Základy zpracování geologických dat

testování statistických hypotéz

R. Čopjaková

Studentův t-test

- Předpoklad normality výběrových souborů
- **Studentův t-test** - často používaná metoda testování statistických hypotéz. V závislosti na situaci, kdy se používá, rozlišujeme 4 typy Studentova t-testu:
- **jednovýběrový t-test**, který slouží k porovnání střední hodnoty výběrového souboru s konstantou (H_0: $x = \mu_0$) (viz. přednáška č. 8)

  jednovýběrový t-test o střední hodnotě
- **dvouvýběrový t-test**, který slouží k porovnání středních hodnot dvou výběrových souborů (H_0: $\mu_1 - \mu_2 =$ *konstanta; nejčastěji 0*);
  - **dvouvýběrový t-test párový – rozsahy obou výběrů jsou stejné** N1=N2; Opakované přeměřování stejných vzorků; Závislost mezi náhodnou veličinou X a Y

    párový t-test shodnosti výsledků
  - **dvouvýběrový t-test nepárový – rozsah výběrů nemusí být stejný** N1 nemusí být rovno N2; Proměřování dvou sad různých vzorků; Nezávislost mezi náhodnou veličinou X a Y

    a) *t-test shodnosti výsledků při rovnosti rozptylů* $\sigma_1^2 = \sigma_2^2$

    b) *t-test shodnosti výsledků při nerovnosti rozptylů* $\sigma_1^2 \neq \sigma_2^2$

Testy shodnosti výsledků dvouvýběrové – T-testy

- Liší se hodnoty naměřené na stejných přístrojích v různých laboratořích? (např. data z EMP v Brně a Barrandově)
- Liší se výsledky získané různými analytickými metodami (např. hodnoty naměřené přenosným terénním gama-spektrometrem a laboratorním gama-spektrometrem)
- Liší se hodnoty naměřené v různých časových intervalech (sezónní vlivy v hydrogeologii)
- Liší se hodnoty naměřené v různých místech (např. srovnání chemického složení – protolitu- ortorul sněžnických a gieraltovských orlicko-kladského krystalinika)
- Byla dekontaminace účinná, snížilo se znečištění?

- **Tedy při objektivním porovnávání výsledků analýz na dvou souborech experimentálně získaných dat**

Testy shodnosti výsledků t-testy pro nepárová data

- Pracujeme se dvěma náhodnými výběry z rozdělení N (μ_1, σ_1^2) a N (μ_2, σ_2^2)

- Nulová hypotéza se formuluje
  Ho: $\mu_1 = \mu_2$ proti alternativní H_A: $\mu_1 \neq \mu_2$ - oboustranná varianta testu
  Ho: $\mu_1 = \mu_2$; H_A: $\mu_1 < \mu_2$ $(\mu_1 > \mu_2)$ -jednostranná varianta testu

- Před provedením dvouvýběrového T-testu (Studentova) je třeba ověřit, zdali-se rozptyly obou výběrů statisticky významně liší či ne.

- Proto je potřeba nejprve provést test na shodu rozptylu dvou výběrů – F-test (Fisher-Snedeckorův test). V závislosti na provedení F-testu volím vhodný test shodnosti výsledků (T-test):

  a) test shodnosti výsledků při rovnosti rozptylů $\sigma_1^2 = \sigma_2^2$
  b) test shodnosti výsledků při nerovnosti rozptylů $\sigma_1^2 \neq \sigma_2^2$

F-test (test shody dvou rozptylů)

- Pro dva soubory dat, za předpokladu normality obou výběrů.
- Klasický *F*-test umožňuje ověření hypotézy $H_0 : \sigma_1^2 = \sigma_2^2$ proti alternativní hypotéze $H_A: \sigma_1^2 \neq \sigma_2^2$ (oboustranný test)
- Testovací kritérium má tvar:

$$F = \frac{s_1^2}{s_2^2}$$

klademe tak, že $s_1^2 > s_2^2$, tedy $F > 1$ (pracujeme s odhady rozptylů),
v Excelu při použití F-testu z Analýzy dat se jako první oblast dat zadává soubor s větším rozptylem (raději nepoužívat).

