Princip testování hypotéz

Jednovýběrové testy parametrických hypotéz

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

doc. Ing. Petr Seďa, Ph.D.

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Co se dnes dozvíte?

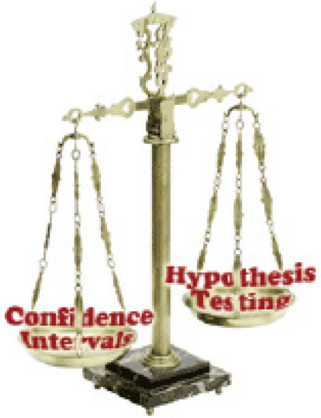
- Základní pojmy z testování hypotéz.
- Hypotéza, testová statistika, významnost testu.
- Test hypotézy o střední hodnotě.
- Jednostranné a oboustranné testy.
- Testy hypotéz o podílu.
- Testy o rozptylu.

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Statistická indukce?

- **Intervalové odhady** (angl. confidence intervals) – umožňují odhadnout nejistotu v odhadu parametru náhodné veličiny.
- **Testování hypotéz** (angl. hypothesis testing) - umožňuje posoudit, zda experimentálně získaná data nepopírají předpoklad, který jsme před provedením testování učinili.

Používáme, chceme-li určit velikost parametru NV, resp. velikost efektu (rozdílu, resp. poměru parametrů dvou NV).



Používáme, chceme-li ověřit platnost předem definované hypotézy (s předem danou hladinou významností).

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Co je to statistická hypotéza?

- **Statistická hypotéza** – tvrzení o vlastnostech **základního souboru** nebo **náhodné veličiny** (obvykle o hodnotě některého **parametru** nebo **rozdělení**), které se snažíme na základě náhodného výběru **potvrdit** nebo **vyvrátit**

Co je zdrojem statistických hypotéz?

- předchozí zkušenosti,
- teorie, kterou je třeba doložit,
- požadavky na kvalitu produktu,
- dohady založené na náhodném pozorování…

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Příklady statistických hypotéz:

- Střední životnost pračky je nižší než výrobcem udávaných 10 let.
- Mortalita pacietů je u laparoskopických operací nižší než u operací tradičně provedených.
- Průměrné výsledky zápočtových testů z předmětu Matematická analýza II u studentů Matematického ústavu závisí na typu absolvované střední školy.
- Pořízený výběrový datový soubor je výběrem z populace mající normální rozdělení pravděpodobnosti atd.

Poznámka: Rozdíl (resp. poměr) parametru náhodné veličiny a jeho očekávané hodnoty, popřípadě rozdíl (resp. poměr) parametrů náhodných veličin nazýváme **efekt**.

Příklad 1 - cholesterol

Domníváme se, že střední hodnota obsahu cholesterolu v krvi je u populace města Opavy 4,9 mmol/l.

$$H_0: \mu = 4{,}9$$
$$H_A: \mu \neq 4{,}9$$

Jak tento předpoklad ověřit?

- Zjistíme údaje o obsahu cholesterolu v krvi u 550 náhodně vybraných obyvatel Opavy.
- Průměrný obsah cholesterolu v krvi **probandů** (tj. jedinců, kteří jsou předmětem zkoumání) byl 5,3 mmol/l.

Jsou tyto výsledky v souladu s naší hypotézou?

- I kdyby byla testovaná hypotéza pravdivá, nelze očekávat, že průměrná hodnota pozorovaná ve výběru bude přesně 4,9 mmol/l.
- Nulovou hypotézu zamítneme, pokud získané uspořádání výběru bude za předpokladu platnosti nulové hypotézy **velmi nepravděpodobné**.

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Jak ověřit, zda je statistická hypotéza pravdivá?

Pravdivost nulové hypotézy **nelze** na základě dat **dokázat**! Pravdivost nulové hypotézy lze na základě dat **pouze vyvrátit**.

Nulová hypotéza:
obžalovaný je nevinen.

Alternativní hypotéza:
obžalovaný je vinen.

Data: výběrový soubor

Testové kritérium: soudce

Princip **presumpce neviny**: Neodsoudí-li soudce obžalovaného, nemusí to znamenat, že je obžalovaný nevinný. Může to znamenat, že neexistuje dostatek důkazů pro jeho odsouzení!

