

Cvičení 7:

Testování statistických hypotéz

Aplikace F-testu a t-testu pro nezávisle proměnné

Statistické metody a zpracování dat 1 (podzim 2015)

Brno 19., 24. a 25.11. 2015

Ambrožová Klára

Úvod do problematiky

- Pepík a Toník se rozhodli porovnat hustotu sněhu v Jeseníkách a Beskydech.



Úvod do problematiky

- Jednoho jasného lednového dne se tedy Pepík vydal Praděd a Toník na Lysou horu. Oba šli po západním svahu a měřili přibližně každých 100 m výškových metrů, přičemž s prvním měřením začali v deset hodin ráno.



Úvod do problematiky

- Ve 12:00 dorazila do Hrubého Jeseníku teplá fronta, která znemožnila Pepíkovi provést všechna měření. Do Beskyd dorazila až o 2 hodiny později, takže Toník všechna měření dokončil.



Úvod do problematiky

- Z Beskyd teď mají 10 měření a z Jeseníku jen 8. Co s tím? Bylo celé měření k ničemu?



Trocha teorie

- Testovat je možno odlišnost i u souborů, které nemají stejný počet prvků
- Aplikujeme F-test a poté t-test

Data: Tabulka1* (10s krát 10ř)

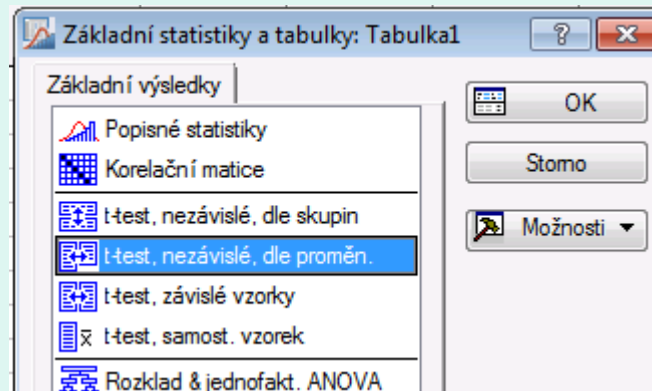
	1	2	
	Lysa hora	Praded	
1	170	230	
2	190	210	
3	110	180	
4	120	140	
5	70	150	
6	30	110	
7	50	80	
8	30		
9	10		
10	20		



Trocha teorie

Jak postupujeme?

1. Určíme, zda máme normální rozdělení (známe z přechozích cvičení)
2. Stanovíme si hladinu významnosti
 - Obvykle se stanovuje 5 % nebo 1 %, popř. 10 %
 - V případě, že stanovíme hladinu významnosti 1 %, pak budeme p-hodnotu porovnávat s 0,01
3. Určíme shodu rozptylů pomocí F-testu



F-test

- umožní nám určit, zda se liší rozptyly testovaných souborů
- nulová hypotéza – rozptyly obou souborů se neliší

$$H_0: \hat{\sigma}_1^2 = \hat{\sigma}_2^2$$

- alternativní hypotéza – liší se dle toho, co chceme testovat, ale v tomto případě chceme použít oboustranný test!



F-test

- alternativní hypotéza – liší se dle toho, co chceme testovat, ale v tomto případě chceme použít oboustranný test!

$$H_1: \hat{\sigma}_1^2 \neq \hat{\sigma}_2^2 \quad (\text{pro oboustranný test})$$

$$H_1: \hat{\sigma}_1^2 < \hat{\sigma}_2^2 \quad \text{nebo} \quad \hat{\sigma}_1^2 > \hat{\sigma}_2^2 \quad (\text{pro jednostranný test})$$

- výpočet: podíl odhadů rozptylů základních souborů

$$F = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$$

$$\hat{\sigma}_1^2 = \frac{n_1}{n_1 - 1} \cdot s_1^2 \quad \text{a} \quad \hat{\sigma}_2^2 = \frac{n_2}{n_2 - 1} \cdot s_2^2$$