- Kritickou hodnotu *F*-testu stanovujeme pro počet stupňů volnosti $f_1 = (n_1-1)$ a $f_2 = (n_2-1)$ a příslušnou hladinu významnosti α – $F_k(1-\alpha/2;f_1,f_2)$
v Excelu stanovíme pomocí funkce F.INV.RT($\alpha/2$; f_1,f_2),
F.INV(1-$\alpha/2$; f_1,f_2)
FINV($\alpha/2$; f_1,f_2) - starší verze MS Office
pozor na pořadí f_1 a f_2 (f_1 - soubor s větším rozptylem, f_2 - soubor s menším rozptylem)

- Srovnáme velikost testovacího kritéria a kritické hodnoty –
platí-li, že $F \leq F_k(1-\alpha/2;f_1,f_2)$, je hypotéza H_0 o shodě rozptylů přijata.
platí-li, že $F > F_k(1-\alpha/2;f_1,f_2)$, je hypotéza H_0 o shodě rozptylů zamítnuta.

T-test při rovnosti rozptylů

- **a) pro $\sigma_1^2 = \sigma_2^2$**
  Ho: $\overline{x}_1 = \overline{x}_2$ má testovací kritérium tvar

$$t = \frac{|x_1 - x_2|}{\sqrt{\frac{(n_1 - 1) * s_1^2 + (n_2 - 1) * s_2^2}{n_1 + n_2 - 2} * \frac{n_1 + n_2}{n_1 * n_2}}}$$

  Pracujeme s výběrovými rozptyly
- Tato testovaná statistika má Studentovo rozdělení s počtem stupňů volnosti $f = n_1 + n_2 - 2$
- Stanovení kritické hodnoty pro Tk $(1-\alpha/2;\ n_1 + n_2 - 2)$ v případě oboustranného testu - v Excelu fce T.INV $(1-\alpha/2;\ n_1 + n_2 - 2)$
- Stanovení kritické hodnoty pro Tk $(1-\alpha;\ n_1 + n_2 - 2)$ v případě jednostranného testu – v Excelu T.INV $(1-\alpha;\ n_1 + n_2 - 2)$
- Platí-li, že $T \leq Tk$ je rozdíl obou průměrů statisticky nevýznamný a hypotéza H_0 se přijímá
- Platí-li, že $T > Tk$ je rozdíl obou průměrů za statisticky významný a hypotéza H_0 se zamítá

T-test při nerovnosti rozptylů

- **Pro $\sigma_1^2 \neq \sigma_2^2$**
  Ho: $\bar{x}_1=\bar{x}_2$ má testovací kritérium tvar

$$t=\frac{|\bar{x}_1-\bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$$

počet stupňů volnosti

$$\nu=\frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}}$$

Pracujeme s výběrovými rozptyly

- Stanovení kritické hodnoty
  Tk ($1-\alpha/2$; ν) – v Excelu T.INV ($1-\alpha/2$; ν) v případě oboustranného testu
  Tk ($1-\alpha$; ν) - v Excelu T.INV ($1-\alpha$; ν) v případě jednostranného testu
- Platí-li, že T ≤ Tk je rozdíl obou průměrů statisticky nevýznamný a hypotéza H_0 se přijímá
- Platí-li, že T > Tk je rozdíl obou průměrů za statisticky významný a hypotéza H_0 se zamítá
- Je-li skupin hodnot (tj. náhodných výběrů) víc než dva, je správnější provést simultánní porovnání pomocí **analýzy rozptylu** (Anova) než opakovanými t-testy po dvojicích.

Párový test shodnosti výsledků

- t-test pro dva závislé resp. korelované výběry
- Opakovaná pozorování znaku (veličiny) (ozn: Y,Z) u stejných statistických jednotek.
- V párovém t-testu ověřujeme, zda rozdíly mezi jednotlivými páry jsou minimální a tedy rozdíl středních hodnot rozdělení pro veličiny Y a Z je roven nule (oboustranný test).

- Vyjdeme ze Studentova testu správnosti výsledků (přednáška 8)
  H_0: $\overline{X} = \mu$
  výpočet testovacího kritéria
  Utvoříme veličinu X, kde $x_i = y_i - z_i$,
  pokud rozdíly mezi jednotlivými páry jsou minimální, pak rozdíl středních hodnot veličiny Y a Z = 0 => označíme $\mu = 0$
  pak můžeme párový test převést na případ jednovýběrového t-testu, přičemž testujeme, veličinu X – rozdíly mezi jednotlivými páry =>
  testovací kritérium je tedy dáno vztahem:

