# STATISTICA Téma 7. Testy na základě více než 2 výběrů

## 1) Test na homoskedasticitu

Nalezneme jej v několika submenu. Omezme se na submenu *Základní statistiky a tabulky* základního menu *Statistika*. V něm vybereme nabídku *Rozklad & jednofakt. ANOVA* - viz obr. 1. Po potvrzení výběru *OK* se otevře okno *Statistiky dle skupin* – viz obr. 2.

#### Obr. 1



#### Obr. 2



V něm vybereme kartu *Jednotlivé tabulky* a zadáme proměnné.

- závislé proměnná musí být číselná (např. počet prodaných kusů),
- grupovací proměnná musí být kategoriální (např. značka).

Bývá užitečné aktivovat tlačítko *Kódy pro grupovací proměnné*. Klikneme na ně a v nabídce, která se objeví, vybereme možnost *Vše* ( zobrazí se všechny kategorie grupovací proměnné, které máme naeditovány). Nastavení potvrdíme tlačítkem *OK*.

Nyní provedeme "formální" výpočet stiskem tlačítka **OK**.

Objeví se okno Statistiky dle skupin - výsledky – viz obr. 3.

0	br.	3

👬 STATISTICA Cz - adnire		_ 8 ×
Soubor Úpravy Zobrazit Vložit Formát Stati	stika Grafy <u>N</u> ástroje <u>D</u> ata <u>O</u> kno Nápo <u>v</u> ěda	
🗈 🛩 🖬 🎒 🕼 👗 🖬 🋍 💅 🗠 🤉	🗵 🏟 💱 Přídat do sešitu 🝷 Přídat do protokolu 🍷 🥔 🎀 🖕	
Arial T 10 T B I	및 특종 클 🌚 ▲·실·II· 🏹 🏢 🎎 🖓 🗰 🖬 🔂 🎝 🐖 Proměnné · Případy · .	
🛗 Data: adnire (10s krát 18ř)		
1 2 3 4 značka prodáno Var3 Var4	5 6 7 8 9 10	
1 A 25	Statistiky dle skupin - vysledky: adnire	
2 A 29	ZÁVISLÉ : 1 proměnnú: prodáno	
3 A 28		
4 A 27 5 Δ 30	CDUDOU - Legnočko (2) - A N B	
6 A 24	BALLER (J). ANK	
7 N 22		
8 N 19	Základ Popisné statistiky ANOVA & testy Post-hoc	
9 N 23		
11 N 30	Hatti Analyza rozprylu De Karegorg, normal, pravo, graty	
12 N 17	l esty homog. rozptylu 🖉 Kategoriz. polo-normál. p-grafy 🔊 Možnosti 🗸	
13 R 24	Leveneovy testy	
14 R 27	Brown-Forsytheovitesty	
16 R 20	Gra <u>í</u> průměrů vs. směr. odch.	
17 R 30		
18 R 28	Graf interakcí	
	Úroveň o pa	
	zvýraznění ,05 🖶 spolehl. průměru 133,00 💌 %	
PS 1* - Leve		
🔼 🖾 Statistiky dle skupin		
Pro nápovědu stiskněte F1	Ř18;S2 28 Filtr - Váhy: VYP	N CAP NUM REC
🔏 Start 🛛 🙋 🚰 🏈 🎲 🔹 🖹 🖉 Centrum	cz E-mail - doruč 🔞 Provádíme testy na rozpt 🕅 STATISTICA Cz - adnire	💐 CS 📲 9:56

V kartě *ANOVA & testy* je sekce testů homogenity rozptylů, tj. testy na homoskedasticitu. Zvolíme nabídku *Levenovy testy*. Prostřednictvím výsledků zjistíme, -viz obr. 4, zda můžeme pokračovat v testech na shodu středních hodnot prostřednictvím Analýzy rozptylu, nebo zda budeme muset provést Kruskal – Wallisův test.

Vyhodnocení provedeme tak, že v tabulce s výsledky prohlédneme číslo v posledním sloupci "p". Je-li p  $\leq 0,05$ , pak H<sub>0</sub> zamítáme - tj. heteroskedasticita je průkazná. V opačném případě heteroskedasticita průkazná není.