Pokud se nepodaří nalézt hodnověrný důvod pro zamítnutí nulové hypotézy H_0 ve prospěch alternativy H_1, pak nezamítáme nulovou hypotézu H_0

Testování hypotéz

Rozdělení hypotéz:

- **Parametrická statistická hypotéza** – tvrzení ohledně **efektu,** vyjadřují určitou hodnotu parametru rozdělení základního souboru (obvykle normálního - např. střední hodnoty).

Příklad: průměrný dosažený věk mužů v ČR je 67 let

  - Hypotézy o parametru **jedné populace** (o střední hodnotě, rozptylu, mediánu, parametru binomického rozdělení, …)
  - Hypotézy o parametrech **dvou populací** (srovnávací testy)
  - Hypotézy o parametrech **více než dvou populací** (např. ANOVA, …)
- **Neparametrická statistická hypotéza** – tvrzení o **jiné vlastnosti rozdělení** náhodné veličiny než o jejím parametru (např. hypotézy o typu rozdělení NV, hypotézy o závislosti NV, …).

Příklad: děti vysokoškoláků jsou obvykle také vysokoškoláci

Poznámka: Rozdíl (resp. poměr) parametru náhodné veličiny a jeho očekávané hodnoty, popřípadě rozdíl (resp. poměr) parametrů náhodných veličin nazýváme **efekt**.

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Rozdělení parametrických hypotéz I:

podle počtu srovnávaných souborů (proměnných)

- o parametrech jednoho znaku

Příklad: průměrný dosažený věk mužů v ČR je 67 let

- o parametrech dvou znaků (tzv. srovnávací testy)

Příklad: průměrný věk žen je vyšší než mužů

- o parametrech více než dvou znaků

Příklad: průměrná cena kávy ve všech sledovaných lokalitách je stejná

Rozdělení parametrických hypotéz II:

podle typu parametrů

- o střední hodnotě μ

Příklad: průměrný dosažený věk mužů v ČR je 67 let

- o rozptylu σ^2 nebo směrodatné odchylce σ

Příklad: diferenciace platů se úpravou tarifních tabulek nezměnila

- o podílu dané vlastnosti v souboru π

Příklad: mobilní telefony používá více než 80% obyvatel ČR

Testování hypotéz

Jak se ve statistice testuje?

Neymann – Pearsonova metoda

výběr ze dvou hypotéz:

nulová hypotéza

rovnovážný stav – **rovnost** $\mu = 13500$

alternativní hypotéza

nerovnost

oboustranná $\mu \neq 13500$

jednostranná $\mu > 13500$

Chyby při rozhodování:

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="2">testování</td></tr>
<tr><td>je zvolena H_0</td><td>je zvolena H_1</td></tr>
<tr><td rowspan="2">skutečnost</td><td>platí H_0</td><td>správné rozhodnutí
pravděpodobnost $1-\alpha$
spolehlivost</td><td>chyba I. druhu
pravděpodobnost α
hladina významnosti</td></tr>
<tr><td>platí H_1</td><td>chyba II. druhu
pravděpodobnost β</td><td>správné rozhodnutí
pravděpodobnost $1-\beta$
síla testu</td></tr>
</table>

- Jelikož výběr na jehož základě rozhodujeme je **náhodný**, nelze se chybám I. a II. druhu vyhnout.
- Chtěli bychom mít k dispozici testy s **nízkou hladinou významnosti** a **vysokou sílou testu.**

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Volba významnosti testu:

V měkkých vědách (ekonomie, sociologie):

$\alpha = 0{,}10$	méně významné
$\alpha = 0{,}05$	významné
$\alpha = 0{,}01$	velmi významné
$\alpha = 0{,}001$	extrémně významné

V technických vědách obvykle $\alpha = 0{,}01$.

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Test hypotézy o střední hodnotě:

Dvojice testových hypotéz:

H_0: $\mu = \mu_0$

H_1: $\mu \neq \mu_0$

Předpoklady testu:

X je náhodný výběr z populace mající **normální rozdělení** s **neznámým** rozptylem.

