9. SEMINÁŘ

INDUKTIVNÍ STATISTIKA

2. TESTOVÁNÍ STATISTICKÝCH HYPOTÉZ

TESTOVÁNÍ STATISTICKÝCH HYPOTÉZ

- **Při testování statistických hypotéz** vycházíme z údajů zjištěných ve výběrovém souboru - jde o **induktivní soud**.

- **Statistická hypotéza =** výrok o statistickém souboru (není nutně totožná s výzkumnou hypotézou!!!)

- K ověření (testování) hypotézy se používá tzv. testů významnosti, které rozhodují mezi:

  – nulovou (testovanou) hypotézou H_0
  – hypotézou alternativní (opačnou) H_A

Příklad: SROVNÁVÁNÍ PRŮMĚRŮ

Je vhodné použít pro hodnocení hladiny cholesterolu různých norem s přihlédnutím k věku?

- **Věcná** (klinická) významnost
- **Statistická** významnost

Příklad:

Je vhodné použít pro hodnocení hladiny cholesterolu různých norem s přihlédnutím k věku?

Muži 20-30 let: $n_1 = 50$ $m_1 = 4{,}57$ $s_1 = 0{,}70$ $SE_1 = 0{,}10$
Muži 40-50 let: $n_2 = 60$ $m_2 = 5{,}42$ $s_2 = 0{,}85$ $SE_2 = 0{,}11$

Statistickou významnost lze odhadnout pomocí intervalů spolehlivosti:

1. Pokud se **intervaly spolehlivosti, které vytvoříme kolem bodových odhadů m_1 a m_2** překrývají, pak rozdíl mezi nimi není statisticky významný. Naopak, pokud se nepřekrývají, je rozdíl statisticky významný.

   $m_1 = 4{,}57$ 95% CI (4,37; 4,77) $m_2 = 5{,}42$ 95% CI (5,18; 5,66)

2. Pro řešení úlohy bychom mohli použít i **intervalový odhad rozdílu průměrů** – pokud CI neobsahuje nulu, je rozdíl statisticky významný.

   95% CI (0,56; 1,14)

Příklad:

Je vhodné použít pro hodnocení hladiny cholesterolu různých norem s přihlédnutím k věku?

Statistickou významnost lze určit testováním statistické hypotézy o rozdílu průměrů $m_1 - m_2$.

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

TESTOVÁNÍ STATISTICKÝCH HYPOTÉZ

1. Stanovíme nulovou a alternativní hypotézu

NULOVÁ A ALTERNATIVNÍ HYPOTÉZA

Při testování hypotéz začínáme tím, že předpokládáme určitou hodnotu parametru základního souboru a potom učiníme závěr týkající se výběrového statistického ukazatele.

Nulová hypotéza H_0 - testovaná

- Předpokládá, že rozdíl mezi parametrem a výběrovým ukazatelem je blízký nule – nulová hypotéza.

  ($\mu = m$; $\mu - m = 0$)

Alternativní hypotéza H_A - opačná

- Předpokládá opak, tj. že rozdíl mezi parametrem a výběrovým ukazatelem je nenulový (hodně vzdálen od nuly).

  ($\mu \neq m$, $\mu - m \neq 0$)

Příklad: SROVNÁVÁNÍ PRŮMĚRŮ

Nulová hypotéza (testovaná)

- vždy předpokládá, že jde o dva náhodné výběry z jednoho základního souboru (rozdíl mezi průměry není statisticky významný)

