Základy zpracování geologických dat

testování statistických hypotéz

R. Čopjaková

Jednovýběrový Studentův t-test o střední hodnotě

Testování přítomnosti soustavné chyby - Test správnosti výsledků

- slouží k porovnání střední hodnoty μ s konstantou (H_0: $\mu = \mu_0$)
- Pracujeme s ***jedním výběrovým souborem***
- Aritmetický průměr výsledků série měření (výběrového souboru) je *správný*, pokud jeho rozdíl od skutečné hodnoty μ s určitou pravděpodobností (na zvolené hladině významnosti α) není statisticky významný.
- Skutečnou hodnotu μ obvykle neznáme a tedy ji nahrazujeme konvenčně správnou hodnotou (tzv. "analytické standardy",) nebo analýzou vzorku se známou koncentrací stanovované složky.
- K testování používáme směrodatnou odchylku - Studentův t-test správnosti výsledků (jeho analog pro malý počet měření ($n < 10$) je Lordův test)

Jednovýběrový Studentův t-test o střední hodnotě

- Formulujeme nulovou hypotézu

  oboustranný test $H_0: x_1 = x_0$, $H_A: x_1 \neq x_0$

  jednostranný test $H_0: x_1 \leq x_0$, $H_A: x_1 > x_0$

- Spočteme testovací kritérium

$$t = \frac{|\bar{x} - \mu|}{\sqrt{\frac{s^2}{n}}} = \frac{|\bar{x} - \mu|}{s}\sqrt{n}$$

- Testovací kritérium má Studentovo rozdělení s stupni volnosti $\nu = n-1$
- Kritickou hodnotu stanovíme jako příslušný kvantil Studentova rozdělení pro n-1 stupnů volnosti

  pro oboustrannou variantu testu $T_k(1-\alpha/2;n-1)$

  pro jednostrannou variantu testu $T_k(1-\alpha;n-1)$

- t srovnám s T_k: pokud $t \leq T_k$, pak Ho – přijímám, rozdíl $|\bar{x} - \mu|$ je způsoben pouze náhodnými chybami a zjištěný výsledek je správný. V opačném případě je výsledek zatížen soustavnou chybou.

Jednovýběrový Studentův t-test o střední hodnotě

Reálný příklad

Proběhlo testování analytických laboratoří – EMP, LA-ICP MS. Máme chemicky homogenní sklo s deklarovaným chemickým složením, v laboratoři provedeme 20 analýz na různých místech tohoto skla a spočteme průměrné koncentrace jednotlivých oxidů.

Deklarovaný obsah Al_2O_3 ve skle je **13,52 hm.%**

Výsledky laboratoře poskytly **průměrný obsah Al_2O_3 ve skle 13,31 hm.% a Sx 0,12**

Otázka je: liší se tato hodnota statisticky významně od hodnoty deklarované? Pracuje naše laboratoř dobře? Pracujeme při **hladině významnosti** $\alpha = 0,05$.

oboustranný test:

H_0 = naměřený obsah Al_2O_3 se významně neliší od deklarovaného obsahu; $\overline{X} = \mu$