#### Obr. 4

📠 STATISTICA Cz - PS 1* - [Leveneův to	est homogenity rozpylů (Tabulka1)]
Soubor Úpr <u>a</u> vy Zobrazit Vložit Eormát	Statistika <u>G</u> rafy <u>N</u> ástroje <u>D</u> ata Pracovní <u>s</u> ešit <u>O</u> kno Nápo <u>v</u> ěda
0 🖆 🖬   🚑 🖪   👗 🛍 🛍 🝼   🗉	🗅 🖓 👫 🕼 Přidat do sešitu 👻 Přidat do protokolu 👻 🥔 📢 🖕
Arial T 10 T B	$I  \underline{\mathbf{U}}  \equiv \equiv \equiv \Box  \underline{\mathbf{\Delta}} \star \underline{\mathbf{\Delta}} \star \underline{\mathbf{\Delta}} \star [\Box \star \nabla_{\mathbf{k}}]  \exists \mathbf{k} \otimes \mathbf{k} \otimes \mathbf{k} \otimes \exists \mathbf{k} \otimes \mathbf$
🎹 Data: Tabulka1* (10s krát 20ř)	
ZNAĆKA PRODÁNO Var3	Var4 Var5 Var6 Var7 Var8 Var9 Var10
10 N 11 N II PS 1* - Leveneův test ho	mogenity rozpylů (Tabulka1)
12 N 🔄 PS 1 13 R 🖻 🐨 Základní statistiky a tabulky	Leveneův test homogenity rozpylů (Tabulka1)
14 R 🖻 🔄 Dialog výsledků rozpad	Uznač. efekty jsou význ. na hlad. p < 1/6000 SČ SV PČ SČ SV PČ F p
16 R	Proměnná efekt efekt efekt chyba chyba chyba
17 R	PRODANO [0,77776] 2 5,300009 58,00007 15 3,777776 1,420471 0,270352
18 R 19	
A Statistiky dle skupin	

#### 2) Testy na shodu středních hodnot Analýza rozptylu

Analýzu rozptylu používáme v případě, že výběry jsou nezávislé a pocházejí z normálních rozdělení, která vykazují homoskedasticitu. Předpokládejme tedy, že máme k disposici více než dva nezávislé výběry a prostřednictvím Levenova testu nebylo možno zamítnout shodu rozptylů.

V takové situaci můžeme využít výsledků *"formálního" výpočtu Analýzy rozptylu* -stačí myší stisknout tlačítko *Statistiky dle skupin* ve spodní části *výsledků Levenova testu* – viz obr. 4

Otevře se opět okno *Statistiky dle skupin*, karta *ANOVA & testy* –viz obr.3 a můžeme pokračovat v práci. Stiskneme tlačítko *Analýza rozptylu a* objeví se tabulka s výsledky –viz obr. 5.



🔼 STATISTICA Cz - PS 1* - [Analýza rozptylu (adnire)]	_ 8 ×
Soubor Úpr <u>a</u> vy Zobrazit Vložit Eormát Statistika Grafy <u>N</u> ástroje <u>D</u> ata Pracovní <u>s</u> ešit <u>O</u> kno Nápověda	
📄 🗅 😹 📓 🚭 💁 💺 ங 🎕 ổ 🗠 🖙 🙀 🍇 Přídat do sešitu 👻 Přídat do protokolu 👻 🥔 👯 🗸	
Arial ▼ 10 ▼ B Z U ≡ Ξ 🗃 Δ·Δ·Δ·Ι· 50 🗰 58 🖓 🛱 🖓 🔂 👌 ≈? 💱 Proměnné * Případy * .	
Data: adnire (10s krát 18ř)	
Image: State in the state	
The Construction of the American	
Image: Sector State     D161     00 202020202222     Dis.     1/4/act MONI     CAR	
	I tour
🙀 Statistica (z - P5 1*) 🚱 Centrum.cz E-mail - doruc 🔤 Provadine testy na rozpt [74] Statistica (z - P5 1*)	10:08

#### Hodnoty uvedené ve sloupcích mají následující význam:

SČ efekt – součet čtverců vlivem (grupovací proměnné) X,

SV efekt – počet stupňů volnosti součtu čtverců,

*PČ efekt* – rozptyl vlivem (grupovací proměnné) X -tzv. průměrný součet čtverců,

- SČ chyba součet čtverců reziduální,
- SV chyba počet stupňů volnosti reziduálního součtu čtverců,
- PČ chyba reziduální rozptyl,
- F hodnota testovacího kritéria,
- *p* minimální hladina významnosti, na níž lze H<sub>0</sub> zamítnout.