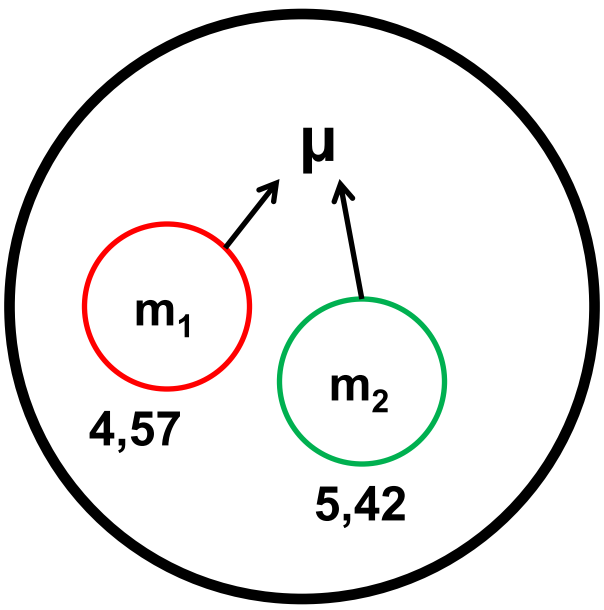
$$H_0: \mu_1 = \mu_2 = \mu$$

$$\mu_1 - \mu_2 = 0$$

Příklad: SROVNÁVÁNÍ PRŮMĚRŮ

Nulová hypotéza (testovaná)

- vždy předpokládá, že jde o dva náhodné výběry z jednoho základního souboru (rozdíl mezi průměry není statisticky významný)



$$H_0: \mu_1 = \mu_2 = \mu$$

$$\mu_1 - \mu_2 = 0$$

Příklad: SROVNÁVÁNÍ PRŮMĚRŮ

Alternativní hypotéza (opačná)

- předpokládá opak, tj. že jde o dva výběry ze dvou různých základních souborů s rozdílnými průměry (rozdíl mezi průměry je statisticky významný)

$$H_A: \mu_1 \neq \mu_2$$

$$\mu_1 - \mu_2 \neq 0$$

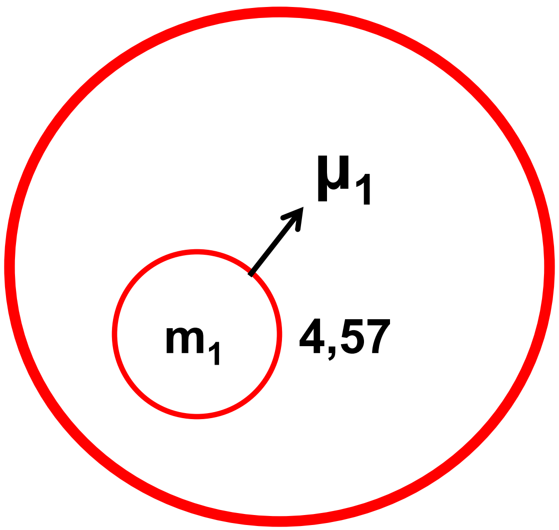
Příklad: SROVNÁVÁNÍ PRŮMĚRŮ

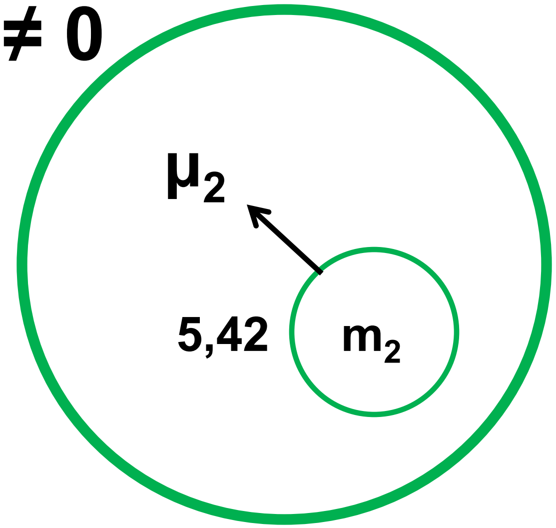
Alternativní hypotéza (opačná)