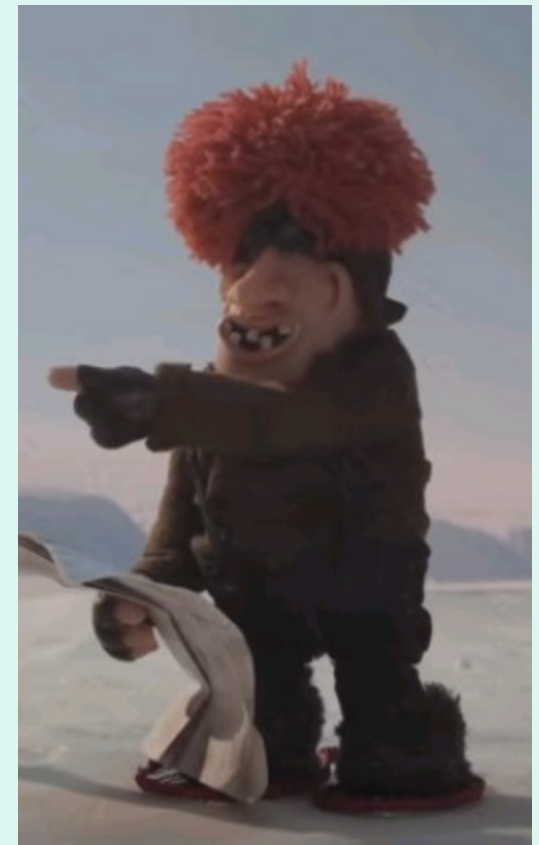
F-test

- Vypočtenou hodnotu porovnáme s kritickou hodnotou (lze určit ze statistických tabulek na základě stupňů volnosti)

NEBO

nám software spočte p-hodnotu, kterou porovnáme s námi zvolenou hladinou významnosti

		T-test pro nezávislé vzorky (Tabulka1)				
		Pozn.: Proměnné byly brány jako nezávislé vzorky				
Skup. 1 vs. skup. 2		Průměr skup. 1	Průměr skup. 2	Hodnota t	sv	p
Lysa_hora vs. Praded		80,00000	157,1429	-2,60073	15	0,020070
...						
...	Poč.plat. skup. 1	Poč.plat. skup. 2	Sm.odch. skup. 1	Sm.odch. skup. 2	F-poměr Rozptyly	p Rozptyly
	10	7	64,29101	53,45225	1,446667	0,673375



F-test

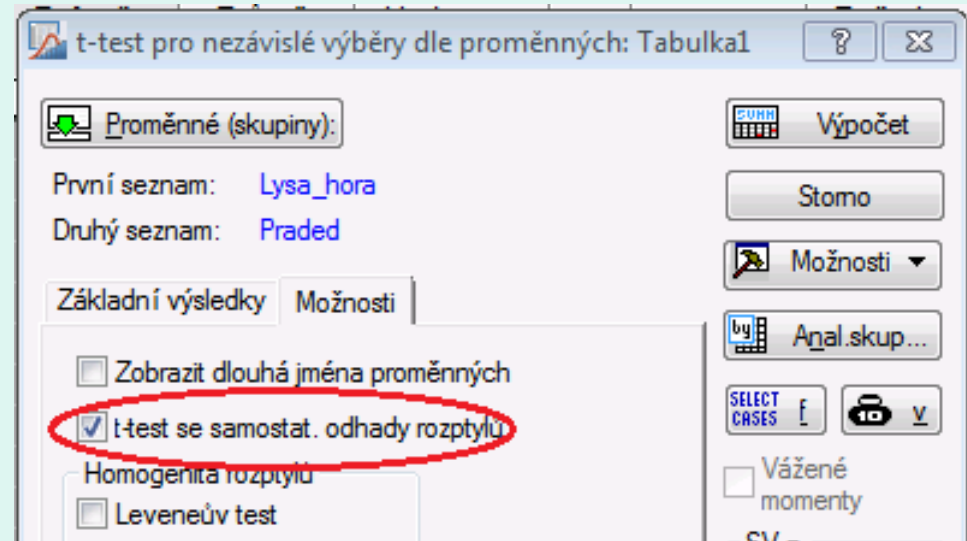


- Jak postupovat dál?