H_A = naměřený obsah se významně liší od deklarovaného; $\overline{X} \neq \mu$

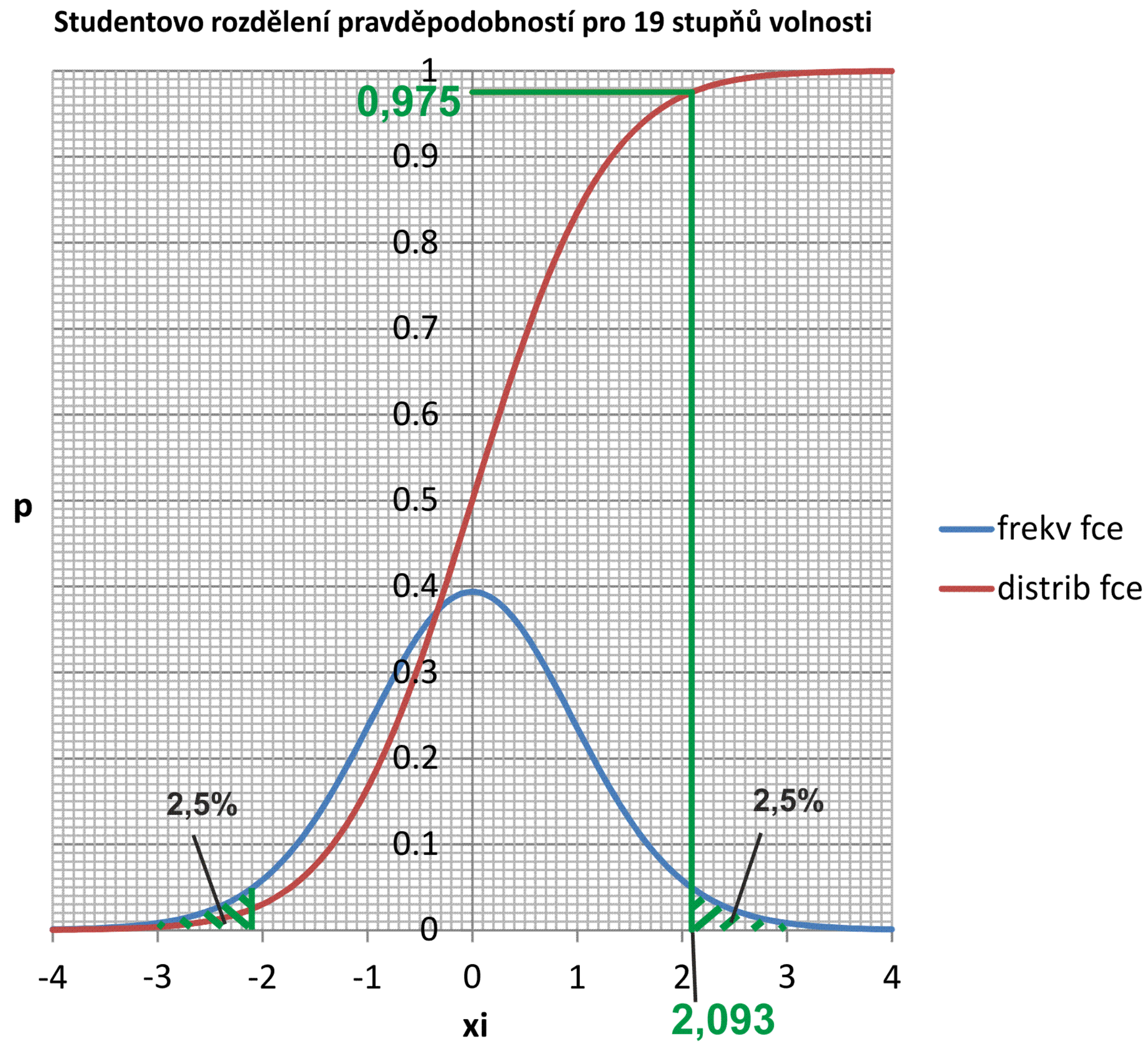
Spočteme testovací kritérium $t = \dfrac{|\bar{x} - \mu|}{\sqrt{\dfrac{s^2}{n}}}$ = 7,83

Testovací kritérium má Studentovo rozdělení se stupni volnosti $\nu = n-1$

Kritická hodnota - stanovíme jako příslušný kvantil Studentova rozdělení pro n-1 stupnů volnosti $T_k(1-\alpha/2;n-1) = t_k(0,975;19) = 2,09$

t srovnám s T_k: pokud 7,83 > 2,09; pak Ho - nulovou hypotézu zamítám

Výsledek je zatížen soustavnou chybou – koncentrace Al_2O_3 stanovené v laboratoři nejsou správné.



kritickou hodnotu v případě oboustranného testu
stanovím jako hodnotu kvantilu $(1-\alpha/2)$ studentova rozdělení pro příslušný stupeň volnosti
pro hladinu významnosti 0,05 - tedy kvantil $_{0,975}$
Tk (kritická hodnota) hodnota kvantilu $_{0,975}$ pro $\nu=19$ je 2,093

Kritickou hodnotu zjistím a) ze statistických tabulek

pro hladinu významnosti 0,05 – tedy hodnotu kvantilu (1-0,05/2)
a počet stupňů volnosti 19 (rozsah souboru = 20 měření)