- předpokládá opak, tj. že jde o dva výběry ze dvou různých základních souborů s rozdílnými průměry (rozdíl mezi průměry je statisticky významný)

$$H_A:\ \mu_1 \neq \mu_2$$

$$\mu_1 - \mu_2 \neq 0$$





TESTOVÁNÍ STATISTICKÝCH HYPOTÉZ

1. Stanovíme nulovou a alternativní hypotézu
2. Zvolíme hladinu významnosti

HLADINA VÝZNAMNOSTI

- Je-li pravděpodobnost nějakého jevu velmi malá, chováme se (většinou) tak, jako by nemohla vůbec nastat.
- Je-li malá pravděpodobnost, že H_0 platí, chováme se tak, jako by neplatila a zamítáme ji.
- Tato malá pravděpodobnost se nazývá **hladina významnosti**, obvykle **α** = 0,05 nebo 0,01. Vyjadřuje riziko nesprávného zamítnutí H_0, **tzv. chyba 1. druhu**.
- **β** ozn. **chybu 2. druhu**, souvisí se silou statistického testu . Nastává, když H_0 nezamítáme, přestože ve skutečnosti neplatí.
- **Síla testu** = 1- β: schopnost zamítnout H_0, když neplatí.

Příklad: SROVNÁVÁNÍ PRŮMĚRŮ

Je vhodné použít pro hodnocení hladiny cholesterolu různých norem s přihlédnutím k věku?

Muži 20-30 let: $n_1 = 50$ $m_1 = 4{,}57$ $s_1 = 0{,}70$ $SE_1 = 0{,}10$

Muži 40-50 let: $n_2 = 60$ $m_2 = 5{,}42$ $s_2 = 0{,}85$ $SE_2 = 0{,}11$

Hladinu významnosti si zvolíme např. $\alpha = 0{,}05$.

TESTOVÁNÍ STATISTICKÝCH HYPOTÉZ

1. Stanovíme nulovou a alternativní hypotézu
2. Zvolíme hladinu významnosti
3. Vybereme vhodný test

TESTY VÝZNAMNOSTI

- Platnost statistických hypotéz prověřujeme pomocí tzv. **testů významnosti**:
  - **Testy pro hodnoty parametrů**
    (měříme vzdálenost pozorované statistiky
    od hypotézou stanovené hodnoty parametru)
  - **Srovnávání rozdílů parametrů**
    (např. test významnosti pro rozdíly středních hodnot
    či pravděpodobností)
  - **Zjišťování typu rozložení četností**
    (test dobré shody, test normality)
  - **Hodnocení závislostí**
    (testy závislosti)

TESTY VÝZNAMNOSTI

Parametrické testy

- Vycházejí ze srovnávání parametrů **μ, σ, π** (zastoupených při srovnávání výběrovými charakteristikami m, p, s).
- Musíme znát typ rozložení testované veličiny, hypotézy se týkají parametrů tohoto rozložení.
- Srovnáváme charakteristiky dvou **nezávislých** výběrů.

TESTY VÝZNAMNOSTI

Neparametrické testy

- Velkou skupinu tvoří např. testy založené na pořadí
- Výhody: jsou početně jednodušší a **nepředpokládají znalost typu rozložení** a lze je použít pro **závislé** výběry
- Nevýhody: mají menší sílu, tzn. mají menší schopnost zamítnout nulovou hypotézu, když ta skutečně neplatí.

VÝBĚR VHODNÉHO TESTU

u-test (z-test):

- Parametrický test
- Normální rozložení
- Vypočítaná testovací charakteristika **u** (někdy ozn. z) se srovnává s kritickými hodnotami normálního rozložení

Příklad: SROVNÁVÁNÍ PRŮMĚRŮ

Je vhodné použít pro hodnocení hladiny cholesterolu různých norem s přihlédnutím k věku?

