

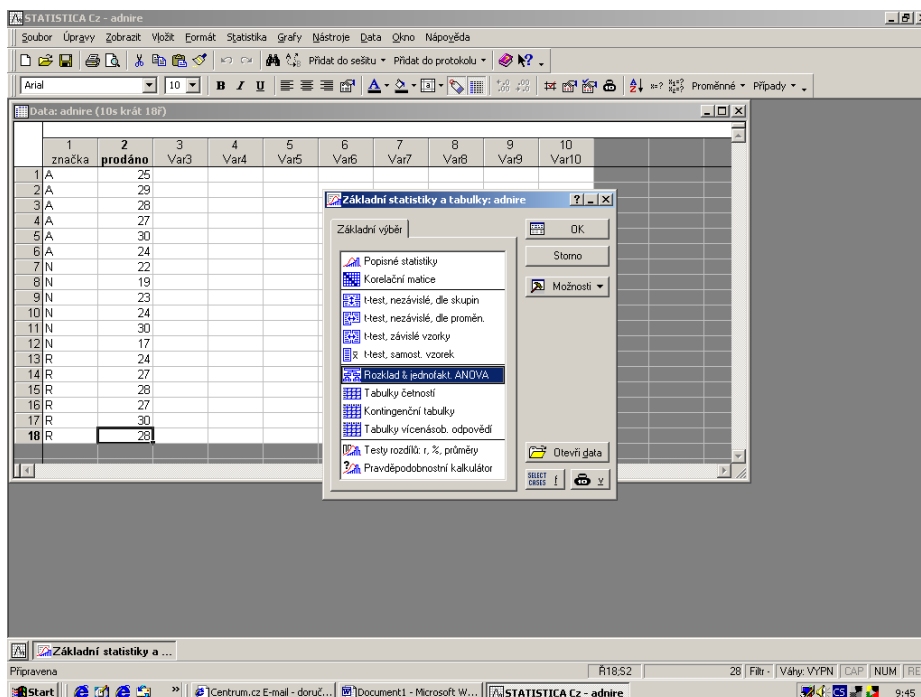
STATISTICA

Téma 7. Testy na základě více než 2 výběrů

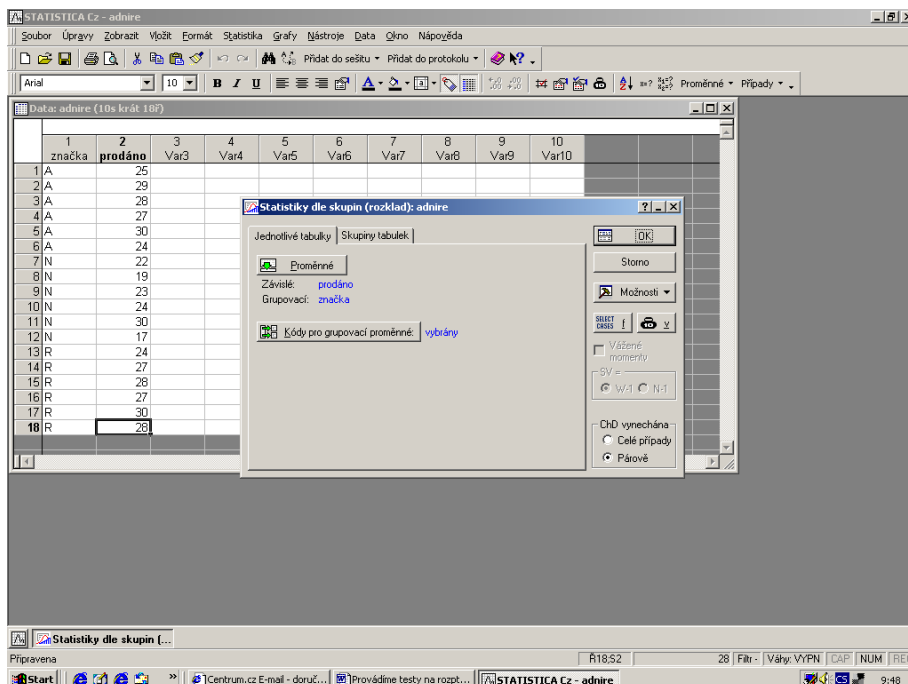
1) Test na homoskedasticitu

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.