$p > 0.01 \rightarrow$ platí H_0 , hodnotíme t-test (ve stejné tabulce)

$p < 0.01 \rightarrow$ platí H_1 (chceme-li dále pokračovat ve výpočtu, tak je v softwaru STATISTICA nutno zaškrtnout volbu „Test se samostat. odhady rozptylů“ na kartě „Možnosti“)

4. t-test



t-test

- lze použít např. pro testování rozdílů dvou výběrových průměrů
- výpočet se liší podle toho, zda jsou či nejsou shodné rozptyly!
 - rozptyly jsou shodné

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

- rozptyly se liší

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}}}$$



t-test

- nulová hypotéza – výběrové průměry obou souborů se neliší

$$H_0: \mu = \mu_0$$

- alternativní hypotéza – liší se dle toho, co chceme testovat, ale v tomto případě chceme použít oboustranný test!

$$H_1: \mu \neq \mu_0 \quad (\text{pro oboustranný test})$$

$$H_1: \mu > \mu_0 \quad \text{nebo} \quad \mu < \mu_0 \quad (\text{pro jednostranný test})$$



t-test

- hodnocení t-testu:

$p > 0.01 \rightarrow$ platí H_0 , nebyl prokázán rozdíl mezi výběrovými průměry

$p < 0.01 \rightarrow$ platí H_1 , tedy bylo zjištěno, že mezi průměry souborů existuje statisticky významný rozdíl



T-test pro nezávislé vzorky (Tabulka1)					
Pozn.: Proměnné byly brány jako nezávislé vzorky					
Skup. 1 vs. skup. 2	Průměr skup. 1	Průměr skup. 2	Hodnota t	sv	p
Lysa_hora vs. Praded	80,00000	157,1429	-2,60073	15	0,020070

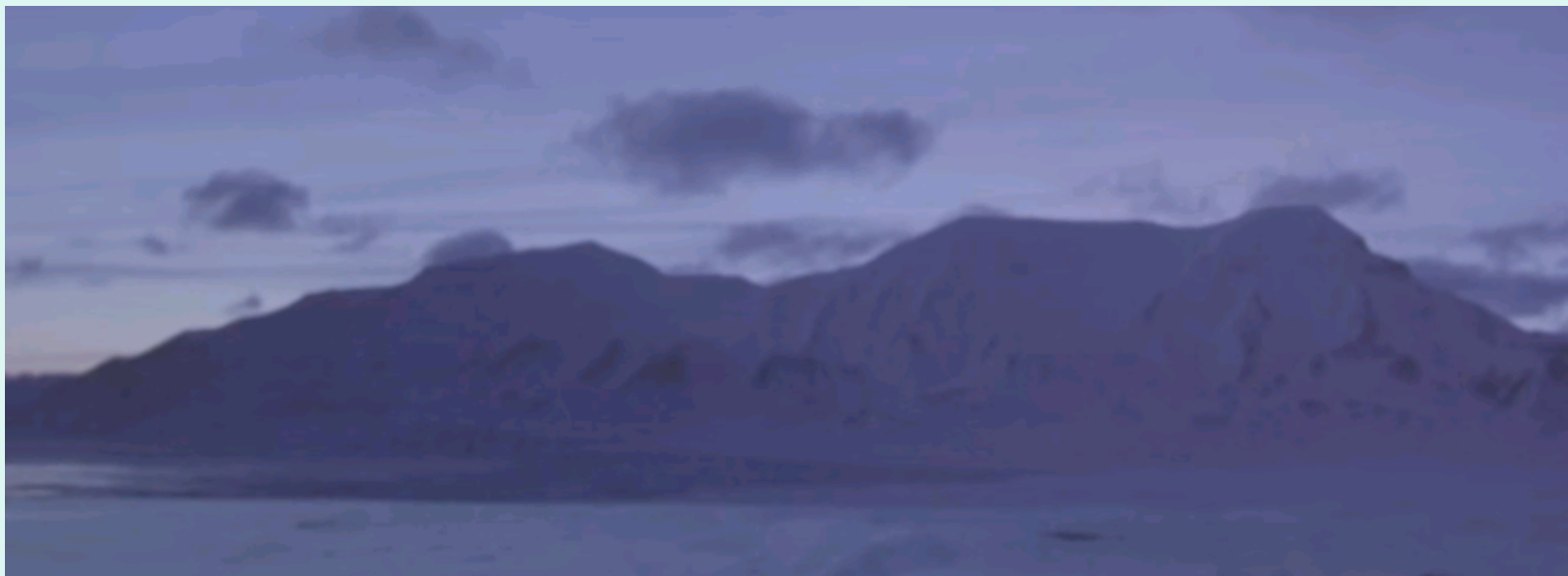
...