Muži 20-30 let: $n_1 = 50$ $m_1 = 4{,}57$ $s_1 = 0{,}70$ $SE_1 = 0{,}10$

Muži 40-50 let: $n_2 = 60$ $m_2 = 5{,}42$ $s_2 = 0{,}85$ $SE_2 = 0{,}11$

Pro srovnání průměrů zvolíme u-test

- **při dostatečně velkých souborech mají rozdíly výběrových průměrů normální rozdělení**

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

PODMÍNKY PRO POUŽITÍ TESTU

Podmínky pro použití u-testu pro srovnávání průměrů:

1. **$n_1 > 30$, $n_2 > 30$**
   – pro menší soubory Studentův t-test (vypočítáme testovací charakteristiku **t** a srovnáme ji s kritickými hodnotami Studentova rozdělení – viz skripta str. 41).
2. **nezávislé výběry** (hodnoty ve srovnávaných souborech se vzájemně neovlivňují)
   – testy pro párované hodnoty
3. **stejné rozptyly**
   – neliší se významně (F-test)

Příklad: SROVNÁVÁNÍ PRŮMĚRŮ

Je vhodné použít pro hodnocení hladiny cholesterolu různých norem s přihlédnutím k věku?

Muži 20-30 let: $n_1 = 50$ $m_1 = 4{,}57$ $s_1 = 0{,}70$ $SE_1 = 0{,}10$

Muži 40-50 let: $n_2 = 60$ $m_2 = 5{,}42$ $s_2 = 0{,}85$ $SE_2 = 0{,}11$

Podmínky pro použití u-testu:

1. **50 > 30; 60 > 30**
2. **soubory jsou nezávislé**
3. **předpokládáme stejné rozptyly**

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

TESTOVACÍ CHARAKTERISTIKA

- Testy významnosti rozhodují mezi $\mathbf{H_0}$ a $\mathbf{H_A}$, a to nejčastěji pomocí výpočtu tzv. **testovací charakteristiky**

- Vymezuje obor hodnot pro zamítnutí a obor hodnot pro nezamítnutí $H_{0.}$

- Pro stanovení takových oborů hodnot je nezbytné, aby měla některé ze známých teoretických rozdělení – umožní to stanovení tzv. **kritických hodnot.**

- Kritické hodnoty vymezují **interval spolehlivosti**, jenž je mírou vzdálenosti od 0. Leží-li hodnota testovací charakteristiky mimo tento interval, zamítáme $H_{0.}$

VZDÁLENOST OD NULY

- Pokud je rozdíl srovnávaných průměrů **rozumně blízko nule**, pak můžeme říct, že rozdíl vznikl náhodou a **nezamítáme nulovou hypotézu.**

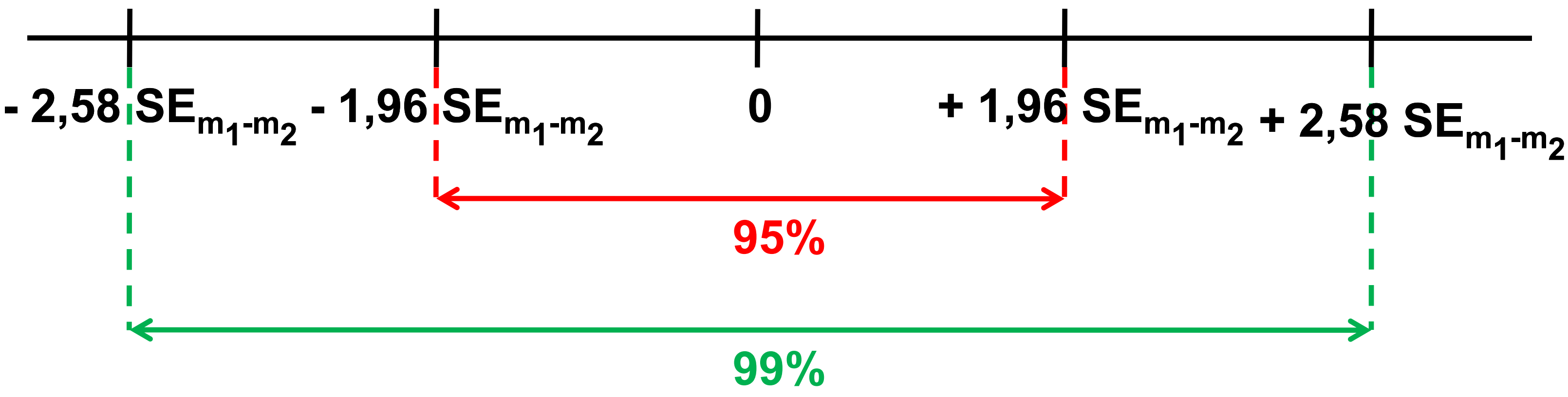
- Je-li rozdíl **hodně vzdálen od nuly**, dáváme přednost alternativní hypotéze, tj. **zamítáme nulovou hypotézu**.

