Testování hypotéz

9. seminář

Testování statistických hypotéz

Statistická hypotéza je výrok o statistickém souboru

Platnost statistických hypotéz se prověřuje pomocí **testů významnosti,** které rozhodují mezi:

- **Hypotézou nulovou** (testovanou) **H_0**

- Hypotézou alternativní (opačnou) **H_A**

Formulace H_0, H_A není nahodilá, je přesně specifikovaná statistikem při odvozování testu významnosti.

! H_0 se volí jako jednoduchá !

$=$ je rovno

$\neq$ není rovno – dvoustranná H

$>$ je větší – jedostranná H

Příklady statistických hypotéz

- Rozložení výšek 10-letých chlapců je normální (Gaussovo) - H_0, H_A není normální.
- 10-letí chlapci jsou větší než 10-letá děvčata - H_A, H_0 mají stejnou výšku $(\mu_1 = \mu_2) \equiv \mu_1 - \mu_2 = 0$
- Lék A je účinnější než lék B při léčbě hypertenze - H_A, H_0 léky jsou stejně účinné. $(\pi_A = \pi_B) \equiv \pi_A - \pi_B = 0$
- Kouření je rizikový faktor pro ICHS, IM, Ca plic - H_A, H_0 kouření není rizikový faktor.
- Existuje závislost mezi nízkou porodní hmotností a kojeneckou úmrtností - H_A, H_0 není závislost.

Příklad:

Je potřeba použít hodnocení hladiny cholesterolu různých norem (standardů) s přihlédnutím k věku?

Výsledky výběrového šetření (muži)

1. (20-30) roků $n_1 = 50$ $m_1 = 4{,}57$
   $s_1 = 0{,}70$ $SE_1 = 0{,}10$
2. (40-50) roků $n_2 = 60$ $m_2 = 5{,}42$
   $s_2 = 0{,}85$ $SE_2 = 0{,}11$

Orientační řešení pomocí CI

20-30 95% CI (4,37; 4,77)

40-50 95% CI (5,18; 5,66)

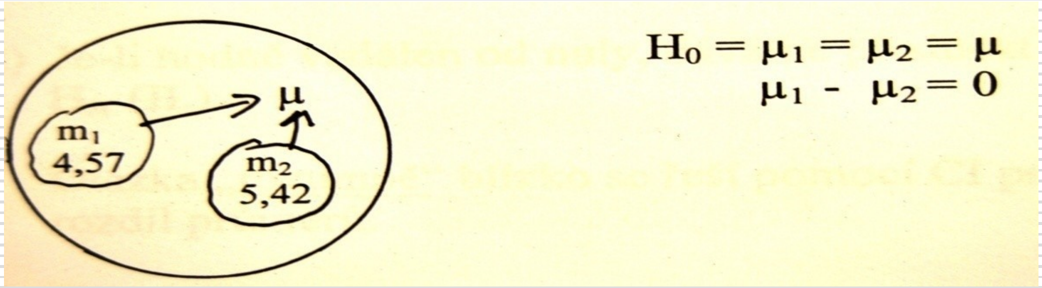
Objektivně lze rozhodnout pomocí testu významnosti <u>pro srovnání dvou průměrů</u>.

Vzhledem k tomu, že oba výběry jsou větší než 30, můžeme vycházet z modelu **normálního** (Gaussova) rozdělení. (μ_1, σ_1) ; (μ_2, σ_2)

H_0 /H_A

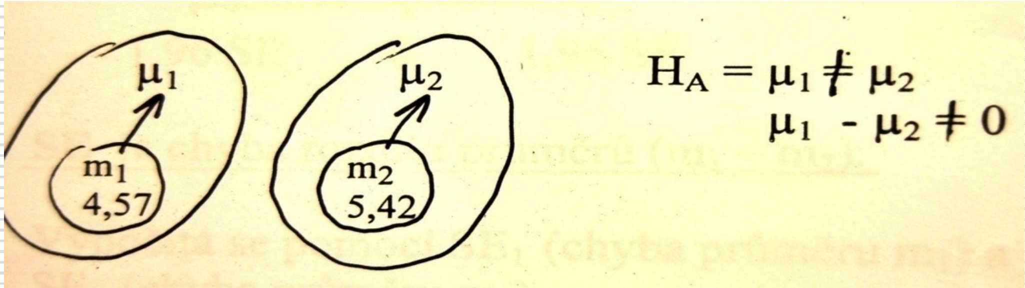
Nulová hypotéza (testovaná)

Předpokládá, že jde o dva náhodné výběry z jednoho základního souboru (rozdíl není).