Kvantily $t_{1-\alpha/2}$ Studentova t rozdělení pro dané stupně volnosti (ν = n-1)

St. volnosti	0,80	0,90	0,95	0,975	0,9875	0,995
1	1,376	3,078	6,314	12,706	25,452	63,657
2	1,061	1,886	2,920	4,303	6,205	9,925
3	0,978	1,638	2,353	3,182	4,176	5,841
4	0,941	1,533	2,132	2,776	3,495	4,604
5	0,920	1,476	2,015	2,571	3,163	4,032
6	0,906	1,440	1,943	2,447	2,969	3,707
7	0,896	1,415	1,895	2,365	2,841	3,499
8	0,889	1,397	1,860	2,306	2,752	3,355
9	0,883	1,383	1,833	2,262	2,685	3,250
10	0,879	1,372	1,812	2,228	2,634	3,169
11	0,876	1,363	1,796	2,201	2,593	3,106
12	0,873	1,356	1,782	2,179	2,560	3,055
13	0,870	1,350	1,771	2,160	2,533	3,012
14	0,868	1,345	1,761	2,145	2,510	2,977
15	0,866	1,341	1,753	2,131	2,490	2,947
16	0,865	1,337	1,746	2,120	2,473	2,921
17	0,863	1,333	1,740	2,110	2,458	2,898
18	0,862	1,330	1,734	2,101	2,445	2,878
19	0,861	1,328	1,729	2,093	2,433	2,861
20	0,860	1,325	1,725	2,086	2,423	2,845
∞	0,8416	1,2816	1,6448	1,9600	2,2414	2,5758

Kritickou hodnotu zjistím b) v Excelu

Oboustranný test

Novější verze MS Office – více typů funkce TINV

- T.INV.2T (0.05;19) = 2,093 - zadám hladinu významnosti s níž testuji a počet stupňů volnosti
- T.INV (0.975;19) = 2,093 – stanovím hodnotu kvantilu 0,975 (tedy 1-α/2) pro daný počet stupňů volnosti

Starší verze MS Office –jen jeden typ funkce TINV

- TINV (0.05;19) = 2,093 - zadám hladinu významnosti s níž testuji a počet stupňů volnosti

Testování odlehlých hodnot

- Je některá hodnota souboru odlehlá? (Mám ji ze souboru vyřadit a nepracovat s ní při výpočtu dalších parametrů). Např. přítomnost nahodilé chyby v analýzách, nebo přítomnost prvku ve výběrovém souboru, který nepochází ze studovaného základního souboru.

- Pro použití v analytické praxi k vyloučení odlehlých výsledků za předpokladu normality výběru je nejvhodnější **Grubbsův test** (parametrický)

- Dále se používá **Dean-Dixonův test** (neparametrický) - univerzální, nejen pro výběry s normálním rozdělením pravděpodobností, nebo neznám-li charakter rozdělení

Grubbsův test

- Při tomto testu se výsledky seřadí podle velikosti tak, že $x_1 < x_2 \ldots < x_n$, testujeme nejmenší i nevětší hodnotu
- Stanovení nulové hypotézy – Ho: hodnota x_1 není odlehlá
  Ho: hodnota x_n není odlehlá
- Výpočet testovacího kritéria:
  pro dolní odlehlou hodnotu

$$T_1 = \frac{\bar{x} - x_1}{S_n}$$

  pro horní odlehlou hodnotu

$$T_n = \frac{x_n - \bar{x}}{S_n}$$

kde Sn je definováno

$$S_n = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}}$$

n je počet měření (do četnosti 100)

Grubbsův test

- Stanovení kritické hodnoty Grubbsova rozdělení ze statistických tabulek Tk (α;n)
- Hodnota T_n a T_1 se porovná s kritickou hodnotou Grubbsova rozdělení T_k (α;n)

	T_α	
n	α = 0,05	α = 0,01
3	1,412	1,416
4	1,689	1,723
5	1,869	1,955
6	1,996	2,130
7	2,093	2,265
8	2,172	2,374
9	2,237	2,464
10	2,294	2,540

- Je-li T_1 nebo $T_n \leq$ Tk, přijmeme nulovou hypotézu Ho, hodnota není odlehlá
- Je-li T_1 nebo T_n >Tk, zamítneme nulovou hypotézu Ho, testovanou hodnotu považujeme za odlehlou a hodnotu vyloučíme ze souboru dat.