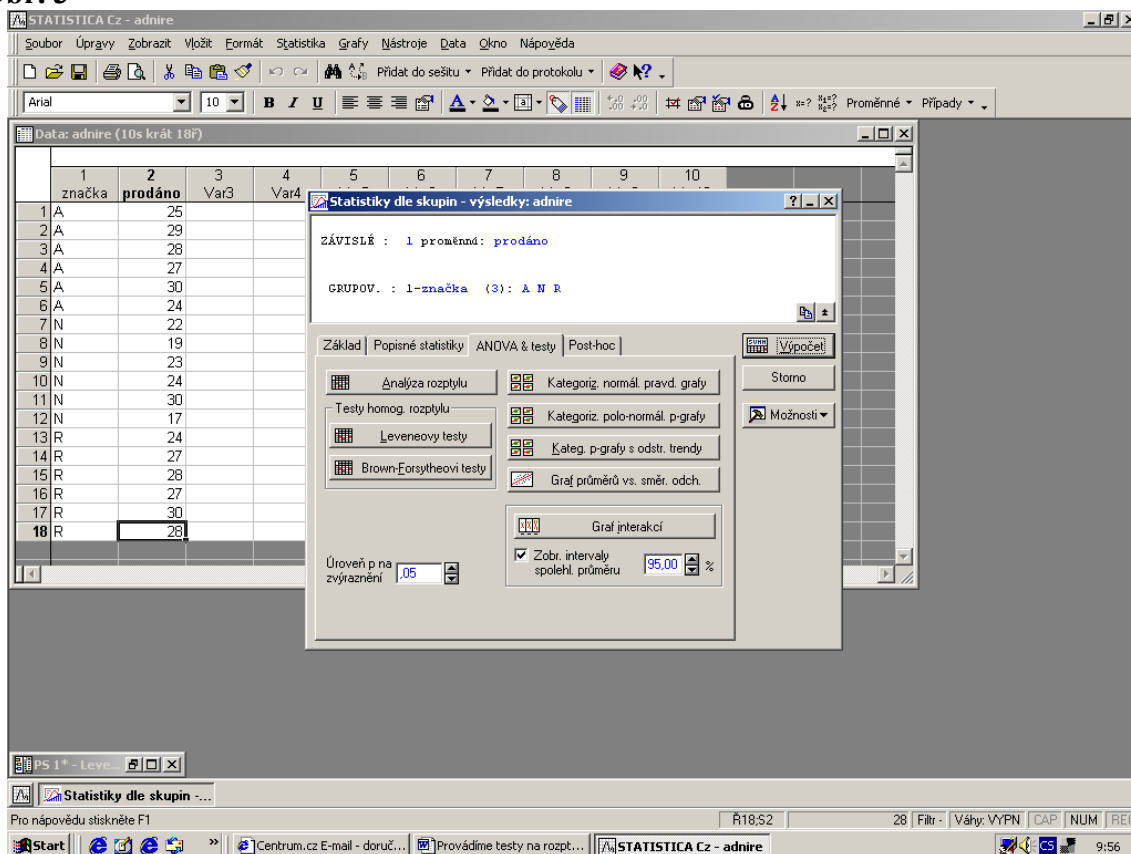
- **závislé** proměnná musí být číselná (např. počet prodaných kusů),
- **grupovací** proměnná musí být kategoriální (např. značka).

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.

Obr. 3



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_0 zamítáme - tj. heteroskedasticita je průkazná. V opačném případě heteroskedasticita průkazná není.

Obr. 4

The screenshot shows the STATISTICA software interface. The main window displays the 'Data: Tabulka1* (10s krát 207)' dataset. A smaller window titled 'PS 1* - Leveneův test homogenity rozpylů (Tabulka1)' is open, showing the results of the Levene's test. The results table is as follows:

Proměnná	SČ efekt	SV efekt	PČ efekt	SČ chyba	SV chyba	PČ chyba	F	p
PRODÁNO	10,77778	2	5,388889	56,66667	15	3,777778	1,426471	0,270932

