Neparametrické testy

Vytvořil Institut biostatistiky a analýz, Masarykova univerzita
J. Jarkovský, L. Dušek

Parametrické vs. neparametrické testy

Parametrické testy

- Mají předpoklady o rozložení vstupujících dat (např. normální rozložení..)
- Při stejném N a dodržení předpokladů mají vyšší sílu testu než testy neparametrické
- Pokud nejsou dodrženy předpoklady parametrických testů, potom jejich síla testu prudce klesá a výsledek testu může být zcela chybný a nesmyslný !

Neparametrické testy

- Nemají předpoklady o rozložení vstupujících dat, lze je tedy použít i při asymetrickém rozložení, odlehlých hodnotách, či nedetekovatelném rozložení
- Snížená síla těchto testů je způsobena redukcí informační hodnoty původních dat, kdy <u>neparametrické testy nevyužívají původní hodnoty, ale nejčastěji pouze jejich **pořadí**</u>

Statistické testy a normalita dat

- **Normalita dat je jedním z předpokladů tzv. parametrických testů (testů založených na předpokladu nějakého rozložení) – např.** *t*-testy
- **Pokud data nejsou normální, neodpovídají ani modelovému rozložení, které je použito pro výpočet (*t*-rozložení) a test tak může lhát**

- **Řešením je tedy:**
  - Transformace dat **za účelem dosažení normality jejich rozložení**
  - Neparametrické testy **– tyto testy nemají žádné předpoklady o rozložení dat**

Typ srovnání	Parametrický test	Neparametrický test
2 skupiny dat nepárově:	Nepárový t-test	Mann Whitney test
2 skupiny dat párově:	Párový t-test	Wilcoxon test, znaménkový test
Více skupin nepárově:	ANOVA	Kruskal- Wallis test
Korelace:	Pearsonův koeficient	Spearmanův koeficient

Neparametrické alternativy nepárového t-testu

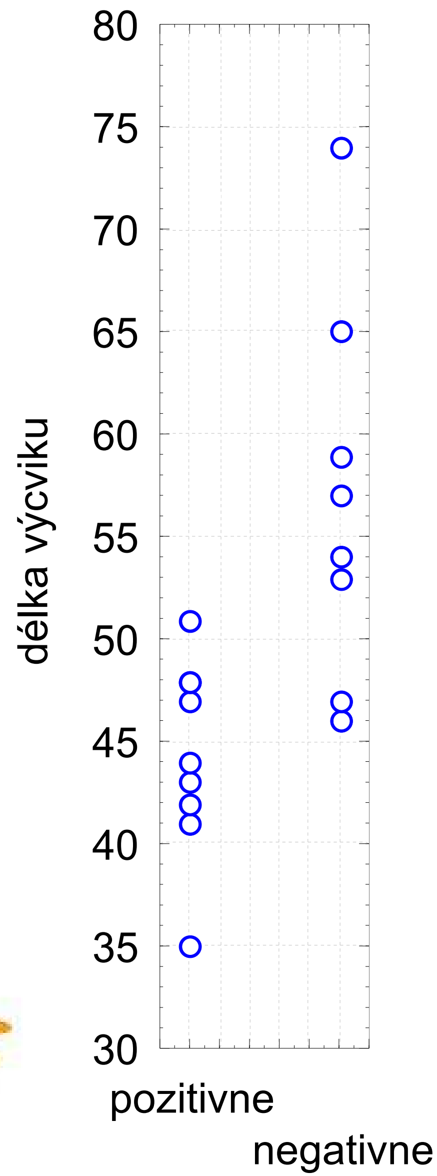
X1	X2	ALL	Rank ALL	X1 rank	X2 rank
27	25	25	5	6	5
35	29	29	7,5	11	7,5
38	31	31	9	13	9
37	23	23	4	12	4
39	18	18	2	14	2
29	17	17	1	7,5	1
41	32	32	10	15	10
	19	19	3		3
		27	6		
		35	11		
		38	13		
		37	12		
		39	14		
		29	7,5		
		41	15		