Vhodná testová statistika ***t*-test**:

$$\boxed{t = \frac{\bar{x} - \mu_0}{s} \sqrt{n}}$$

normovaná hodnota výběrového průměru

Nulové rozdělení: **Studentovo rozdělení** s $n - 1$ stupni volnosti.

Testování hypotéz

Test hypotézy o střední hodnotě:

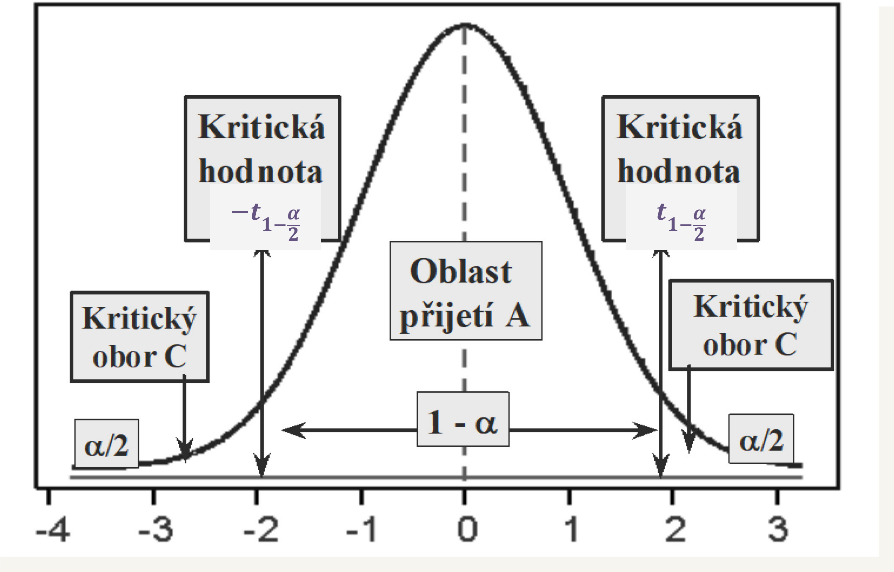
kritický kvantil $t_{1-\frac{\alpha}{2}}(n-1)$

přijetí nulové hypotézy:

$|t| < t_{1-\frac{\alpha}{2}}(n-1)$

zamítnutí nulové hypotézy:

$|t| > t_{1-\frac{\alpha}{2}}(n-1)$



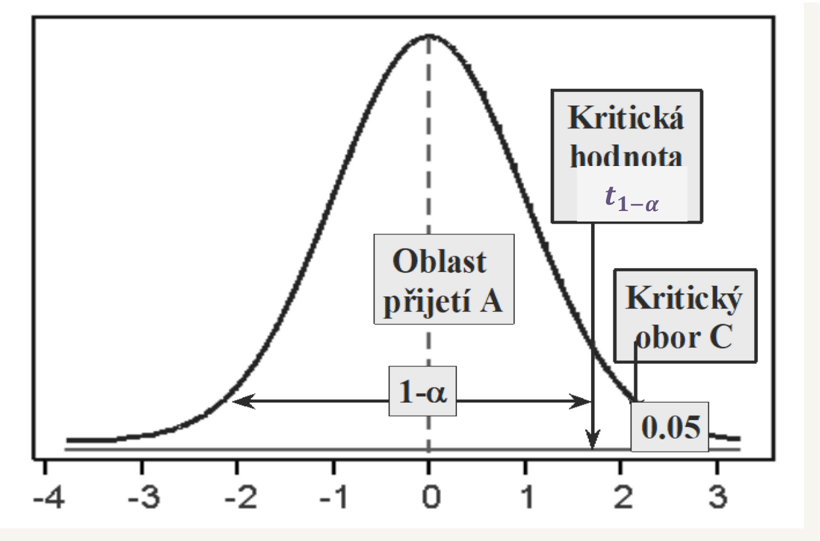
SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Jednostranné testy:

$H_0\colon \mu = \mu_0$
$H_1\colon \mu > \mu_0$
pro pravostrannou alternativu

$H_0\colon \mu = \mu_0$
$H_1\colon \mu < \mu_0$
pro levostrannou alternativu

kritický kvantil $t_{1-\alpha}$



zamítnutí nulové hypotézy:

$H_1\colon \mu > \mu_0$, $t > t_{1-\alpha}$

$H_1\colon \mu < \mu_0$, $t < -t_{1-\alpha}$

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Jak postupovat při testování hypotéz?