Alternativní hypotéza (opačná)

Předpokládá, že jde o dva náhodné výběry ze dvou základních souborů s rozdílnými průměry.



Rozhodování mezi H_0 a H_A se zakládá na rozdílu $m_1 - m_2$ (5,42-4,57 = 0,85)

Rozhodování

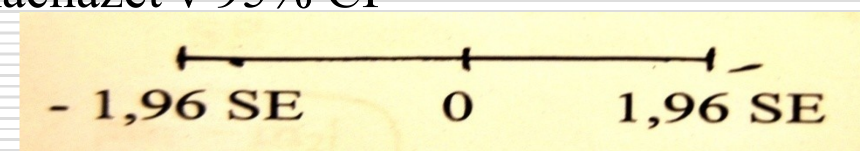
1) Pokud je rozdíl $m_1 - m_2$ (5,42-4,57= 0,85)

?? **rozumně** blízko nule, tzn., že se dá vysvětlit **<u>náhodou,</u>** rozhodujeme se pro H_0 (I.).

2) Je-li hodně vzdálen od nuly, dáváme přednost H_A (II.)

Otázka „**rozumně**“ blízko se řeší pomocí CI pro rozdíl průměrů.

Pokud H_0 platí ($\mu_1 = \mu_2 = \mu$), pak s pravděpodobností 0,95 by se měl rozdíl $m_1 - m_2$ nacházet v 95% CI



! SE je chyba rozdílu průměrů ($m_1 - m_2$).

Vypočítá se pomocí SE_1 (chyba průměru m_1) a SE_2 (chyba průměru m_2).

! Pro **nezávislé** výběry platí $SE^2 = SE_1^{\ 2} + SE_2^{\ 2}$ **!**

$$\frac{s_1^2}{n_1-1}+\frac{s_2^2}{n_2-1}$$

Rozhodování

Jak rozhodujeme?

1) Pokud je $m_1 - m_2$ mimo CI, **H_0 zamítáme**

$|m_1 - m_2| > 1{,}96.\ SE$

2) Pokud $m_1 - m_2$ padne do CI, **H_0 nezamítáme**

$|m_1 - m_2| \leq 1{,}96.\ SE$

! Nezamítnutí H_0 neznamená její **přijetí**

Formální úprava zápisu

ad1) $|m_1 - m_2| / SE > 1{,}96$ **H_0 zamítáme**

ad 2) $|m_1 - m_2| / SE \leq 1{,}96$ **H_0 nezamítáme**

= u = testovací charakteristika = > u-test

V anglické literatuře se používá i označení z ≥ **z-test**

Příklad: Cholesterol

Podmínky použitelnosti

1) n1, n2 > 30
2) nezávislé výběry = > **u-test**

   u = 5,70

Závěr:

u = 5,70 >2,58 = > H0 zamítáme na 1% HV, tzn., že je hodně malá pravděpodobnost, že se mýlíme, když přisuzujeme významný vliv věku.

Jak rozhodujeme?

Zamítnutí – pravděpodobnost, že rozdíl mezi průměry je způsoben náhodou je tak malá, že tuto možnost zamítáme – **dáváme přednost alternativní hypotézu**

Nezamítnutí – rozdíly nepřesahují velikost rozdílů způsobených náhodou

! Testování statistických hypotéz !