Mann Whitney U-test

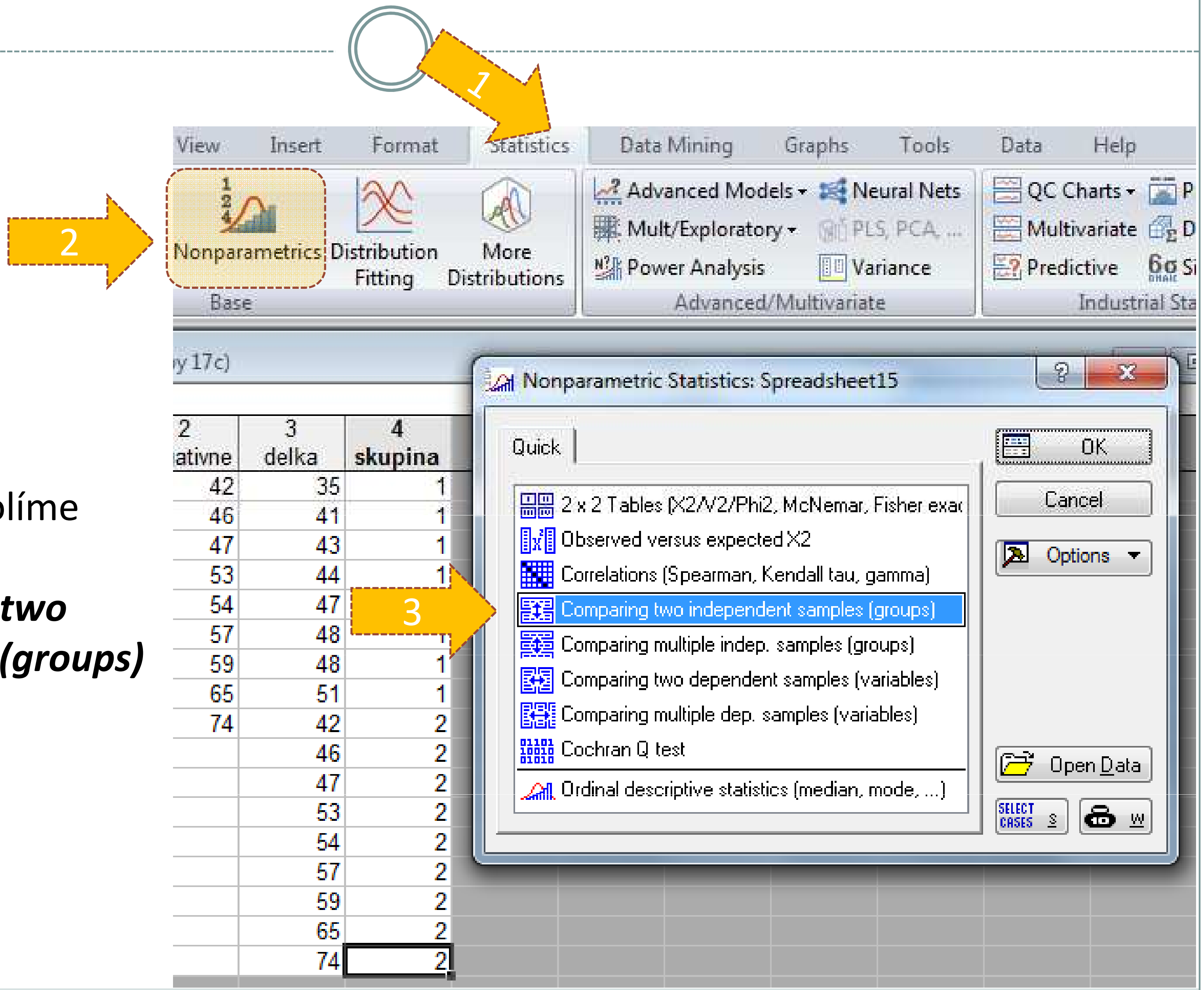
- Stejně jako řada jiných neparametrických testů počítá i tento test s pořadím dat v souborech namísto s originálními daty. Jde o neparametrickou obdobu nepárového t-testu a z těchto neparametrických testů má nejvyšší sílu testu (95% párového t-testu).
- V případě Mann-Whitney testu jsou nejprve čísla obou souborů sloučena a je vytvořeno jejich pořadí v tomto sloučeném souboru, pak jsou hodnoty vráceny do původních souborů a nadále se pracuje již jen s jejich pořadím.
- Pro oba soubory je tedy vytvořen součet pořadí a menší z obou součtů je porovnán s kritickou hodnotou testu, pokud je tato hodnota menší než kritická hodnota testu, zamítáme nulovou hypotézu shody distribučních funkcí obou skupin.
- Podobným způsobem je počítán i **Wilcoxon rank sum test** (pozor, existuje ještě Wilcoxnův párový test!!!)

