1, \dots, X_n$, jejichž funkční předpisy nezávisí na θ.

Číslo $\alpha \in (0,1)$ volíme, v praxi se většinou pracuje s $\alpha = 0{,}05$.

Obecná konstrukce intervalu spolehlivosti pro parametrickou funkci $g(\theta)$**.** Najdeme funkci náhodného výběru a parametrické funkce $g(\theta)$, tj. náhodnou veličinu $h\big(X_1, \dots, X_n; g(\theta)\big)$, jejíž rozdělení nezávisí na θ. Nechť $h_{\alpha/2}$ a $h_{1-\alpha/2}$ jsou kvantily tohoto rozdělení. Pak jistě platí

$$\mathsf{P}_\theta\Big(h_{\alpha/2} < h\big(X_1, \dots, X_n; g(\theta)\big) < h_{1-\alpha/2}\Big) = 1-\alpha \quad \text{pro všechna } \theta \in \Theta. \tag{11.4}$$

Je-li možné nerovnosti v závorce převést ekvivalentními úpravami na takový tvar, že uprostřed stojí $g(\theta)$ a vlevo i vpravo je něco, co na θ nezávisí, sestrojili jsme intervalový odhad.

Speciální případy.

- Intervalové odhady pro parametry **normálního rozdělení** jsou shrnuty v tabulkách 5.1 a 5.2 ve skriptech.
- **Asymptotické intervalové odhady** (intervalové odhady s asymptotickou spolehlivostí $1-\alpha$) je možné zkonstruovat pomocí CLV - tabulka 5.3 ve skriptech.
  Intervalové odhady pro speciální případ **alternativního rozdělení** jsou rozepsány v tabulce 5.4.

Testování hypotéz je naivně řečeno ověřování platnosti nějakého výroku.

- **Hypotéza** je výrok (o nějaké populaci), o jehož platnosti chceme rozhodnout na základě nasbíraných dat.
- **Předpokládaný model**: $X_1, \dots, X_n$ náhodný výběr z určitého rozdělení F_θ, kde $\theta \in \Theta$ neznáme. Naše data jsou pak realizací takového výběru.
- Testuje se vždy tzv. **nulová hypotéza** H_0: $\theta \in \Theta_0$ proti **alternativní hypotéze** H_1: $\theta \in \Theta_1$, kde Θ_0 a Θ_1 jsou disjunktní.
- **Test** je rozhodovací pravidlo (postup), na jehož základě zamítáme nebo nezamítáme H_0.
  Možná rozhodnutí:
  - zamítáme H_0 ve prospěch H_1 (*naše data svědčí proti H_0, prokazujeme platnost H_1*)
  - nezamítáme H_0 (*na základě našich dat nelze H_0 zamítnout, naše data nejsou v rozporu s H_0*)

  $\leftrightsquigarrow$ **nesymetrie** mezi H_0 a H_1!
- Většinou nemůžeme rozhodnout s absolutní jistotou, která z hypotéz je platná $\rightsquigarrow$ můžeme se dopustit **chyby**. Mohou nastat tyto možnosti:

	Skutečnost	
Rozhodnutí	H_0 platí	H_1 platí
zamítáme H_0	chyba 1.druhu	OK
nezamítáme H_0	OK	chyba 2.druhu

- **Chyba 1.druhu je závažnější** (falešně něco prokazujeme) $\rightsquigarrow$ její pravděpodobnost chceme kontrolovat.
  Zvolíme α = maximální přípustná pravděpodobnost chyby 1.druhu (většinou $\alpha = 0{,}05$ nebo $0{,}01$) a chceme
  $$\mathsf{P}(\text{chyba 1. druhu}) = \mathsf{P}(\text{zamítáme } H_0 \mid H_0 \text{ platí}) \leq \alpha. \tag{11.5}$$
- Test je popsán **kritickým oborem** W = množina výsledků pokusů, pro které H_0 zamítáme
  - Je-li $(X_1, \dots, X_n) \in W$, pak H_0 zamítáme.
  - Je-li $(X_1, \dots, X_n) \notin W$, pak H_0 nezamítáme.

  Musí platit:
  $$\mathsf{P}_\theta\big((X_1, \dots, X_n) \in W\big) \leq \alpha \quad \text{pro všechna } \theta \in \Theta_0 \tag{11.6}$$
  a mluvíme pak o testu na **hladině** α.
- Testy o parametrech normálního rozdělení (tabulky 6.1, 6.2 a 6.3 ve skriptech), testy založenými na CLV (tabulka 6.4 ve skriptech).