Poč.plat. skup. 1	Poč.plat. skup. 2	Sm.odch. skup. 1	Sm.odch. skup. 2	F-poměr Rozptyly	p Rozptyly
10	7	64,29101	53,45225	1,446667	0,673375

Co z toho vyplývá?

(aneb co napíšou Pepík a Toník do závěru...)

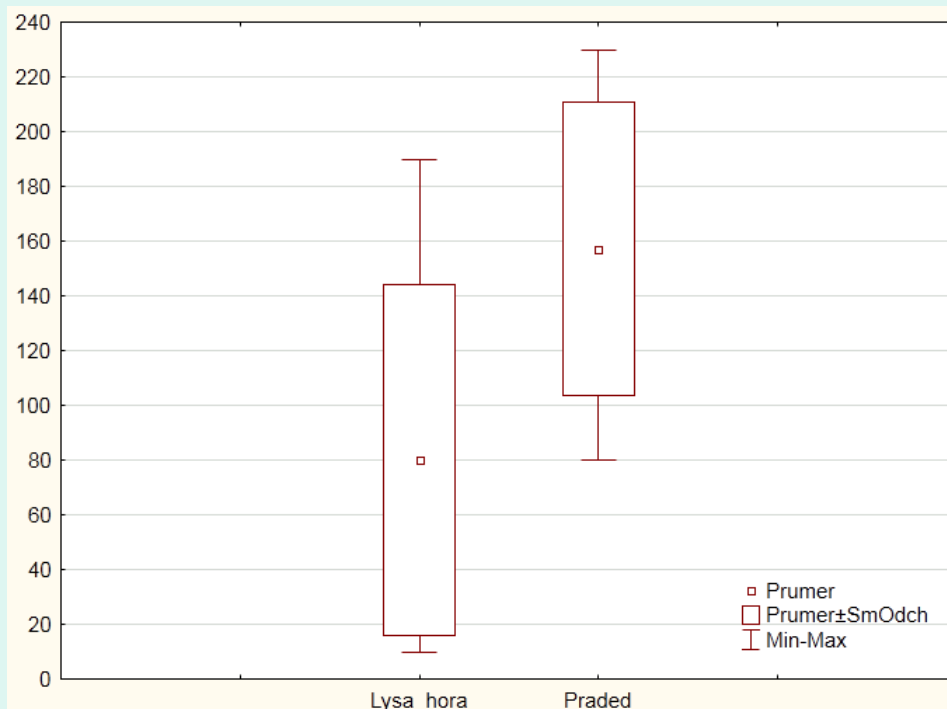
- Pepík a Toník si zvolili hladinu významnosti 1 %, no a vyšlo jim, že rozptyly hustot sněhu na Pradědu a Lysé hoře se statisticky významně neliší, a výběrové průměry také ne



Co z toho vyplývá?

(aneb co napíšou Pepík a Toník do závěru...)

- Toník se podíval na krabicový graf a konstatoval, že průměry obou souborů se liší o téměř 80 kg/m^3 a také minimální a maximální hodnota jsou zřetelně odlišné, a že se mu to nějak nezdá...



Co z toho vyplývá?

(aneb co napíšou Pepík a Toník do závěru...)