VZDÁLENOST OD NULY

Chyba rozdílu průměrů

- Rozdíly průměrů mají normální rozdělení s parametry **μ** a **σ**; σ odhadujeme pomocí SE

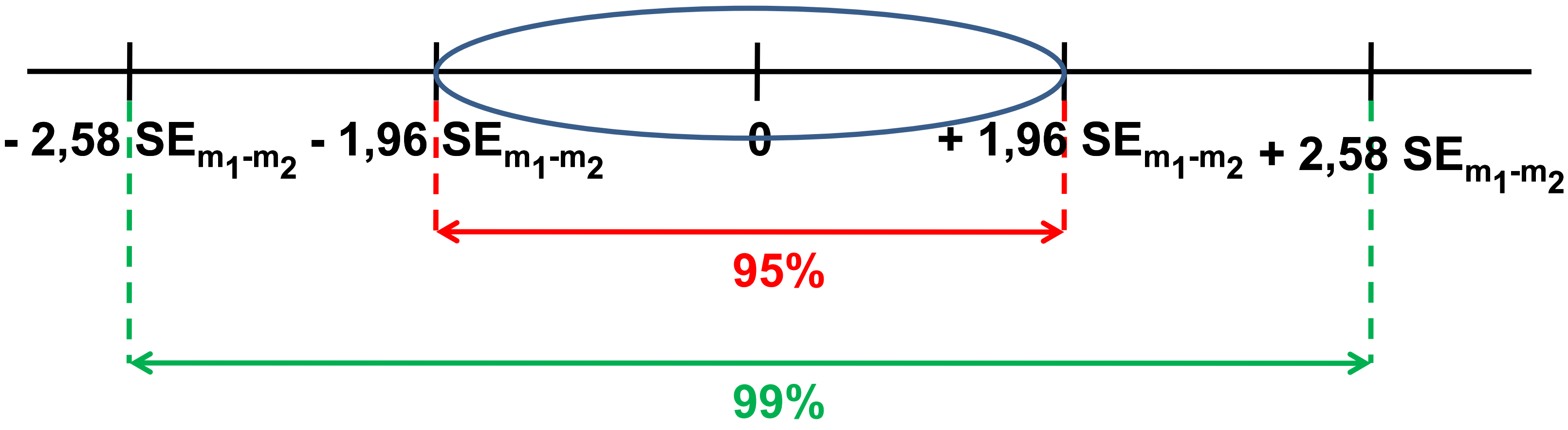
- $\mathbf{SE_{m_1\text{-}m_2}}$ = chyba rozdílu průměrů ($m_1 - m_2$), přičemž pro nezávislé výběry platí: $\mathbf{SE^2_{m_1\text{-}m_2} = SE^2_{m_1} + SE^2_{m_2}}$



VZDÁLENOST OD NULY

- Řeší se pomocí intervalu spolehlivosti 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% intervalu spolehlivosti.



JAK ROZHODUJEME?