2) Testy na shodu středních hodnot

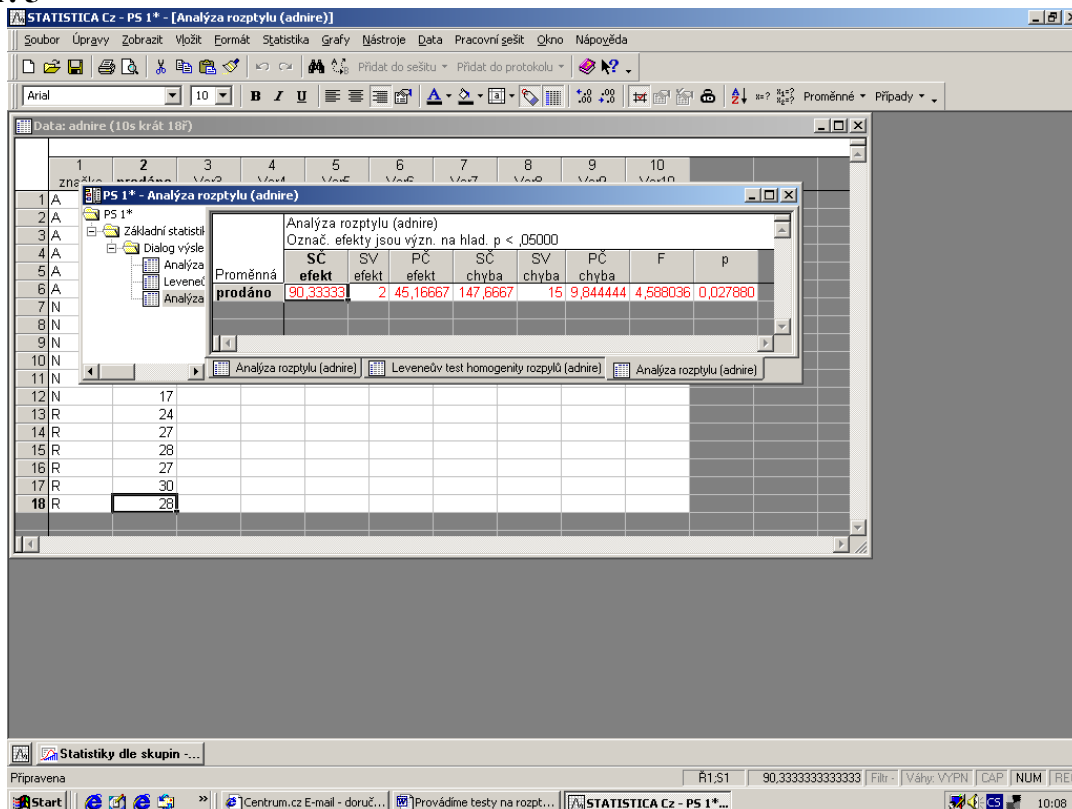
Analýza rozptylu

Analýzu rozptylu používáme v případě, že výběry jsou nezávislé a pocházejí z normálních rozdělení, která vykazují homoskedasticitu. Předpokládejme tedy, že máme k dispozici 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 myši 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.

Obr. 5



Hodnoty uvedené ve sloupcích mají následující význam:

SC 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,

SC chyba – součet čtverců reziduální,

SV chyba – počet stupňů volnosti reziduálního součtu čtverců,

PČ chyba – reziduální rozptyl,

F – hodnota testovacího kritéria,

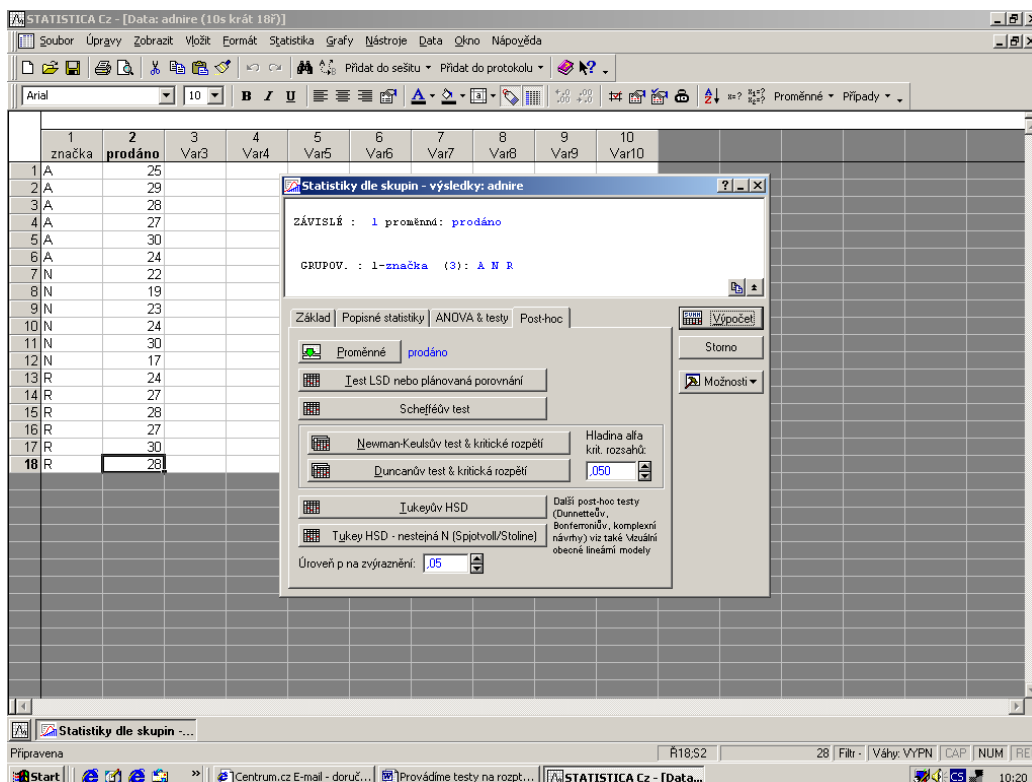
p – minimální hladina významnosti, na níž lze H_0 zamítnout.