Dean-Dixonův test

- Při tomto testu se výsledky seřadí podle velikosti tak, že $x_1 < x_2 \ldots < x_n$, testujeme nejmenší i nevětší hodnotu
- Stanovení nulové hypotézy – Ho: hodnota x_1 není odlehlá
  Ho: hodnota x_n není odlehlá
- Výpočet testovacího kritéria:
  pro dolní odlehlou hodnotu pro horní odlehlou hodnotu

$$Q_1 = \frac{x_2 - x_1}{x_n - x_1} = \frac{x_2 - x_1}{R}$$

$$Q_n = \frac{x_n - x_{n-1}}{x_n - x_1} = \frac{x_n - x_{n-1}}{R}$$

kde R je variační rozpětí souboru dat

- Použití testu do četnosti souboru $n \leq 30$

Dean-Dixonův test

- Stanovení kritické hodnoty Dean-Dixonova rozdělení ze statistických tabulek Qk (α;n)

	Q_α	
n	α = 0,05	α = 0,01
3	0,941	0,988
4	0,765	0,889
5	0,642	0,760
6	0,560	0,698
7	0,507	0,637
8	0,468	0,590
9	0,437	0,555
10	0,412	0,527

- Hodnota Q_n a Q_1 se porovná s kritickou hodnotou Dean-Dixonova rozdělení Q_k (α;n)
- Je-li Q_1 nebo $Q_n \leq Qk$, přijmeme nulovou hypotézu Ho, hodnota není odlehlá
- Je-li Q_1 nebo $Q_n > Qk$, zamítneme nulovou hypotézu Ho, testovanou hodnotu považujeme za odlehlou a hodnotu vyloučíme ze souboru dat.

Příklad testování odlehlých hodnot; Dean-Dixonův test

- Máme soubor 10 měření. Ověřte, zda je některá hodnota odlehlá:
  2,1 2,9 3,1 3,3 3,3 3,4 3,5 3,5 3,6 3,9
- Ho – hodnota 2,1 není odlehlá
- Spočtení testovacího kritéria

$$Q_1 = \frac{x_2 - x_1}{x_n - x_1} = \frac{x_2 - x_1}{R}$$

- Q1 = (2,9-2,1)/(3,9-2,1) = 0,8/1,8 = 0,444
- Qk (α;n) = Qk (0,05;10) = 0,412
- 0,444 > 0,412 tedy Q1 > Qk, nulovou hypotézu zamítáme, hodnotu považujeme za odlehlou a ze souboru ji vyloučíme

- testujeme dále pro nový soubor dat po odstranění odlehlé hodnoty
- 2,9 3,1 3,3 3,3 3,4 3,5 3,5 3,6 3,9
- Ho – hodnota 2,9 není odlehlá
- Spočtení testovacího kritéria

- Q1 = (3,1-2,9)/(3,9-2,9) = 0,2/1 = 0,2
- Qk (α;n) = Qk (0,05;9) = 0,437
- 0,2 ≤ 0,437 tedy Q1 ≤ Qk, nulovou hypotézu přijmeme, hodnotu nepovažujeme za odlehlou

- testujeme dále zda je v souboru dat horní odlehlá hodnota
- 2,9 3,1 3,3 3,3 3,4 3,5 3,5 3,6 3,9
- Ho – hodnota 3,9 není odlehlá
- Spočtení testovacího kritéria

$$Q_n = \frac{x_n - x_{n-1}}{x_n - x_1} = \frac{x_n - x_{n-1}}{R}$$

- Qn = (3,9-3,6)/(3,9-2,9) = 0,3/1 = 0,3
- Qk (α;n) = Qk (0,05;9) = 0,437
- 0,3 ≤ 0,437 tedy Qn ≤ Qk, nulovou hypotézu přijmeme, hodnotu nepovažujeme za odlehlou