Testovací charakteristika „u“

- Pokud leží rozdíl mimo interval spolehlivosti, pak **zamítáme** nulovou hypotézu.

$$|m_1 - m_2| \geq 1{,}96 SE_{m_1 - m_2} = \underbrace{\frac{|m_1 - m_2|}{SE_{m_1 - m_2}}}_{u} \geq 1{,}96$$

- Pokud leží rozdíl v intervalu spolehlivosti, pak nulovou hypotézu **nezamítáme**.

$$|m_1 - m_2| < 1{,}96 SE_{m_1 - m_2} = \frac{|m_1 - m_2|}{SE_{m_1 - m_2}} < 1{,}96$$

- **Nezamítnutí nulové hypotézy neznamená její přijetí!!!**

Příklad: SROVNÁVÁNÍ PRŮMĚRŮ

Je potřeba použít pro hodnocení hladiny cholesterolu různých norem s přihlédnutím k věku?

Muži 20-30 let: $n_1 = 50$ $m_1 = 4,57$ $s_1 = 0,70$ $SE_{m_1} = 0,10$

Muži 40-50 let: $n_2 = 60$ $m_2 = 5,42$ $s_2 = 0,85$ $SE_{m_2} = 0,11$

Výpočet testovací charakteristiky u:

$m_1 – m_2 = 4,57 – 5,42 = -0,88$

$SE_{m_1-m_2}{}^2 = 0,10^2 + 0,11^2 = 0,0221$

$SE_{m_1-m_2} = 0,15$

$u = 0,88 : 0,15 = 5,66$

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

Příklad: SROVNÁVÁNÍ PRŮMĚRŮ

Je vhodné použít pro hodnocení hladiny cholesterolu různých norem s přihlédnutím k věku?

Muži 20-30 let: $n_1 = 50$ $m_1 = 4{,}57$ $s_1 = 0{,}70$ $SE_1 = 0{,}10$

Muži 40-50 let: $n_2 = 60$ $m_2 = 5{,}42$ $s_2 = 0{,}85$ $SE_2 = 0{,}11$

Srovnání testovací charakteristiky s kritickou hodnotou:

- $|5{,}66| > 1{,}96$
- Testovací charakteristika je větší než kritická hodnota pro $\alpha = 0{,}05$, tzn. leží mimo 95% CI.
- **Malé soubory – t-test: kritické hodnoty Studentova t-rozdělení (skripta str. 41).** Stupně volnosti $f = (n_1 + n_2 - 2)$.

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

JAK ROZHODUJEME?

- **Nezamítnutí H_0**– rozdíly nepřesahují velikost rozdílů způsobených náhodou, **ale** mohla nastat tzv. chyba druhého typu.

- **Zamítnutí H_0**– pravděpodobnost, že rozdíl mezi průměry je způsoben náhodou je tak malá, že tuto možnost zamítáme – a přijímáme alternativní hypotézu (riziko chyby prvního typu).

JAK ROZHODUJEME?

	Skutečnost	
Naše rozhodnutí	H_0 neplatí	H_0 platí
Zamítáme H_0	Správné rozhodnutí	Chyba I. typu
Nezamítáme H_0	Chyba II. typu	Správné rozhodnutí

JAK ROZHODUJEME?

P-value

- udává pravděpodobnost, že hodnocený rozdíl je způsoben náhodou
- pokud je menší než zvolená hladina významnosti, nulovou hypotézu zamítáme, pokud je větší nulovou hypotézu nezamítáme
- Např.: α = 5% (pravděpodobnost platnosti H_0)

  p-value = 0,00073, zamítáme H_0

  p-value = 0,07300, nezamítáme H_0

Příklad: SROVNÁVÁNÍ PRŮMĚRŮ

Je vhodné použít pro hodnocení hladiny cholesterolu různých norem s přihlédnutím k věku?