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.

Všimněte 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



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

Obr. 7

PS 3* - LSD test; proměnná: PRODÁNO (Tabulka1)

LSD test; proměnná: PRODÁNO (Tabulka1)
Označ. rozdíly jsou významné na hlad. $p < ,05000$

ZNAČKA	{1}	{2}	{3}
A {1}	M=27,167	0,021075	0,927911
N {2}	0,021075		0,017543
R {3}	0,927911	0,017543	

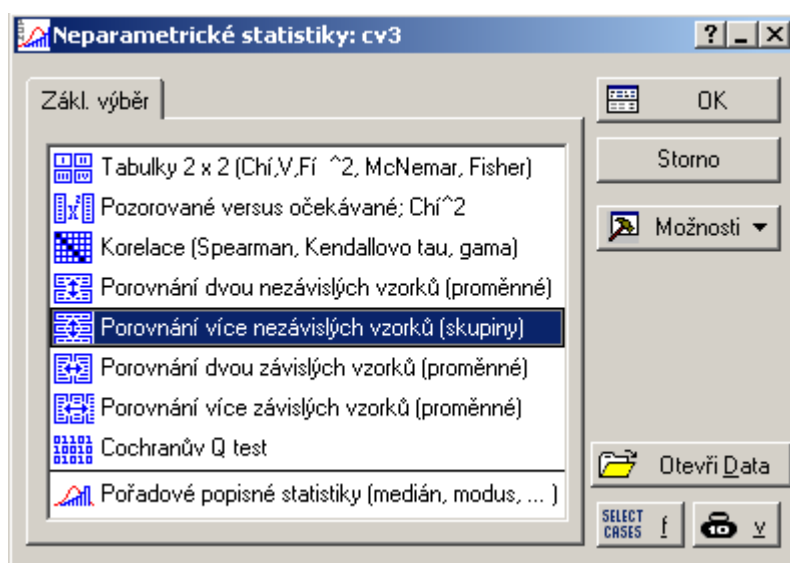
LSD test; proměnná: PRODÁNO (Tabulka1)

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.)

Kruskal – Wallisův test

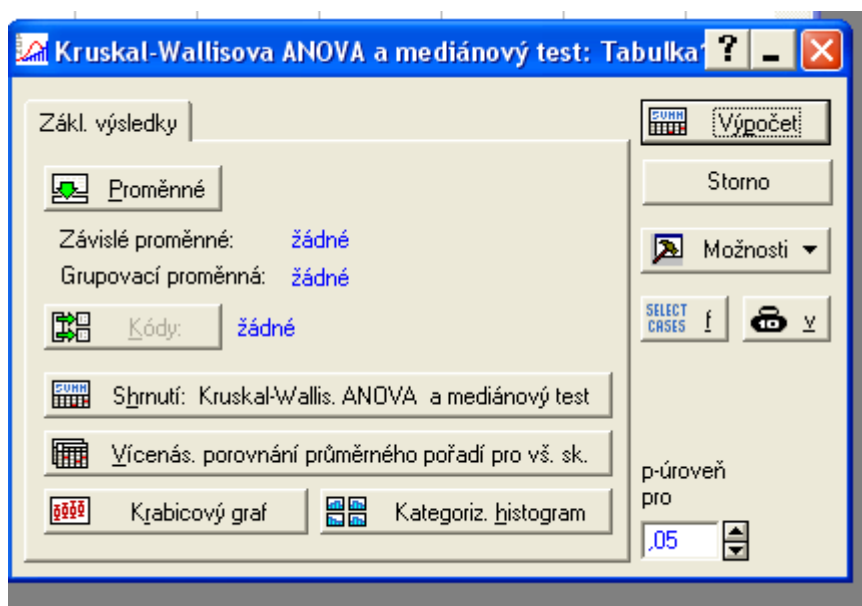
Kruskal – Wallisův test používáme v případě, že výběry jsou nezávislé a pocházejí z normálních rozdělení, která vykazují heteroskedasticitu. Předpokládejme tedy, že máme k dispozici více než dva nezávislé výběry a prostřednictvím Levenova testu jsme zamítli hypotézu o shodě rozptylů. Kruskal – Wallisův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