1. Formulujeme **nulovou a alternativní hypotézu**.
2. Zvolíme tzv. **testovou statistiku**, tj. výběrovou charakteristiku, jejíž rozdělení závisí na testovaném parametru. (Rozdělení testové statistiky za předpokladu platnosti nulové hypotézy nazýváme nulové rozdělení.)
3. Ověříme **předpoklady** testu!
4. Určíme **kritický obor $\boldsymbol{C}$**, tj. množinu, v níž se, za předpokladu platnosti H_0, hodnoty testové statistiky vyskytují s **velmi malou pravděpodobností**. Velikost kritického oboru zvolíme tak, abychom platnou hypotézu zamítali nejvýše s pravděpodobností α.
   - Doplňkem k C je tzv. **obor přijetí $\boldsymbol{A}$**.
   - Hranici mezi kritickým oborem a oborem přijetí označujeme jako **kritická hodnota** testu t_{krit}.
5. Na základě konkrétní realizace výběru určíme **pozorovanou hodnotu** testové statistiky.
6. Na základě vztahu mezi **pozorovanou hodnotu** testové statistiky a t_{krit} **rozhodneme o výsledku** testu („Zamítáme H_0.“ nebo „Nezamítáme H_0.“)

Příklad 2 – platy

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Podle tvrzení agentury ALFA je průměrný plat pracovníků počítačových firem 30 000 Kč. Průzkum zkoumal platy 14 náhodně vybraných zaměstnanců počítačových firem a zjistil tyto údaje: průměrný plat: 31 500 Kč, směrodatná odchylka: 620 Kč. Lze na základě tohoto průzkumu na hladině významnosti 5% potvrdit hypotézu agentury ALFA?

Pravostranný test střední hodnoty:

$H_0: \mu = \mu_0 = 30\,000$

$H_1: \mu > 30\,000$

$$t = \frac{\bar{x} - \mu_0}{s}\sqrt{n} = \frac{31500 - 30000}{620}\sqrt{14} = 8{,}72$$

kritická hodnota: $t_{0,95}(13) = 2{,}160$

$t > t_{0,95}(13)$ → hodnota leží v kritickém oboru → **zamítáme nulovou hypotézu**

Na hladině 0,05 průzkum nepotvrdil názor agentury ALFA, že průměrný plat v počítačových firmách je 30 000 Kč. Nulovou hypotézu tedy zamítáme.

Test hypotézy o podílu:

Dvojice testových hypotéz:

H_0: $\pi = \pi_0$
H_1: $\pi \neq \pi_0$
$\pi > \pi_0$
$\pi < \pi_0$

Vhodná testová statistika (***z***-test):

$$z = \frac{p - \pi_0}{\sqrt{\pi_0 (1 - \pi_0)}} \sqrt{n}$$

Musí být splněna pravidla pro aproximaci:

$$n \cdot \pi_0 \cdot (1 - \pi_0) > 9$$

Nulové rozdělení: **Normované normální rozdělení** s n stupni volnosti.

Příklad 3 – internet

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Na sídlišti Modrý vrch bylo osloveno 200 domácností, zda mají zájem o připojení k vysokorychlostnímu internetu. Zájem projevilo 87 z dotázaných domácností. Pro firmu AGS je instalace internetových rozvodů zajímavá, pokud jej bude využívat alespoň 40% domácností na sídlišti. Je pro firmu AGS sídliště Modrý vrch z tohoto hlediska zajímavé?

test hypotézy o podílu:

$H_0: \pi = \pi_0 = 0{,}4$

$H_1: \pi > 0{,}4$ (pravostranný test)

pravidla pro aproximaci:

$$n \cdot \pi_0 \cdot (1 - \pi_0) = 200 \cdot 0{,}4 \cdot 0{,}6 = 48 > 9$$ splněno

$$z = \frac{p - \pi_0}{\sqrt{\pi_0 (1 - \pi_0)}} \sqrt{n} = \frac{\frac{87}{200} - 0{,}4}{\sqrt{0{,}4 \cdot 0{,}6}} \sqrt{200} = 1{,}01$$