Muži 20-30 let: $n_1 = 50$ $m_1 = 4{,}57$ $s_1 = 0{,}70$ $SE_1 = 0{,}10$
Muži 40-50 let: $n_2 = 60$ $m_2 = 5{,}42$ $s_2 = 0{,}85$ $SE_2 = 0{,}11$

Zamítnutí nebo nezamítnutí nulové hypotézy:

$|5{,}66| > 1{,}96$

Na 5% hladině významnosti zamítáme nulovou hypotézu a přijímáme hypotézu alternativní, tj. rozdíl mezi mladšími a staršími muži je statisticky významný.

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: SROVNÁVÁNÍ PRŮMĚRŮ

Je vhodné použít pro hodnocení hladiny cholesterolu různých norem s přihlédnutím k věku?

Muži 20-30 let: $n_1 = 50$ $m_1 = 4{,}57$ $s_1 = 0{,}70$ $SE_1 = 0{,}10$

Muži 40-50 let: $n_2 = 60$ $m_2 = 5{,}42$ $s_2 = 0{,}85$ $SE_2 = 0{,}11$

Interpretace výsledků:

Na 5% hladině významnosti jsme prokázali, že existuje statisticky významný rozdíl v průměrných hodnotách hemoglobinu u dvou srovnávaných věkových skupin.

K odpovědi na uvedenou otázku je třeba posoudit i **věcnou (klinickou) významnost** zjištěného rozdílu.

SHRNUTÍ PŘÍKLADU

1. H_0: $\mu_1 = \mu_2 = \mu$; $\mu_1 - \mu_2 = 0$
   H_A: $\mu_1 \neq \mu_2$; $\mu_1 - \mu_2 \neq 0$
2. $\alpha = 0{,}05$
3. u-test
4. $n_1 > 30$; $n_2 > 30$; nezávislé soubory; stejné rozptyly
5. $|u| = 5{,}66$
6. $5{,}66 > 1{,}96$
7. Zamítáme nulovou hypotézu a přijímáme hypotézu alternativní.
8. Rozdíl mezi průměrnými hodnotami cholesterolu je statisticky významný.

Příklad: SROVNÁNÍ PRAVDĚPODOBNOSTÍ

Byl sledován výskyt alergií u studentů LF.

Muži: $n_1 = 105$ $k_1 = 21$ $p_1 = 0{,}20$ (20%)

Ženy: $n_2 = 195$ $k_2 = 19$ $p_2 = 0{,}097$ (9,7%)

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

Příklad: SROVNÁNÍ PRAVDĚPODOBNOSTÍ

- **Podmínka pro použití u-testu:**

  $n_1 \cdot p_1 \cdot (1 - p_1) > 9$

  $n_2 \cdot p_2 \cdot (1 - p_2) > 9$

- **Standardní chyba rozdílu pravděpodobností**

$$SE^2 = SE_1^2 + SE_2^2 = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}$$

- **Testovací charakteristika u;** $u = \frac{p_1 - p_2}{SE}$

Řešení

1. H_0: $\pi_1 = \pi_2 = \pi$; $\pi_1 - \pi_2 = 0$
   H_A: $\pi_1 \neq \pi_2$; $\pi_1 - \pi_2 \neq 0$
2. a) $\alpha = 0{,}05$
   b) $\alpha = 0{,}01$
3. u-test
4. velikost souboru: $n_1 > 30$; $n_2 > 30$
   platnost nerovnosti: $16{,}8 > 9$; $17{,}1 > 9$
   nezávislé soubory
5. $u = 2{,}34$
6. a) $2{,}34 > 1{,}96$
   b) $2{,}34 < 2{,}58$
7. a) Na 5% hladině významnosti nulovou hypotézu zamítáme a přijímáme hypotézu alternativní.

   b) Na 1% hladině významnosti nulovou hypotézu nezamítáme.
8. a) Častější výskyt alergie u mužů je statisticky významný (riziko chyby 1.druhu).

   b) Nepodařilo se prokázat, že by alergie byly častější u mužů (riziko chyby 2. druhu).