Je-li tedy  $p \leq 0,05$ , zamítá se hypotéza o shodě středních hodnot. V takovém případě je třeba zjistit, které střední hodnoty jsou rozdílné. Pokud hypotézu o shodě středních hodnot nelze zamítnout, další testy neděláme.

Testy mnohonásobného porovnávání najdeme na kartě *Post – hoc*. –viz obr. 6.

Všimněne si, že na kartě **Post – hoc** lze zvolit úroveň **p** pro zvýraznění. Jedná se o výhodnou optickou pomůcku. Jakmile bude rozdíl ve výsledné tabulce menší nebo roven nastavené hodnotě **p**, bude číslo ve výsledné tabulce červené. Červená barva pak signalizuje, že rozdíl mezi středními hodnotami příslušných faktorů je průkazný.

### Obr.6

<b>M</b> STATISTICA	Cz - [Data: adnire (10s l	krát 18ř)]	_ 8 ×
Soubor Up	r <u>a</u> vy <u>Z</u> obrazit Vložit <u>F</u> i	ormát Statistika <u>G</u> rafy <u>N</u> ástroje <u>D</u> ata <u>O</u> kno Nápo <u>v</u> ěda	_ 8 ×
🗅 🗲 🔛   🤅	3 🖪   👗 🖻 🛍 🝼	' 🗠 🖓 👫 🕼 Přidat do sešitu 🔹 Přidat do protokolu 🔹 🥔 💦 🖕	
Arial	▼ 10 ▼	B I U 토 吾 君 🗃 💁 ▲ • 실 • 团 • 📎 🏢 128 429 材 🔂 🔂 👌 斜 #? 🐉 Proměnné • Případy • •	
1	<b>2</b> 3	4 5 6 7 8 9 10	
značka 1 A	prodáno Var3	Var4 Var5 Var6 Var7 Var8 Var9 Var10	
2 4	25	Statistiky dle skupin - výsledky: adnire	
3 A	28		
4 A	27	ZÁVISLÉ : 1 proměnná: prodáno	
5 A	30		
6 A	24	GRUPOV. : 1-značka (3): A N R	
7 N	22		
9 N	23		
10 N	24	Základ Popisné statistiky ANOVA & testy Post-hoc	
11 N	30	Stomo Stomo	
12 N	17		
13 R	24	Itest LSD nebo plánovaná porovnání 🔊 🔊 Možnosti 🗸	
14 R	27	Etter Scheffeitutest	
16 P	20		
17 R	30	Mewman-Keulsův test & kritické rozpětí Hladina alfa krit rozsahů	
18 R	28	Duncanův test & kniická rozpětí	
		Iukeyűv HSD     Další post-hoo testy (Dumetteðu,	
		Bonferroniův, komplexní	
		obecné lineámí modely	
		Uroveň p na zvýraznění: 1.05	
			<b>_</b>
1			Þ
🜆 💁 Statisti	ky dle skupin		
Připravena		Ř18;S2 28 Filtr - Váhy: VYPN	CAP NUM REC
🛃 Start 🛛 🙆	🗹 🏉 😂 🔹 🖉	Centrum.cz E-mail - doruč 🛛 Provádíme testy na rozpt 🛛 🕅 STATISTICA Cz - [Data	CS 📲 10:20

Prostřednictvím tlačítka vybereme test- např. *LSD test* - a provedeme výpočet. Výsledná tabulka je uvedena na obr. 6



PS 3* - LSD test	t; proměnna	á: PRODÁN	) (Tabulka	1)			
🔄 PS 3* È 🔂 Základní statisl		LSD test; p Označ. roz	oroměnná: F díly jsou vý:	PRODÁNO ( znamné na	(Tabulka1) hlad. p < ,i	05000	
LSD te	ZNAĆKA	{1} M=27,167	{2} M=22,500	{3} M=27,333			
	A {1}	0.021075	0,021075	0,927911			
	R {3}	0,927911	0,017543	0,011,040			
		× ,					
<	LSD test;	proměnná: Př	RODANO (Tal	oulka1)			

Provedeme vyhodnocení zvoleného testu mnohonásobného porovnání a tím naše činnost končí. (Uvnitř tabulky jsou uvedeny hodnoty p důležité pro vyhodnocení průkaznosti rozdílu mezi příslušnými kategoriemi grupovací proměnné.)

