Testování statistických hypotéz (1)

V lékařském výzkumu → neustále prověřujeme **předpoklady** a **domněnky** (cíl: potvrdit nebo vyvrátit)

možno formulovat jako **tzv. statistické hypotézy**

STAT metody: urychlují a objektivizují třídění domněnek (hypotéz) na:

- mylné
- správné

Testování statistických hypotéz (2)

Pojmy:

- Testovaná (nulová) hypotéza
- Alternativní hypotéza
- Hladina významnosti
- Kritické hodnoty

atd.

Testování statistických hypotéz (3)

STATISTICKÁ HYPOTÉZA = výrok o statistickém souboru

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

- **Hypotézou nulovou** (testovanou) $\mathbf{H_0}$
- **Hypotézou alternativní** (opačnou) $\mathbf{H_A}$

Formulace H_0, H_A <u>není nahodilá</u> x je přesně specifikovaná statistikem při odvozování testu významnosti.

! $\mathbf{H_0}$ se volí jako jednoduchá !

Příklady statistických hypotéz

1. H_0: Rozložení výšek 10-letých chlapců je normální (Gaussovo).
   H_A: Není normální.
2. H_A: 10-letí chlapci jsou větší než 10-letá děvčata.
   H_0: Mají stejnou výšku.
   $(\mu 1 = \mu 2) \equiv \mu 1 - \mu 2 = 0$
3. H_A: Lék A je účinnější než lék B při léčbě hypertenze.
   H_0: Léky jsou stejně účinné.
   $(\pi A = \pi B) \equiv \pi A - \pi B = 0$
4. H_A: Kouření je rizikový faktor pro ICHS, IM, Ca plic.
   H_0: Kouření není rizikový faktor. (RR = 1)
5. H_A: Existuje závislost mezi nízkou porodní hmotností a kojeneckou úmrtností.
   H_0: Není závislost.

Příklad (1):

Je potřeba použít pro 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	$n_2 = 60$	$m_2 = 5{,}42$
		$s_2 = 0{,}85$	$SE_2 = 0{,}11$

Příklad (2):

Orientační řešení pomocí CI

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

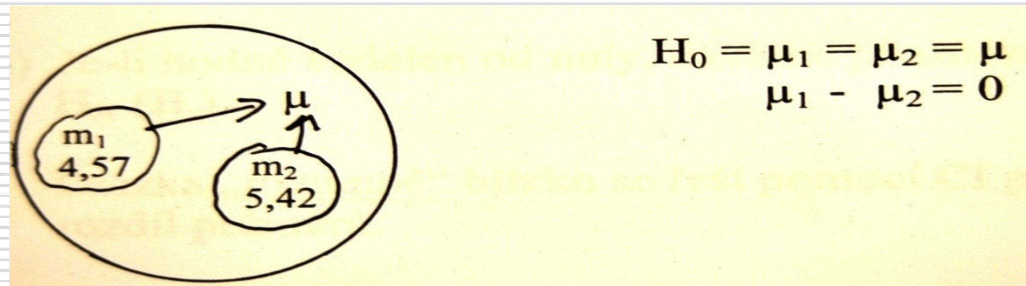
40-50 95% CI (5,20; 5,64)

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

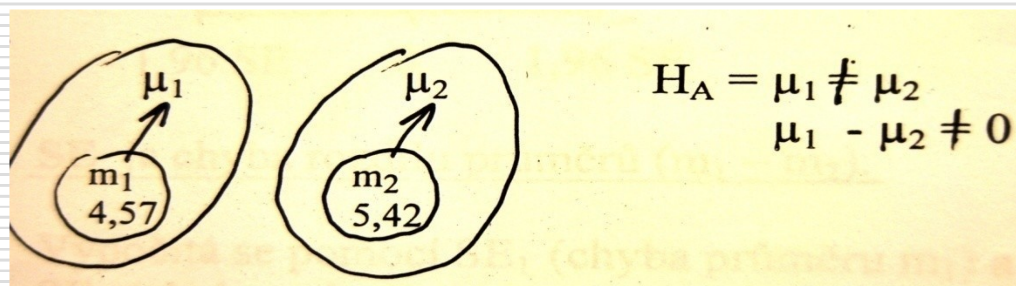
Vzhledem k tomu, že oba výběry jsou <u>větší než 30</u>, můžeme vycházet z modelu **normálního** (Gaussova) rozdělení. (μ_1, σ_1) ; (μ_2, σ_2)

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)

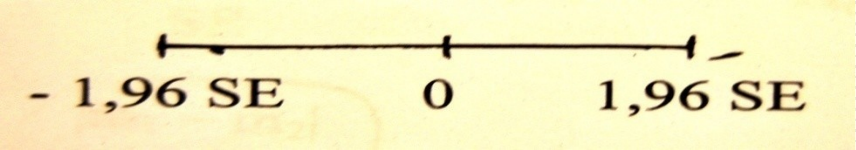
1) ?? **rozumně** blízko nule, tzn., že se dá vysvětlit **<u>náhodou</u>** → rozhodujeme se pro $\mathbf{H_0}$

2) Je-li hodně vzdálen od nuly, dáváme přednost $\mathbf{H_A}$

Rozhodování (2)

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

Pokud H0 platí ($\mu 1 = \mu 2 = \mu$), pak s pravděpodobností 0,95 by se měl rozdíl m1 – m2 nacházet v 95% CI



! SE je chyba rozdílu průměrů (m1 – m2).