Příklad 3 – internet

Na sídlišti Modrý vrch bylo osloveno 200 domácností, zda mají zájem o připojení k vysokorychlostnímu internetu. Zájem projevilo 87 z dotázaných domácností. Pro firmu AGS je instalace internetových rozvodů zajímavá, pokud jej bude využívat alespoň 40% domácností na sídlišti. Je pro firmu AGS sídliště Modrý vrch z tohoto hlediska zajímavé?

p-hodnota testu: $z_{0,95} = 1{,}645$

$z < z_{0,95}$ → hodnota leží v oboru přijetí → **nezamítáme** nulovou hypotézu

Nelze tvrdit, že na sídlišti bude mít více než 40% domácností zájem o internet → pro firmu AGS je sídliště Modrý vrch nezajímavé.

Testování hypotéz

Test hypotézy o rozptylu:

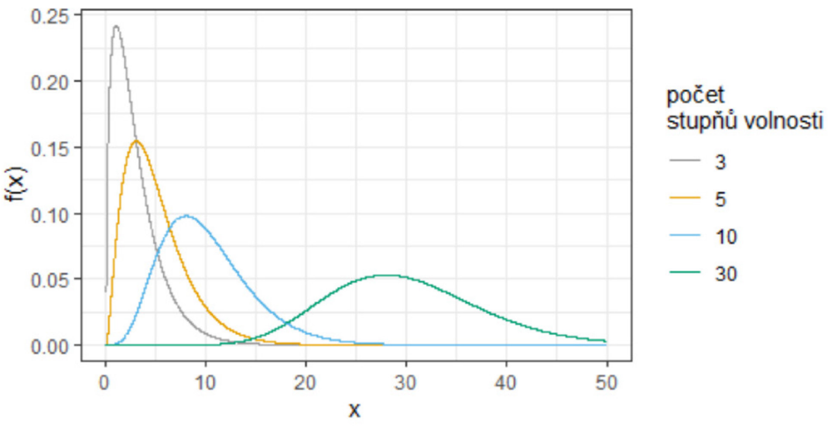
Jedná se o **Pearsonův χ^2-test:**

$H_0: \sigma^2 = \sigma_0^2$

$H_1: \sigma^2 \neq \sigma_0^2$

$\sigma^2 > \sigma_0^2$

$\sigma^2 < \sigma_0^2$



Předpoklady testu:

X je náhodný výběr z populace mající **normální rozdělení** s **neznámou** střední hodnotou.

Vhodná testová statistika (χ^2-test): $\chi^2 = \dfrac{s^2}{\sigma^2}(n-1)$

Nulové rozdělení: $\boldsymbol{\chi^2}$ **rozdělení** s $n-1$ stupni volnosti.

Příklad 4 – internet

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Internetové připojení pro 10 bytových jednotek má směrodatnou odchylku 12,2 Mb. Firma provedla úpravu pro zlepšení stability připojení. Náhodně bylo otestováno 15 bytových jednotek byla naměřena odchylka 10,9 Mb. Testujte na hladině významnosti 0,05 hypotézu, že úpravy zvýšily stabilitu připojení.

test hypotézy o rozptylu:

$H_0: \sigma^2 = 12{,}2^2$

$H_1: \sigma^2 < 12{,}2^2$ (levostranný test)

$$\chi^2 = \frac{s^2}{\sigma^2}(n-1) = \frac{10{,}9^2}{12{,}2^2}(14) = 11{,}175$$

kritická hodnota: $\chi^2_{0,05}(14) = 6{,}571$

$\chi^2 > \chi^2_{0,05}(14) \rightarrow$ hodnota neleží v kritickém oboru $\rightarrow$ **nezamítáme nulovou hypotézu**

Na hladině 0,05 průzkum nepotvrdil, že se zlepšila stabilita internetového připojení. Nulovou hypotézu tedy nezamítáme.

Literatura

1. Janáček J. *Statistika jednoduše*. Praha: Grada, 2022. **(kapitola 5).**
2. Ramík J. a Čemerková Š. *Statistika A*. Opava, Karviná: SLU, 2000. **(kapitola 8).**

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Děkuji za pozornost.