Úvod do testování hypotéz
a
vybrané parametrické testy hypotéz

—

Supplementum ke cvičení 4ST201 Statistika

Lubomír Štěpánek[1, 2]

[1]Oddělení biomedicínské statistiky
Ústav biofyziky a informatiky
1. lékařská fakulta
Univerzita Karlova, Praha

[2]Katedra biomedicínské informatiky
Fakulta biomedicínského inženýrství
České vysoké učení technické v Praze

Obsah

Příklad

- V zásilce je 80 kusů zboží. Náhodně vybereme osm kusů a postupně je zvážíme. Získané hmotnosti (v kilogramech) jsou

$$8 \quad 8 \quad 9 \quad 7 \quad 6 \quad 8 \quad 9 \quad 9.$$

  Určeme bodový odhad střední hodnoty hmotnosti jednoho kusu zboží a odhad rozptylu hodnoty hmotnosti jednoho kusu zboží. Měly-li dva z vybraných osmi kusů zboží vadu, jaký je bodový odhad populační četnosti zboží s vadou?

Příklad

- Změřili jsme tělesnou výšku šesti žen a získali jsme následující hodnoty (v centimetrech)

$$163 \quad 175 \quad 177 \quad 165 \quad 171 \quad 174.$$

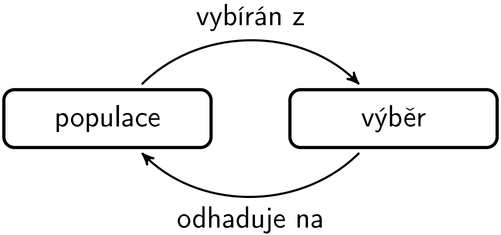
Předpokládáme, že tělesná výška ženy sleduje normální rozdělení s (populačním) rozptylem 49 cm^2. Určeme

(i) bodový odhad populační střední hodnoty tělesné výšky ženy;
(ii) oboustranný 95% interval spolehlivosti pro populační střední hodnotu tělesné výšky ženy;
(iii) nutný počet žen, jejichž tělesnou výšku bychom museli změřit, abychom snížili přípustnou chybu odhadu[1] populační střední hodnoty tělesné výšky ženy na 1 cm.

[1] jde o poloviční délku intervalu spolehlivosti pro střední hodnotu

Vztah populace a výběru

- z populace je vybírán výběr
- z charakteristik výběru je usuzováno na parametry populace



Statistické usuzování

- statistické usuzování je založeno na statistických hypotézách
- statistická hypotéza je výrok o parametru pravděpodobnostního rozdělení zkoumaného znaku
- k rozhodování o pravdivosti statistických hypotéz používáme obvykle hodnoty ze statistického výběru
  - populace (základní soubor) je prakticky vždy alespoň částečně nedostupná

princip

- základním principem testování hypotéz je formulace testované (vyvrátitelné, znegovatelné) hypotézy, kterou při nalezení dostatečně významného protipříkladu[2] vyvrátíme ve prospěch jejího opaku; jinak ji považujeme za platnou (odpovídající realitě)

[2] ve statistice znamená nalezení protipříkladu sesbírání dostatku dat odporujících testované hypotéze

Nulová hypotéza

- testovanou statistickou hypotézu, kterou lze eventuálně vyvrátit (negovat), nazýváme *nulovou* hypotézou a značíme ji H_0
  - např. populační střední hodnota μ je rovna konkrétnímu μ_0, tedy

    $$H_0 : \mu = \mu_0,$$

    kde μ_0 je nějaká konstanta
  - nulová hypotéza má obvykle tvar rovnosti[3] populační charakteristiky a konstanty nebo dvou (různých) populačních charakteristik

[3] konkrétní rovnost ($=$) lze na rozdíl od obecné nerovnosti ($\neq$) přímočaře negovat