1. Stanovíme nulovou a alternativní hypotézu
2. Zvolíme hladinu významnosti
3. Vybereme vhodný test
4. Ověříme, zda jsou splněny podmínky pro použití testu
5. Vypočítáme testovací charakteristiku
6. Srovnáme ji s odpovídajícími kritickými hodnotami
7. Zamítneme nebo nezamítneme nulovou hypotézu
8. Výsledky interpretujeme

Příklad:

Srovnejte výšku tříletých brněnských chlapců a děvčat na podkladě výběrového šetření náhodně vybraných dětí:

CH: $n_1 = 80$ $m_1 = 97{,}4$ $s_1 = 3{,}8$

D: $n_2 = 80$ $m_2 = 96{,}3$ $s_2 = 3{,}7$

Řešení

Řešení:

1) $H_0 \equiv \mu_{CH} = \mu_D$ $H_A \equiv \mu_{CH} \neq \mu_D$
2) Zvolíme hladinu významnosti $\alpha = 0{,}05$ eventuálně 0,01
3) Ověříme podmínky použitelnosti u-testu

   $n_1, n_2 > 30$

   výběry jsou nezávislé

Příklad

4) Vypočítáme testovací charakteristiku

$$u = \frac{97{,}4 - 96{,}3}{\sqrt{\frac{3{,}8^2}{79} + \frac{3{,}7^2}{79}}} = 1{,}84$$

5) Závěr: protože u = 1,84 $<$ 1,96 H_0 nezamítáme.

6) Interpretace: Nezamítnutí H_0 neznamená její přijetí.

Správná formulace: Rozdíl v průměrných výškách tříletých chlapců a děvčat **není statisticky významný**. Nebylo by správné tvrdit, že rozdíl neexistuje.

Nezamítnutí H_0

Nezamítnutí H_0 ($\mu_1 = \mu_2$) představuje rozhodnutí dvojznačné. Buď nulová hypotéza platí, nebo neplatí, avšak na základě zjištěných výsledků se ji nepodařilo zamítnout.

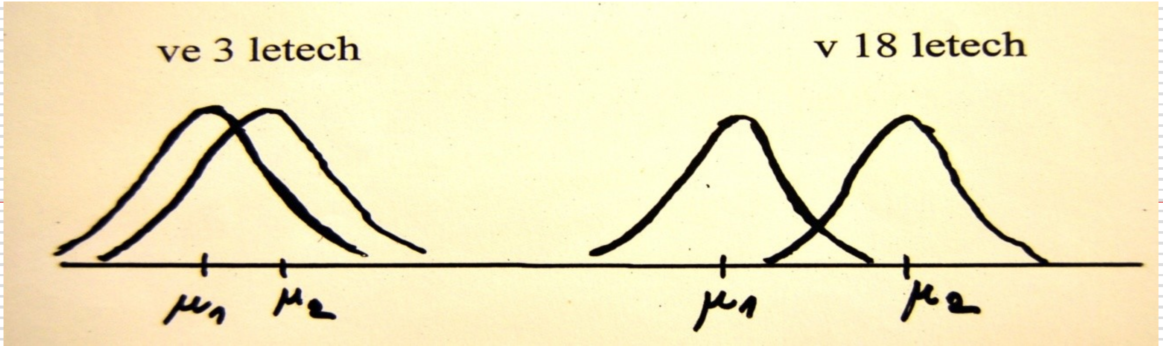
Příklad: s výškou chlapců a děvčat (skripta str. 23),

$u = 2{,}70$, což vede k zamítnutí H_0 ($n_1 = 170$, $n_2 = 172$)

<u>Rozdíl ve výškách chlapců a děvčat 1,1 cm se jako významný prokázal při větším počtu změřených dětí.</u>

! Závěr: Prokázání relativně malého rozdílu v průměrech vyžaduje větší počet měření.

Rozložení výšek chlapců a děvčat



Příklad na srovnání pravděpodobností

Byl sledován výskyt alergií u studentů LF:

Male: $n_1 = 105$ $k = 21$ $p = 0{,}20$ (20%)

Female: $n = 195$ $k = 19$ $p = 0{,}097$ (9,7%)

Otázka: Je rozdíl mezi pravděpodobností výskytu alergie u mužů a u žen způsoben náhodou, anebo lze odvodit, že alergie postihují častěji muže?