#### <u>Kruskal – Wallisův test</u>

Kruskal – Wallisův test používáme v případě, že výběry jsou nezávislé a pocházejí z normálních rozdělení, která vykazují heteroskedasticitu. Předpokládejme tedy, že máme k disposici více než dva nezávislé výběry a prostřednictvím Levenova testu jsme zamítli hypotézu o shodě rozptylů. Kruskal – Wallisův test najdeme v menu *Statistika*, submenu *Neparametrická statistika*, nabídka *Porovnání více nezávislých vzorků* –viz obr.8

#### **Obr. 8**/

🎑 Neparametrické statistiky: cv3	?X
Zákl. výběr	СК
Tabulky 2 x 2 (Chí,V,Fí ^2, McNemar, Fisher)	Storno
x Pozorované versus očekávané; Chí^2	🔊 Možnosti 🖛
Korelace (Spearman, Kendallovo tau, gama)	Moznosu +
📰 Porovnání dvou nezávislých vzorků (proměnné)	
🗱 Porovnání více nezávislých vzorků (skupiny)	
Porovnání dvou závislých vzorků (proměnné)	
Porovnání více závislých vzorků (proměnné)	
11111 Cochranův Q test	🕞 Otevři Data
Pořadové popisné statistiky (medián, modus, )	SELECT f & Y

Potvrdíme – li výběr klávesou **OK**, otevře se okno pro **Kruskal – Wallisův a mediánový test** – viz obr. 9.

#### Obr. 9

🕼 Kruskal-Wallisova ANOVA a mediánový test: Ta	bulka' ? 💶 🔀
Zákl. výsledky	Výpočet
💀 Proměnné	Storno
Závislé proměnné: žádné Grupovací proměnná: žádné	🔊 Možnosti 🔻
Žie Kódy: žádné	SELECT <u>f</u> <u>∞</u> ⊻
Shrnutí: Kruskal-Wallis, ANOVA a mediánový test	
	p-úroveň
trabicový graf	pro

Natavíme proměnné, popřípadě kódy ( postup je podrobně uveden v kapitole o analýze rozptylu ) a provedeme výpočet. Objeví se výsledná sestava –viz obr. 10.

PS 4* - Medián	ový test, celk. medián =	27,0000; P	RODÁNO (	Tabulka1)		×				
Neparametrick	Mediánový test, celk. medián =27,0000; PR( Nezávislá (grupovací) proměnná : ZNAĆKA Chí-kvadrát = 1,870130 , sv = 2 , p =,3926									
Mediár	Závislá: PRODÁNO	Α	N	R	Celkem					
	<= Medián: pozorov.	3,000000	5,00000	3,000000	11,00000					
	očekáv.	3,666667	3,66667	3,666667						
	pozoč.	-0,6666667	1,33333	-0,6666667						
	> Medián: pozorov.	3,000000	1,00000	3,000000	7,00000					
	očekáv.	2,333333	2,33333	2,333333						
	pozoč.	0,6666667	-1,33333	0,666667						
	Celkem: oček.	6,000000	6,00000	6,000000	18,00000					
						-1				
	•				Þ					
<	IIII Mediánový test, celk. me	dián =27,0000;	PRODÁNO (	(Tabulka1)	ŀ	( ) ·				

#### Obr.10

V popředí je mediánový test. Musíme proto levé straně výstupní sestavy *přepnout Kruskal – Wallisův test* ( klepneme na tlačítko ). Výstupní sestava pro tento test je uvedena na obr. 11.





Ve výstupní sestavě je uvedena kritická hodnota testovacího kritéria ( Kruskal – Wallisův test H), celkový rozsah, tj. celkový počet hodnot (N), počet stupňů volnosti (počet výběrů – 1,N), minimální hladina významnosti, pro níž lze  $H_0$  zamítnout (p) a ve sloupcích pak je uveden počet hodnot statistického znaku v každém výběru a součet pořadí těchto hodnot. Není – li rozdíl mezi středními hodnotami průkazný, práci jsme skončili.

Je- li rozdíl mezi středními hodnotami průkazný, musíme provést testy mnohonásobného porovnání. Stačí klepnout myší na tlačítko Kruskal – Wallisova AN ve spodní části výstupní sestavy – viz Obr. 11. Otevře se opět okno pro Kruskal – Wallisův a mediánový test –viz obr. 9. Stiskem tlačítka Vícenás. porovnání průměrného pořadí pro vš. sk. test provedeme a výstupní sestavu –viz obr. 12 – vyhodnotíme. (Uvnitř tabulky jsou uvedeny hodnoty pdůležité pro vyhodnocení průkaznosti rozdílu mezi součty pořadí příslušných kategorií grupovací proměnné.)