Alternativní hypotéza

- při testování hypotéz se vždy nakonec pro platnost některé hypotézy rozhodneme, proto doplňujeme nulovou hypotézu ještě o její negaci, tzv. *alternativní* statistickou hypotézu značenou H_1
  - např. populační střední hodnota μ není rovna (je větší než, je menší než) konkrétnímu μ_0, tedy

$$H_1 : \mu \neq \mu_0 \qquad \text{(dvoustranná varianta)}$$

anebo

$$H_1 : \mu > \mu_0 \qquad \text{(jednostranná varianta)}$$

anebo

$$H_1 : \mu < \mu_0, \qquad \text{(jednostranná varianta)}$$

kde μ_0 je vždy nějaká konstanta

Vyvrácení nulové hypotézy a možné závěry testování

- protože usuzujeme na populační charakteristiky pomocí hypotéz postavených pouze na výběrových hodnotách, může dojít i k chybným závěrům

		závěr testu (pravděpodobnost)	
		H_0 se nezamítá	H_0 se zamítá
skutečnost	H_0 je pravdivá	správné rozhodnutí $(1-\alpha)$	chyba I. typu (α)
	H_0 není pravdivá	chyba II. typu (β)	správné rozhodnutí $(1-\beta)$

Obecný postup testování hypotéz

(i) formulace hypotéz
- začíná volbou porovnávaných parametrů, tj. měr střední hodnoty, variability či tvaru

(ii) volba testového kritéria
- testové kritérium je skalární proměnná, která převede test hypotézy coby rozhodovací pravidlo pouze na úroveň jednoho čísla
- obvykle se pro zvolené porovnávané parametry využívá ustálených algoritmů (testů hypotéz) a výpočtů testového kritéria

(iii) sestrojení kritického oboru
- kritický obor je interval hodnot testového kritéria, pro které zamítneme nulovou hypotézu ve prospěch alternativní

(iv) výpočet hodnoty testového kritéria

(v) formulace závěrů testu

Parametrické testy hypotéz

- testem hypotézy rozumíme obvykle ustálené rozhodovací pravidlo založené na testování statistické hypotézy určitého typu
  - jednotlivé testy hypotéz se od sebe liší především tím, jaké parametry základního souboru a výběru, resp. dvou výběrů mezi sebou porovnávají
- parametrické testy hypotéz předpokládají platnost některých apriorních *parametrů* výběru či základního souboru
  - hodnoty výběru či základního souboru tvoří kvantitativní znak
  - mnohdy předpokládáme normální rozdělení zkoumané statistické veličiny (viz dále)

Test hypotézy o střední hodnotě

- testuje nulovou hypotézu $H_0 : \mu = \mu_0$ o tom, zda se průměr μ základního souboru (populace) rovná určité hodnotě μ_0
- testovým kritériem je v případě známého populačního rozptylu σ^2

$$U = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n} \sim \mathcal{N}(0, 1^2)$$

kde $\bar{X}$ je výběrový průměr a n rozsah výběru

- a v případě známého populačního rozptylu

$$U = \frac{\bar{X} - \mu_0}{s_x}\sqrt{n} \sim \mathcal{N}(0, 1^2),$$

kde $\bar{X}$ je výběrový průměr, n rozsah výběru a nakonec výraz $s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})$ je výběrový rozptyl

Test hypotézy o střední hodnotě

- pro malé výběry do $n < 30$ nahrazujeme testové kritérium statistikou

$$t = \frac{\bar{X} - \mu_0}{s_x}\sqrt{n} \sim t(n-1),$$

kde $\bar{X}$ je výběrový průměr, n rozsah výběru, s_x^2 výběrový rozptyl a $t(n-1)$ je kvantil Studentova t-rozdělení o $n-1$ stupních volnosti

Test hypotézy o střední hodnotě

- alternativní hypotézou a kritickým oborem tak může být

$$H_1 : \mu \neq \mu_0, \quad W_\alpha = \{u : |u| \geq u_{1-\alpha/2}\} \quad \text{(dvoustranný test)}$$

anebo

$$H_1 : \mu > \mu_0, \quad W_\alpha = \{u : u \geq u_{1-\alpha}\} \quad \text{(jednostranný test)}$$

anebo

$$H_1 : \mu < \mu_0, \quad W_\alpha = \{u : u \leq -u_{1-\alpha}\} \quad \text{(jednostranný test)}$$

Příklad

- Ředitel tvrdí, že průměrná mzda v jeho firmě je 25000 korun. Chceme ověřit jeho tvrzení, neboť se domníváme, že údaj o průměrné mzdě nadhodnotil. Máme údaje o 49 mzdách, ze který jsme spočítali průměr 23929 korun a výběrovou směrodatnou odchylku 5603 korun. Předpokládáme, že mzdy mají normální rozdělení. Hladinu významnosti volme 0,05.

