Cvičení z NSTP022
12. týden cvičení (6. – 10. 5. 2013)

Testování hypotéz

1. Zabýváme se opět otázkou IQ žáků 8. třídy. Provedli jsme 16 měření IQ náhodně vybraných žáků 8.třídy a obdrželi jsme hodnoty, kterým odpovídá $\bar{X}_{16} = 109{,}875$. Stejně jako minule budeme předpokládat, že data pochází z normálního rozdělení s neznámou střední hodnotou a rozptylem 9. 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 $\alpha = 0{,}05$. Interpretujte výsledek.

2. Provádíme průzkum, zda v nejmenované hospodě okrádají své hosty. Na minulém cvičení jsme proto zakoupili 10 piv, změřili jejich objem a obdrželi jsme hodnoty (v litrech)

$$\bar{X}_{10} = 0{,}4893, \quad S_{10}^2 = 0{,}0003893444.$$

   Předpokládáme, že datům odpovídají nezávislé náhodné veličiny s normálním rozdělením.
   (a) Otestujte, zda je hostinský skutečně nepoctivý: Zformulujte nulovou a alternativní hypotézu, vyberte vhodný test, napište jeho kritický obor a proveďte jej na hladině $\alpha = 0{,}05$.
   (b) Jaká je souvislost mezi výsledkem tohoto testu a intervalovým odhadem, který jsme minule konstruovali?

3. Chceme porovnat průměrnou výšku 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:

$$\bar{X}_{25} = 180{,}6 \quad \bar{Y}_{20} = 164{,}9 \quad S_X^2 = 6{,}5 \quad S_Y^2 = 9{,}3.$$

   Zjistěte, zda lze tvrdit, že je průměrná výška chlapců o více než 10 cm větší než průměrná výška dívek.
   Jaký model zde musíme předpokládat? Zformulujte nulovou a alternativní hypotézu, kterou chceme testovat. Vyberte a proveďte vhodný test a uveďte jeho předpoklady.

4. 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.

5. 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 - Testování hypotéz

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
  $$\mathsf{P}(\text{chyba 1. druhu}) = \mathsf{P}(\text{zamítáme } H_0 |\ H_0 \text{ platí}) \leq \alpha. \tag{12.1}$$
- 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:
  $$\mathsf{P}_\theta\big((X_1, \dots, X_n) \in W\big) \leq \alpha \quad \text{pro všechna } \theta \in \Theta_0 \tag{12.2}$$
  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).

Příklad. Nechť $X_1, \dots, X_n$ je náhodný výběr z rozdělení $\mathsf{N}(\mu, \sigma^2)$, kde $\sigma^2 > 0$ je známé. Chceme testovat např.
$$H_0 :\ \mu = \mu_0 \quad \text{proti} \quad H_1 :\ \mu \neq \mu_0. \tag{12.3}$$
Příslušný test na hladině α je popsán kritickým oborem
$$\frac{\sqrt{n}\,|\bar{X}_n - \mu_0|}{\sigma} \geq u_{1-\alpha/2}. \tag{12.4}$$
Pro naše konkrétní data $x_1, \dots, x_n$ spočteme levou stranu (12.4). V případě, že nerovnost (12.4) platí, pak H_0 zamítáme a prokázali jsme, že skutečná hodnota μ se nerovná μ_0. Neplatí-li tato nerovnost, nelze na základě našich dat H_0 zamítnout.

Výsledky

1. Označení X_i = IQ žáka i, $i = 1, \ldots, 16$.
   Model: $X_1, \ldots, X_{16}$ nezávislé náhodné veličiny s $\mathsf{N}(\mu, 9)$ rozdělením.
   (a) na hladině 0,05 nelze zamítnout $H_0 : \mu = 110$ ve prospěch $H_1 : \mu \neq 110$; tj. naše data nejsou v rozporu s tvrzením A.
   (b) na hladině 0,05 zamítáme $H_0 : \mu = 115$ ve prospěch $H_1 : \mu < 115$, tj. prokázali jsme tvrzení B.
   (c) na hladině 0,05 nelze zamítnout $H_0 : \mu = 110$ ve prospěch $H_1 : \mu > 110$, tj. nelze prokázat tvrzení C.

2. Označení X_i = objem piva i, $i = 1, \ldots, 10$.
   Model: $X_1, \ldots, X_{10}$ nezávislé náhodné veličiny s $\mathsf{N}(\mu, \sigma^2)$ rozdělením.
   Testujeme $H_0 : \mu = 0{,}5$ proti $H_1 : \mu \neq 0{,}5$ (nebo proti $H_1 : \mu < 0{,}5$).
   Nezamítáme H_0, tj. na hladině 0,05 nelze prokázat, že je hostinský nepoctivý.

3. Model:
   - $X_1, \ldots, X_{25}$ výška chlapců v cm, náhodné veličiny s $\mathsf{N}(\mu_1, \sigma^2)$,
   - $Y_1, \ldots, Y_{20}$ výška dívek v cm, náhodné veličiny s $\mathsf{N}(\mu_2, \sigma^2)$,
   - všechny veličiny $X_1, \ldots, X_{25}$, $Y_1, \ldots, Y_{20}$ jsou nezávislé.

   Testujeme $H_0 : \mu_1 - \mu_2 = 10$ proti $H_1 : \mu_1 - \mu_2 > 10$.
   Na hladině 0,05 zamítáme H_0, tj. můžeme tvrdit, že jsou chlapci v průměru o více než 10 cm vyšší než dívky.

4. X_i=stisk levé ruky osoby i, Y_i=stisk pravé ruky osoby i, $i = 1, \ldots, 10$.
   Veličiny (X_i, Y_i) jsou závislé, proto definujeme $Z_i = X_i - Y_i$.
   Model: $Z_1, \ldots, Z_{10}$ nezávislé náhodné veličiny s $\mathsf{N}(\mu, \sigma^2)$ rozdělením.
   Testujeme $H_0 : \mu = 0$ proti $H_1 : \mu > 0$.
   Na hladině 0,05 zamítáme H_0, tj. data potvrzují domněnku, že je u leváků levá ruka silnější.

5. X_i=indikátor toho, zda u i-té osoby došlo ke snížení hmotnosti.
   Model: $X_1, \ldots, X_{250}$ nezávislé náhodné veličiny s alternativním rozdělením $\mathsf{Alt}(p)$.
   Testujeme $H_0 : p = 0{,}5$ proti $H_1 : p > 0{,}5$.
   Na asymptotické hladině 0,01 nezamítáme H_0, tj. na základě našich dat nelze prokázat tvrzení na sloganu.