Párový t-test shodnosti výsledků

- Stanovím kritickou hodnotu Studentova rozdělení pro n-1 stupnů volnosti a $1-\alpha/2$ v případě oboustranného testu Tk ($1-\alpha/2$; n-1) a v případě jednostranného testu Tk ($1-\alpha$; n-1).
- V excelu tedy jako T.INV ($1-\alpha/2$; n-1) u oboustranného testu nebo T.INV ($1-\alpha$; n-1) u jednostranného testu
- Srovnám hodnotu testovacího kritéria s kritickou hodnotou (hodnotu testovacího kritéria uvažuji jako velikost – beru v absolutní hodnotě).
- pokud

  $| t | \leq Tk$ → přijímám H_0

  $| t | > Tk$ → zamítám H_0

testy v Analýze dat

- **Dvouvýběrový F-test pro rozptyl** (z Analýzy dat raději nepoužívat; záleží na pořadí souborů, jako první vkládám ten s větším rozptylem)
- **Dvouvýběrový t-test s rovností rozptylů** (volí se v závislosti na výsledku F-testu; používat tento test z Analýzy dat)
- **Dvouvýběrový t-test s nerovností rozptylů** (volí se v závislosti na výsledku F-testu; používat tento test z Analýzy dat)
- **Dvouvýběrový párový t-test na střední hodnotu** (umět i podle vzorce i z Analýzy dat)

Výsledkem těchto testů v Analýze dat je tabulka, kde najdu všechny potřebné hodnoty

	Soubor 1	*Soubor 2*
Stř. hodnota	332.8127	334.9451
Rozptyl	170.0737	146.9383
Pozorování	28	17
Společný rozptyl	161.4652	
Hyp. rozdíl stř. hodnot	0	
Rozdíl	**43**	
t Stat	**-0.5458**	
P(T<=t) (1)	0.294013	
t krit (1)	**1.681071**	
P(T<=t) (2)	0.588026	
t krit (2)	**2.016692**	

Např. výsledná tabulka t-testu s rovností rozptylu

Aritmetický průměr a rozptyl obou souborů

Rozsah souborů

testovaný rozdíl mezi střední hodnotou souborů

počet stupňů volnosti

testovací kritérium, beru ji jako kladné číslo, nebo jako 1. volím soubor s větším průměrem

kritická hodnota studentova rozdělení - pro jednostrannou variantu testu - hodnota kvantilu pro p=0.95 a 43 stupňů volnosti

kritická hodnota studentova rozdělení - pro oboustrannou variantu testu - hodnota kvantilu pro p=0.975 a 43 stupňů volnosti

Chí-kvadrát test dobré shody ($\chi 2$ – test)

- Pro testování shody rozdělení pravděpodobnosti náhodného výběru s teoretickým, očekávaným rozdělením pravděpodobností.

- Tedy ptáme-li se na otázku?
  Má soubor dat normální rozdělení? Tam kde požadována normalita dat
  Má soubor dat logaritmicko-normální rozdělení?
  Má soubor dat rovnoměrné rozdělení?
  ⋮

- Provedeme n nezávislých opakování pokusu. Výsledky rozdělíme do tříd a sledujeme vztah mezi intervalovým rozdělením četností $n_1, \ldots, n_k$ (kde k představuje označení třídy) souboru dat a očekávaným, teoretickým rozdělením, podle něhož očekáváme, že se soubor dat má chovat.

- Podmínky užití testu: žádný interval s nulovou četností, < 20% intervalů s četností menší než 5; možnost sloučit intervaly

- V případě testování zda má soubor dat normální rozdělení - **Testování „normality" souboru dat je značně nespolehlivé pokud je počet měření malý** (n méně 100). Proto je vždy vhodné ověřit si rozložení dat souboru **vizuální kontrolou** – kontrolou histogramu rozdělení četností.

Chí-kvadrát test dobré shody ($\chi 2$ – test)

- Chí kvadrát test je založen na tom, že náhodnou veličinu s určitým rozdělením pravděpodobností lze transformovat na veličinu mající přibližně rozdělení chí kvadrát

- Ho : $p_{e1} = p_{o1}, \ . \ . \ . \ , \ p_{ek} = p_{ok}$ pro všechny intervaly
  H_A : $p_{ej} \neq p_{oj}$ alespoň pro některý interval

- Výpočet testovacího kritéria

$$\chi 2 = \sum_{j=1}^{k} \frac{(n_{ej} - n_{oj})^2}{n_{oj}}$$

n_{ej} - experimentální četnosti v j-té třídě
n_{oj} – očekávané četnosti v j-té třídě

- Pokud má testovaná náhodná veličina předpokládané rozdělení, má náhodná veličina χ^2 přibližně rozdělení chí-kvadrát