Test hypotézy o relativní četnosti

- testuje nulovou hypotézu $H_0 : \pi = \pi_0$ o tom, že relativní četnost π určité varianty znaku v základním souboru (populaci) se rovná určité hodnotě π_0
- testovým kritériem je

$$U = \frac{P - \pi_0}{\sqrt{\pi_0(1 - \pi_0)}} \sqrt{n} \sim \mathcal{N}(0, 1^2)$$

kde P je výběrová relativní četnost dané varianty znaku

Test hypotézy o relativní četnosti

- alternativní hypotézou a kritickým oborem tak může být

$$H_1 : \pi \neq \pi_0, \quad W_\alpha = \{u : |u| \geq u_{1-\alpha/2}\} \quad \text{(dvoustranný test)}$$

anebo

$$H_1 : \pi > \pi_0, \quad W_\alpha = \{u : u \geq u_{1-\alpha}\} \quad \text{(jednostranný test)}$$

anebo

$$H_1 : \pi < \pi_0, \quad W_\alpha = \{u : u \leq -u_{1-\alpha}\} \quad \text{(jednostranný test)}$$

Příklad

- V malé nemocnici se za poslední měsíc narodilo 90 dívek a 60 chlapců. Na hladinách významnosti (i) 0,05 a (ii) 0,01 chceme otestovat hypotézu, zda se dívky a chlapci rodí se stejnou pravděpodobností.

Test hypotézy o shodě dvou středních hodnot

- testuje nulovou hypotézu $H_0 : \mu_1 = \mu_2$ o tom, zda se průměr μ_1 jednoho základního souboru rovná průměru μ_2 druhého základního souboru
- předpokládáme, že známe rozptyly σ_1^2 a σ_2^2 v obou souborech
- testovým kritériem je

$$U = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim \mathcal{N}(0, 1^2)$$

kde $\bar{X}_1$ a $\bar{X}_2$ jsou výběrové průměry v obou výběrech z daných souborů, n_1 a n_2 jsou rozsahy výběrů

Test hypotézy o shodě dvou středních hodnot

- alternativní hypotézou a kritickým oborem tak může být

$$H_1 : \mu_1 \neq \mu_2, \quad W_\alpha = \{u : |u| \geq u_{1-\alpha/2}\} \quad \text{(dvoustranný test)}$$

anebo

$$H_1 : \mu_1 > \mu_2, \quad W_\alpha = \{u : u \geq u_{1-\alpha}\} \quad \text{(jednostranný test)}$$

anebo

$$H_1 : \mu_1 < \mu_2, \quad W_\alpha = \{u : u \leq -u_{1-\alpha}\} \quad \text{(jednostranný test)}$$

Příklad

- Bylo vybráno 73 polí stejné kvality. Na 38 z nich se zkoušel nový způsob hnojení, zbývajících 35 bylo ošetřeno původním způsobem. Průměrné výnosy pšenice při novém způsobu hnojení byly 5,4 tun na hektar a výběrový rozptyl byl 0,27. Průměrné výnosy pšenice při původním způsobu hnojení byly 4,7 tun na hektar a výběrový rozptyl byl 0,24. Chceme zjistit, zda nový způsob hnojení má vliv na výnosy pšenice. Hladinu významnosti volme 0,05.

Literatura

- Hindls, Richard, Stanislava Hronová, Jan Seger a Jakub Fischer. *Statistika pro ekonomy*. Praha: Professional Publishing, 2007. ISBN: 978-80-86946-43-6.
- Marek, Luboš. *Statistika v příkladech*. Praha: Professional Publishing, 2015. ISBN: 978-80-7431-153-6.

Děkuji za pozornost!

lubomir.stepanek@vse.cz
lubomir.stepanek@lf1.cuni.cz
lubomir.stepanek@fbmi.cvut.cz

https://github.com/LStepanek/4ST201_Statistika