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 (3)

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 <u>neznamená její **přijetí**</u>

Rozhodování (4)

Formální úprava zápisu:

ad1) $|m_1 - m_2|$ /SE > 1,96 → **H0 zamítáme**

ad 2) $\underline{|m_1 - m_2| \ /SE} \leq 1,96$ → **H0 nezamítáme**

u = testovací charakteristika = > u-test

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

Příklad: Cholesterol

Podmínky použitelnosti:

1) n_1, $n_2 > 30$
2) nezávislé výběry = > **u-test**

   u = 5,70

Závěr:

u = 5,70 >2,58 = > **H_0 zamítáme na 1% HV**, tzn., že je hodně malá pravděpodobnost, že se mýlíme, když přisuzujeme významný vliv věku.

Jak rozhodujeme? (1)

Zamítnutí –pravděpodobnost, že rozdíl mezi průměry je způsoben náhodou, je tak malá, že tuto možnost (H_0) zamítáme – a **přijímáme alternativní hypotézu (H_A)**

- riziko chyby prvního typu

Nezamítnutí – rozdíly nepřesahují velikost rozdílů způsobených náhodou, ale mohla nastat tzv. **chyba druhého typu.**

	Skutečnost	
Naše rozhodnutí	**H_0 neplatí**	**H_0 platí**
Zamítáme H_0	Správné rozhodnutí	Chyba 1. typu α
Nezamítáme H_0	Chyba 2. typu β	Správné rozhodnutí

Jak rozhodujeme? (2)

IF **spolehlivost rozhodování**: 95% , resp. 99%

→ **riziko (nesprávnost) rozhodování**:   5%, resp. 1%

Je to pravděpodobnost, že rozhodnutí je špatné → pravděpodobnost **chyby 1.druhu** nebo tzv. **HLADINA VÝZNAMNOSTI α**

pro 95% → konstanta 1,96 → 5 % kritická hodnota

Pro 99% → konstanta 2,58 → 1% kritická hodnota

Jak rozhodujeme? (3)

- rozhodování doprovázeno **chybou 2.druhu β**, kt.vyjadřuje nesprávné zamítnutí H_A

→ stará se o ni statistik při odvozování testu

! Testování statistických hypotéz !

1. Stanovíme nulovou a alternativní hypotézu
2. Zvolíme hladinu významnosti *(HV = 5% nebo 1%)*
3. Vybereme vhodný test *(u-test; t-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:	$n_2 = 80$	$m_2 = 96{,}3$	$s_2 = 3{,}7$

Nezamítnutí H_0

Nezamítnutí H0 (μ1 = μ2) 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),

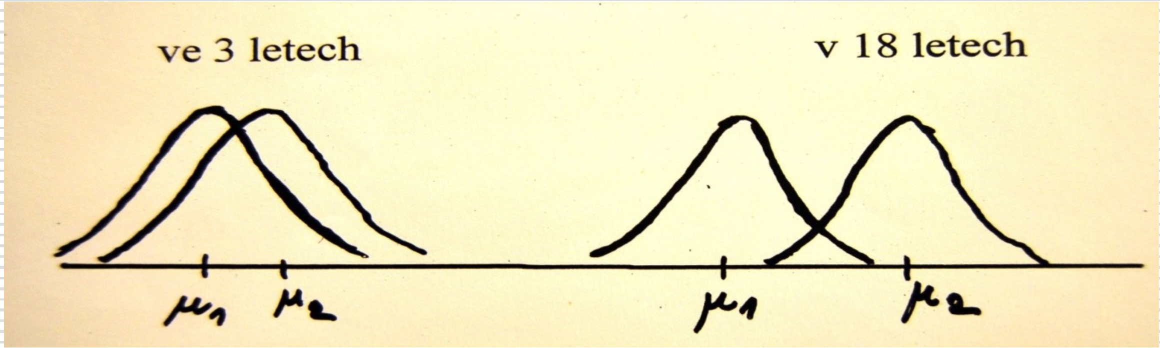
u = 2,70, což vede k zamítnutí H_0 (n1 = 170, n2 = 172)

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

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

Nezamítnutí H_0

Rozložení výšek chlapců a děvčat



Příklad na srovnání pravděpodobností (1)

Byl sledován <u>výskyt alergií</u> u studentů LF:

Muži:	$n_1 = 105$	$k = 21$	$p = 0{,}20$ (20%)
Ženy:	$n_2 = 195$	$k = 19$	$p = 0{,}097$ (9,7%)

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

Příklad na srovnání pravděpodobností (2)

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

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%)

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. 25

Příklad

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

1) Nekuřačky: $n_1 = 15$ $m_1 = 3,59$ $s_1 = 0,37$
2) Silné kuřačky: $n_2 = 15$ $m_2 = 3,20$ $s_2 = 0,49$