- Pepík po chvíli přemýšlení řekl, že i jemu ty soubory připadají dost odlišné, a že by mohli do závěru napsat, že měření byla pravděpodobně ovlivněna tím, že z vrcholových částí Jeseníků nebylo k dispozici žádné měření, a že by tudíž asi bylo vhodnější zvolit jinou hladinu významnosti (tedy 5 %) nebo provést pokus ještě alespoň jednou



Zdroje

- BRÁZDIL, Rudolf. Statistické metody v geografii :cvičení. 3. vyd. Brno: Vydavatelství Masarykovy univerzity, 1995. 177 s. ISBN 80-210-1260-9.
- BUDÍKOVÁ, Marie. Parametrické úlohy o dvou nezávislých náhodných výběrech z normálních rozložení (přednáška). Brno: Masarykova univerzita, 17.11.. 2015.
- DOBROVOLNÝ, Petr. Z1069 Statistické metody a zpracování dat: V. Testování statistických hypotéz (přednáška) Brno: Masarykova univerzita, 17.11.. 2015.
- Obrázky pochází ze seriálu Sval & Bard: <http://www.svalandbard.com/>



Dodatek 1 : Co je co ve výsledných tabulkách?

- **Nebyl prokázán statisticky významný rozdíl mezi rozptyly:**
(pozor, tabulka je rozpůlená)

2) Hodnota t: hodnota spočteného testovacího kritéria pro t-test,
p: p-hodnota příslušející t-testu (menší než 0,05, takže byl prokázán rozdíl mezi střední hodnotou souborů!)

Tohle jsou průměry souborů ze Strážnice a z Klementina

Sv: stupně volnosti, zde spočtené jako $(30+120)-2$

Počet členů v souboru ze Strážnice a z Klementina

Směrodatné odchyly obou souborů

T-test pro nezávislé vzorky (Tabulka1)					
Pozn.: Proměnné byly brány jako nezávislé vzorky					
Skup. 1 vs. skup. 2	Průměr skup. 1	Průměr skup. 2	Hodnota t	sv	p
Straznice vs. Klementinum2	8,991111	9,821667	-4,68277	148	0,000006

Poč.plat. skup. 1	Poč.plat. skup. 2	Sm.odch. skup. 1	Sm.odch. skup. 2	F-poměr Rozptyly	p Rozptyly
30	120	0,671983	0,910451	1,835717	0,060320

1) F-poměr: hodnota spočteného testovacího kritéria, p-rozptyly: p-hodnota příslušející tomuto F-testu (zde je větší než 0,05 → nebyl prokázán rozdíl mezi rozptyly)

Dodatek 1 : Co je co ve výsledných tabulkách?

- **Byl prokázán statisticky významný rozdíl mezi rozptyly :**
- zde je tabulka po spočtení znovu se zakliknutým „t-test se samost. Odhady rozptylů“(pozor, tabulka je rozpůlená)

Ve střední části tabulky přibyla část pro interpretaci, kde „t samost. Odh. Rozp.“ je hodnota testovacího kritéria t-testu, a „p oboustr.“ je příslušná p-hodnota

		T-test pro nezávislé vzorky (Tabulka1)							
		Pozn.: Proměnné byly brány jako nezávislé vzorky							
Skup. 1 vs. skup. 2	Průměr skup. 1	Průměr skup. 2	Hodnota t	sv	p	t samost. odh. rozp.	sv	p oboustr.	
Straznice vs. Klementinum 1	8,991111	9,397500	-2,10663	148	0,036837	-2,65712	65,16241	0,009901	

Poč. plat. skup. 1	Poč. plat. skup. 2	Sm. odch. skup. 1	Sm. odch. skup. 2	F-poměr Rozptyly	p Rozptyly
30	120	0,671983	1,000375	2,216201	0,015160

Dodatek 2: Proč se mohou soubory z cv7 lišit?

- a) U klimatologických dat hodně záleží na geografické poloze (zeměpisná šířka a délka, nadmořská výška) → je třeba se podívat, zda se poloha mého města (Aš, Bylnice...) neliší od polohy Klementina
- b) Data mohou pocházet z jiného období – zatímco období 1961–1990 bylo v ČR velmi teplé, tak třeba období 1830–1930 patřilo v ČR spíše k chladnějším obdobím (i když tohle je hodně zjednodušené) → jak moc se mi liší období z Klementina od období 1961–1990
- c) Soubor, který obsahuje více prvků, je reprezentativnější a data z Klementina jsou v tomto případě 4x delší než dat z druhého místa, je tedy možné, že v Klementinu bylo zachyceno více extrémních situací
- d) Určitě na něco přijdete sami...