Chí-kvadrát test dobré shody ($\chi 2$ – test)

Stanovení kritické hodnoty

kritická hodnota $\chi 2_\kappa$ se stanovuje pro jako příslušný kvantil α chí-kvadrát rozdělení pravděpodobností pro 1-α a ν = k-s-1 stupňů volnosti,

kde k je počet tříd (intervalů) náhodného výběru a s je počet parametrů daného rozdělení

v excelu $\chi 2_\kappa$ (1-α; k-s-1) stanovujeme pomocí funkce

CHISQ.INV(0.95; k-s-1) *nové MS Office*

CHISQ.INV.RT(0.05; k-s-1) *nové MS Office*

CHIINV (0.05; k-s-1) *staré MS Office*

pro rovnoměrné rozdělení - $\chi 2_\kappa$ (1-α; k-0-1)

pro normální rozdělení N (μ,σ^2) - $\chi 2_\kappa$ (1-α; k-2-1)

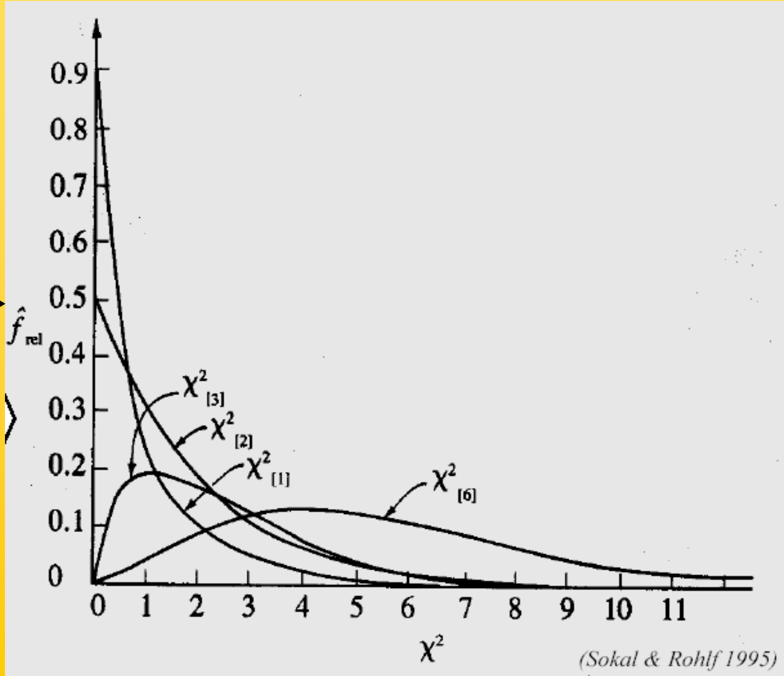
pro binomické rozdělení Bi (n, p) - $\chi 2_\kappa$ (1-α; k-2-1)

Pearsonovo ($\chi 2$) rozdělení

funkce, s intervalem hodnot <0,+∞)

má 1 parametr ν - stupně volnosti

hustota pravděpodobnosti pro $\chi 2$ rozdělení s 1, 2, 3 a 6 stupni volnosti.



Chí-kvadrát test dobré shody - příklad

Chceme ověřit, zda je hrací kostka pravidelná. Hodíme kostkou 120krát a sledujeme četnosti jednotlivých hodnot. Pracujte při hladině významnosti 5%.

Při pravidelné kostce je pravděpodobnost každého čísla 1/6, tedy všechny hodnoty od 1 do 6 mají očekávanou četnost 20.

Ho: hrací kostka je pravidelná, $n_{ei} = n_{oi}$

Následující tabulka uvádí skutečné (experimentální) n_e, očekávané četnosti n_o a výpočet testovacího kritéria.

hod	n_e	p_o	n_o	$(n_e-n_o)^2/n_o$
1	16	1/6	20	0,80
2	22	1/6	20	0,20
3	15	1/6	20	1,25
4	28	1/6	20	3,20
5	21	1/6	20	0,05
6	18	1/6	20	0,20
suma	120	1	120	5,70

kritická hodnota se stanovuje pro 1-α; k-s-1 stupňů volnosti, kde počet parametrů rovnoměrného rozdělení je 0 a tedy $\chi 2_{\kappa}$ (0,95; 5) = 11,07

testovací kritérium	5,70
kritická hodnota	11,07

5,7 < 11,07 => Ho platí

Děkuji za pozornost