Příklad 1: Mann – Whitney U test

- 17 štěňat bylo trénováno v chození na záchod metodou pozitivního posilování (pochvala, když jde na záchod venku) nebo negativního (trest, když jde na záchod doma). Jako parametr bylo měřeno, za kolik dní je štěně vycvičeno.
- nulová hypotéza je, že není rozdíl v metodách tréninku, tedy, že oběma metodami je štěně vycvičeno za stejnou dobu.
- po srovnání rozložení + malý počet hodnot je vhodné použít neparametrický test
- je vytvořeno pořadí sloučených hodnot
- pořadí hodnot v jednotlivých skupinách dat je sečteno a menší ze součtů je použit pro srovnání s kritickou hodnotou testu
- výsledkem testu je p<α, nulovou hypotézu tedy zamítáme a výsledkem testu je, že pozitivní působení při výcviku štěňat dává lepší výsledky



Příklad 1: Řešení v softwaru Statistica I



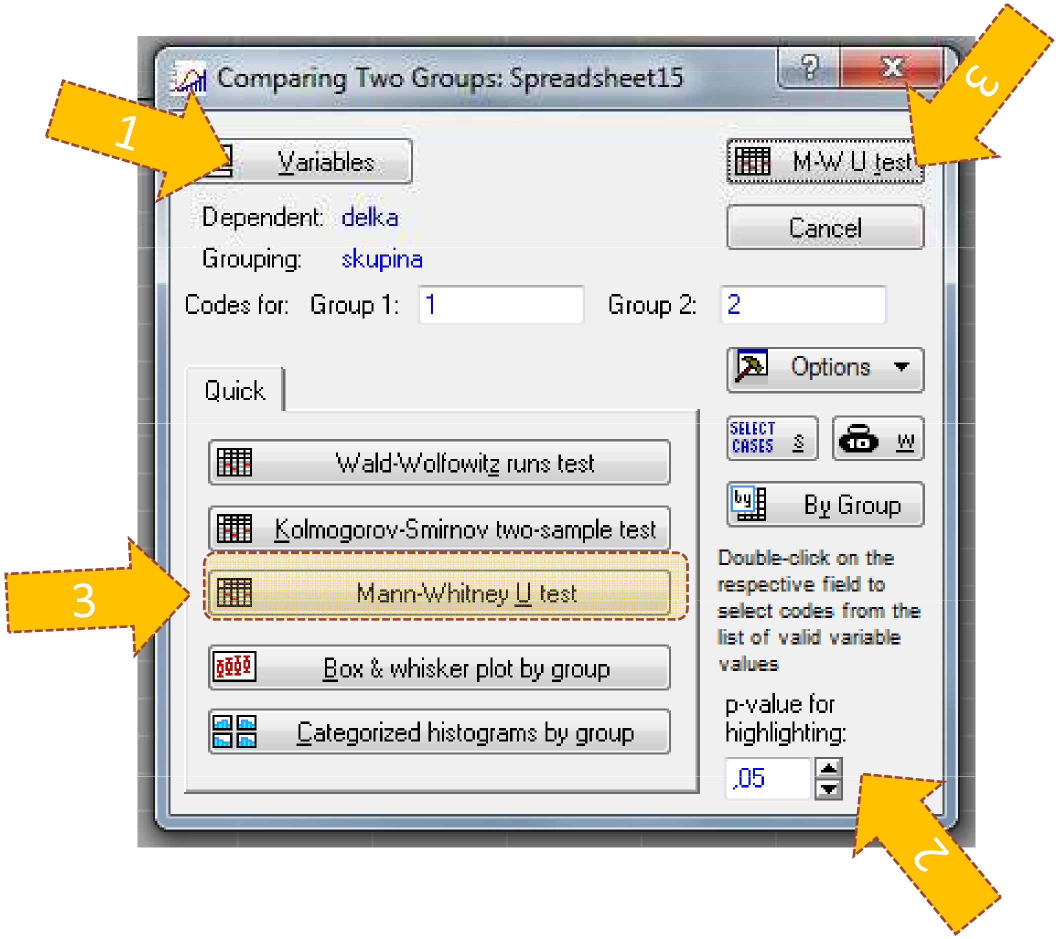
- V menu ***Statistics*** zvolíme ***Nonparametrics ,*** vybereme ***Comparing two independent samples (groups)***