Obr. 8/



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

Obr. 9



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

Obr.10

PS 4* - Mediánový test, celk. medián =27,0000; PRODÁNO (Tabulka1)

Mediánový test, celk. medián =27,0000; PRODÁNO (Tabulka1)
Nezávislá (grupovací) proměnná : ZNAČKA
Chí-kvadrát = 1,870130 , sv = 2 , p = ,3926

Závislá:
PRODÁNO

	A	N	R	Celkem
<= Medián: pozorov.	3,000000	5,00000	3,000000	11,00000
očekáv.	3,666667	3,66667	3,666667	
poz.-oč.	-0,666667	1,33333	-0,666667	
> Medián: pozorov.	3,000000	1,00000	3,000000	7,00000
očekáv.	2,333333	2,33333	2,333333	
poz.-oč.	0,666667	-1,33333	0,666667	
Celkem: oček.	6,000000	6,00000	6,000000	18,00000

Mediánový test, celk. medián =27,0000; PRODÁNO (Tabulka1)

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.

Obr. 11

PS 4* - Kruskal-Wallisova ANOVA založ. na poř.; PRODÁNO (Tabulka1)

Kruskal-Wallisova ANOVA založ. na poř.; PRODÁNO (Tabulka1)
Nezávislá (grupovací) proměnná : ZNAČKA
Kruskal-Wallisův test: H (2, N= 18) =5,137461 p =,0766

Závislá:
PRODÁNO

	Kód	Počet platných	Součet pořadí
A	101	6	69,00000
N	102	6	33,00000
R	103	6	69,00000

Kruskal-Wallisova ANOVA založ. na poř.; PRODÁNO (Tabulka1)

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 *p* důležité pro vyhodnocení průkaznosti rozdílu mezi součty pořadí příslušných kategorií grupovací proměnné.)

Obr. 12

Vícenásobné porovnání p hodnot (oboustr.); c (cv3)

Nezávislá (grupovací) proměnná : s

Kruskal-Wallisův test: $H(2, N=26) = 9,919486$ $p = ,0070$

Závislá:	s	p	d
c	R:14,450	R:18,111	R:6,2143
s		1,000000	0,001250
p	1,000000		0,000089
d	0,001250	0,000089	

Friedmanův test

Friedmanův test používáme v případě, že výběry jsou závislé. Předpokládejme tedy, že máme k dispozici 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

	1 Vitana	2 Knorr	3 Maggi	4 Var4	5 Var5	6 Var6	7 Var7	8 Var8	9 Var9	10 Var10
1	5	3	2							
2	5	1	3							
3	5	4	1							
4	5	4	3							
5	3	5	1							
6	4	5	2							
7	5	4	1							
8	5	1	3							
9	5	1	2							
10	4	2	1							
11	5	3	3							
12	3	1	1							
13	5	1	3							
14	3	5	2							
15	3	5	3							

Obr. 14

Friedmanova ANOVA založená na pořadí: kono

Zákl. výsledky

Proměnné: **žádné**

Výpočet: Friedmanova ANOVA & Kendallova shoda

Křabicový graf všech proměnných

Výpočet Storno Možnosti SELECT CASES

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. 15

PS 1*		Friedmanova ANOVA a Kendallův koeficient shody (kono)				
Neparametrická		ANOVA chí-kv. (N = 15, sv = 2) = 14,14035 p < ,00085				
Dialog Fried		Koeficient shody = ,47135 prům. pořadí r = ,43358				
Friedma						
Proměnná	Průměrné pořadí	Součet pořadí	Průměr	Sm.Odch.		
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		