#### Friedmanův test

Friedmanův test používáme v případě, že výběry jsou závislé. Předpokládejme tedy, že máme k disposici více než dva závislé výběry. Friedmanův test najdeme v menu Statistika, submenu Neparametrická statistika, nabídka Porovnání více závislých vzorků. Po potvrzení výběru tlačítkem OK se objeví okno Friedmanova testu - viz obr.14.

Pozor! Pro Friedmanův test jsou hodnoty každého znaku editovány jako zvláštní proměnná. – viz. obr.13.

#### **Obr. 13**

/∖ st	ATISTICA Cz - [Da	ta: Tabulka6*	(10s krát 15ř)	)]										_ 8 ×
	Soubor Úpr <u>a</u> vy <u>Z</u> e	obrazit Vļožit	Eormát Statist	ika <u>G</u> rafy	Nástroje [	⊇ata <u>O</u> kno	Nápo <u>v</u> ěda							_8×
	🗋 🖻 🚰 🛃 🎒 🕼 🕺 🕸 🏥 🖑 🔛 🕬 構 🎎 Přidat do sešitu * Přidat do protokolu * 🥔 🎌 🗸													
Aria	el	• 10 •	] <b>B</b> <i>I</i> <u>U</u>		= 😭 🔺	- 👌 - 🗃	- 🇞 📗	+.0 .00 0.+ 00.	Ħ 🗗 🗑	68 <b>2</b> ↓	<=? <sup>8</sup> 1 <sup>=</sup> ? Pro	měnné 🔹 Případy 🔹	-	
	1 Vitana	2 Knorr	3 Maggi	4 ∀ar4	5 Var5	6 Var6	7 Var7	8 Var8	9 Var9	10 Var10				
1	5	3	2								_			
2	5	4	1				L;							
4	5	4	3											
5	3	5	1											
6	4	5	2											
7	5	4	1											
9	5	1	2											
10	4	2	1											
11	5	3	3											
12	3	1	1											
14	3	5	2											
15	3	5	3											
														Þ
14														
Připrav	/ena									Ř2;\$7		Filtr - Váhy	: VYPN   CAP   NU	JM REC

#### **Obr. 14**

🙀 Friedmanova ANOVA založená na pořadí: kono	?_×
Zákl. výsledky	Výpočet
Proměnné: žádné	Storno
Výpočet: Friedmanova ANOVA & Kendallova shoda	🔈 Možnosti 🔻
	CRSES <u>f</u>
Krabicový graf všech proměnných	

Nastavíme proměnné, provedeme výpočet ( tlačítkem *Výpočet nebo Výpočet: Friedmanova ANOVA&Kendallova shoda* ). Výstupní sestava je uvedena na obr.15. V sestavě je uvedena kritická hodnota testovacího kritéria ( *ANOVA chí-kv* ), počet hodnot v každém výběru ( *N* ), počet stupňů volnosti ( *sv =počet výběrů – 1* ), minimální hladina významnosti, pro níž lze  $H_0$  zamítnout ( *p* ). V prvních dvou sloupcích jsou uvedeny hodnoty *průměrných pořadí* a *součtů pořadí hodnot* statistického znaku v každém z výběrů. V posledních dvou sloupcích jsou pak uvedeny *průměry* a *směrodatné odchylky* hodnot statistického znaku v každém z výběrů. Není však nabízen žádný test mnohonásobného porovnání. Ten musíme provést sami. Např. použijeme Nemenyiho metodou. Pro rychlejší dotestování využijeme informace ve výstupní sestavě – obr.15.

Obr. 1	15
--------	----

PS 1* 		Friedmanova ANOVA a Kendallův koeficient shody (kono) ANOVA chí-kv. (N = 15, sv = 2) = 14,14035 p < ,00085 Koeficient shody = ,47135 prům. pořadí r = ,43358					
		Průměrné	Součet	Průměr	Sm.Odch.		
	Proměnná	pořadí	pořadí				
	Vitana	2,700000	40,50000	4,333333	0,899735		
	Knorr	1,933333	29,00000	3,000000	1,690309		
	Maggi	1,366667	20,50000	2,066667	0,883715		