Příklad 1: Řešení v softwaru Statistica II

- Vybereme proměnné, které chceme testovat

- ***p-value for highlighting-*** Úroveň p lze změnit

- Kliknutím na ***Mann-Whitney U test,*** nebo na M-W U test získáme výstupy:



Řešení: Mann-Whitney test v Statistica III

Součet pořadí T_1

Součet pořadí T_2

Hodnota asymptotické testové statistiky

Mann-Whitney U Test (Spreadsheet15)
By variable skupina
Marked tests are significant at p <,05000

variable	Rank Sum Group 1	Rank Sum Group 2	U	Z	p-value	Z adjusted	p-value	Valid N Group 1	Valid N Group 2	2*1sided exact p
delka	49,50000	103,5000	13,50000	-2,11695	0,034265	-2,11955	0,034045	8	9	0,027396

Hodnota testové statistiky

Asymptotická p- hodnota

Přesná p- hodnota
(označení 2*1sided exact p- použít, jestliže rozsah výběru je menší než 30)

Neparametrická obdoba párového t-testu

Párový Wilcoxonův test

- Jsou vytvořeny diference mezi soubory, je vytvořeno jejich pořadí bez ohledu na znaménko a poté je sečteno pořadí kladných a pořadí záporných rozdílů. Menší z těchto dvou hodnot je srovnána s kritickou hodnotou testu a pokud je menší než kritická hodnota testu, pak zamítáme hypotézu shody obou souborů hodnot. Pro test existuje aproximace na normální rozložení, ale pouze pro velká n>25.

$$t = \frac{Menší\_suma\_diferencí - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

Před zásahem	Po zásahu	Změna	Absolutní pořadí
6	2	4	10
2,5	3	-0,5	1,5
6,3	5	1,3	6
8,1	9	-0,9	5
1,5	2	-0,5	1,5
3,4	4	-0,6	3
2,5	1	1,5	8
1,11	2	0,89	4
2,6	4	-1,4	7
1	3	-2	9

Příklad 2: Wilcoxonův párový test

člověk	A	B	diference	pořadí
1	142	138	4	4,5
2	140	136	4	4,5
3	144	147	-3	3
4	144	139	5	7
5	142	143	-1	1
6	146	141	5	7
7	149	143	6	9,5
8	150	145	5	7
9	142	136	6	9,5
10	148	146	2	2