Postup:

1) Pro soubor mužů i pro soubor žen zjistit, zda je splněna podmínka pro použití u-testu.
2) Vypočítat SE rozdílů pravděpodobností

$$SE^2_{p_1-p_2} = \frac{p_1 \cdot (1-p_1)}{n_1} + \frac{p_2 \cdot (1-p_2)}{n_2}$$

3) Vypočítat testovací charakteristiku a porovnat ji s příslušnou kritickou hodnotou

Řešení

1/ $H_0 = \pi_1 = \pi_2$

$H_A = \pi_1 \neq \pi_2$

2/ HV: 5% (1%)

3/ PODM.: $n \cdot p \cdot (1-p) > 9$

$105 \cdot 0{,}2 \cdot 0{,}8 = 16{,}5 > 9$

$195 \cdot 0{,}097 \cdot 0{,}903 = 17{,}1 > 9$

4/ $$SE = \sqrt{\frac{0{,}2 \cdot 0{,}8}{105} + \frac{0{,}097 \cdot 0{,}903}{195}} =$$

$$= 0{,}044376 \doteq 0{,}044$$

5/ test. charakteristika

$$u = \frac{|0{,}2 - 0{,}097|}{0{,}044} = \underline{\underline{2{,}34}}$$

6/ $2{,}34 > 1{,}96 \Rightarrow H_0$ zamítáme (5%)

$2{,}34 < 2{,}58 \Rightarrow H_0$ nezamítáme (1%)

7/ INT. je stat. významný (5%)

není stat. významný (1%)

Srovnání pravděpodobností u-testem

Příklad:

V souboru 200 náhodně vybraných studentů LF byla zjištěna zraková vada u 80 studentů (p_1 = 80/200= 0,40,ev. 40%)

U 250 nestudujících stejného věku byla zraková vada zjištěna u 85 vyšetřovaných (p_2= 0,34, ev. 34%)

Srovnání pravděpodobností u-testem

Řešení: $H_0 \equiv \pi_1 = \pi_2$ není rozdíl ve výskytu zrakové vady u stud. a nestud. mládeže

$H_A \equiv \pi_1 \neq \pi_2$ je rozdíl

<u>Podmínky použití u-testu</u>

- Nezávislé výběry
- Konvergence binomického rozdělení k normálnímu (n.p.(1-p) > 9)

200 . 0,4 . (1-0,4) = 48 > 9, 250 . 0,34 . (1-0,34) = **56,1 > 9**

$$u = \frac{0{,}40 - 0{,}34}{\sqrt{\frac{0{,}40(1-0{,}40)}{200} + \frac{0{,}34(1-0{,}34)}{250}}} = 1{,}31 < 1{,}96$$

Závěr: Nulovou hypotézu nezamítáme, nepodařilo se prokázat, že by nestudující mládež měla významně méně zrakových vad než studenti LF.

Studentovo rozdělení t

Podmínka: n_1, $n_2 < 30$

Testovací charakteristika t =

$$t = \frac{m_1 - m_2}{s_{m_1 - m_2}} = \frac{m_1 - m_2}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2} \times \frac{n_1 + n_2}{n_1 \times n_2}}}$$

Počet stupňů volnosti f = $n_1 + n_2 - 2$

Kritické hodnoty viz skripta str. 23

Příklad

Srovnejte průměrnou porodní hmotnost u novorozenců matek silných kuřaček a nekuřaček na podkladě výběrového šetření u 30 novorozenců.

1) Nekuřačky: $n1 = 15$ $m1 = 3,59$ $s1 = 0,37$
2) Silné kuřačky: $n2 = 15$ $m2 = 3,20$ $s2 = 0,49$

Řešení

Řešení:

- Výběry jsou nezávislé
- n1, n2 < 30 => Studentův t-test

$$t = \frac{3{,}59 - 3{,}20}{\sqrt{\frac{15.0{,}37^2 + 15.0{,}49^2}{15+15-2} \cdot \frac{15+15}{15.15}}} = \frac{0{,}39}{0{,}1641}$$

$$= \underline{2{,}38} > t_{0,05}\,(28) = 2{,}05$$
$$< t_{0,01}\,(28) = 2{,}76$$

Závěr: Nulovou hypotézu ($\mu_1 = \mu_2$) zamítáme na 5% HV. Riziko, že se mýlíme, je mezi (1-5)%.

Děkuji za pozornost