A…….parametr krve před podáním léku

B…….parametr krve po podání léku

W_+ …..součet pořadí přes kladné hodnoty rozdílů = 51

W_- …..součet pořadí přes záporné hodnoty rozdílů = 4

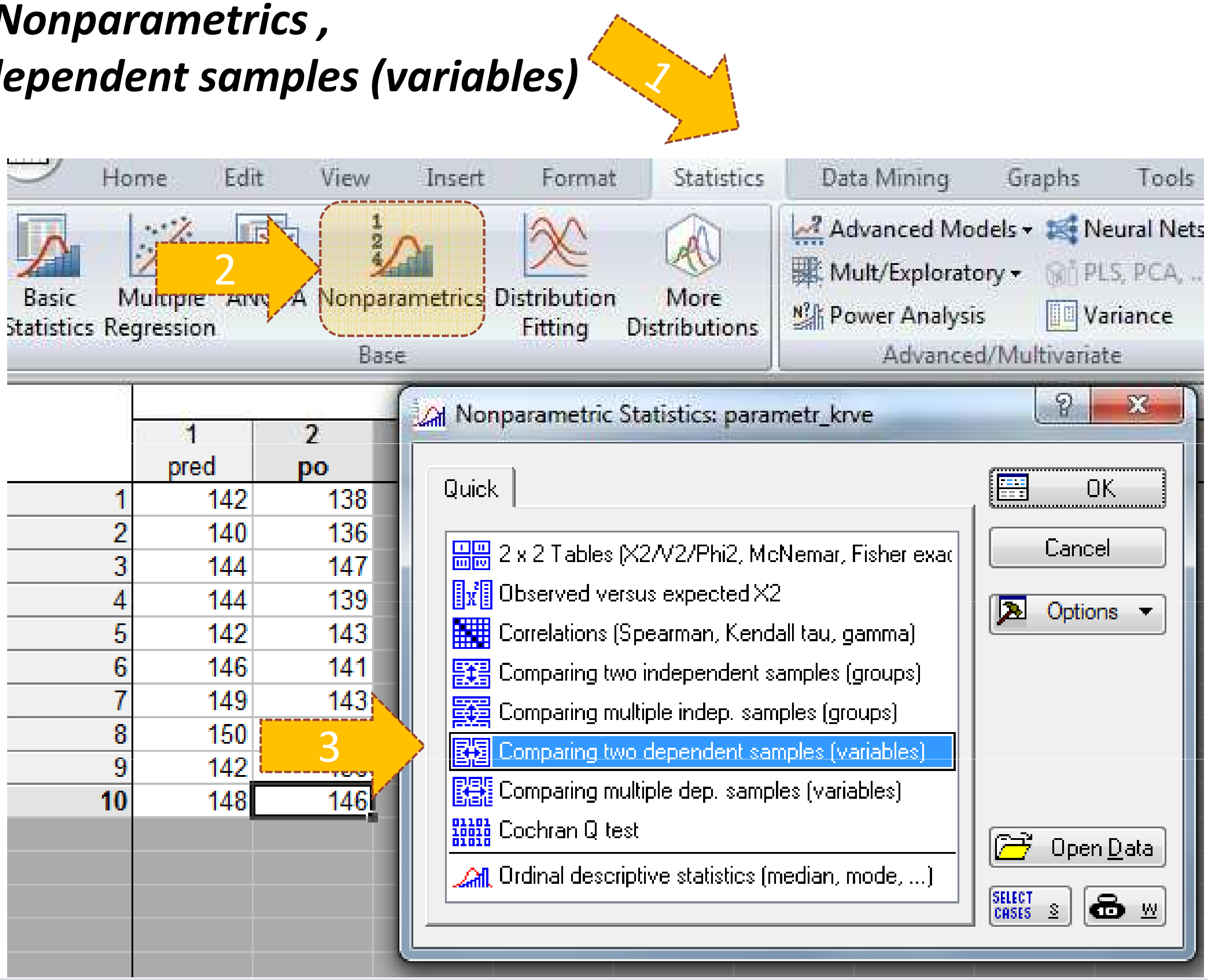
$W = \min(W_+;W_-) = 4$

počet párů = n = 10

Pokud je **W** menší než kritická hodnota testu, pak zamítáme hypotézu shody distribučních funkcí obou skupin.

Příklad 2: Řešení v softwaru Statistica I

- V menu ***Statistics*** zvolíme ***Nonparametrics ,***
vybereme ***Comparing two dependent samples (variables)***



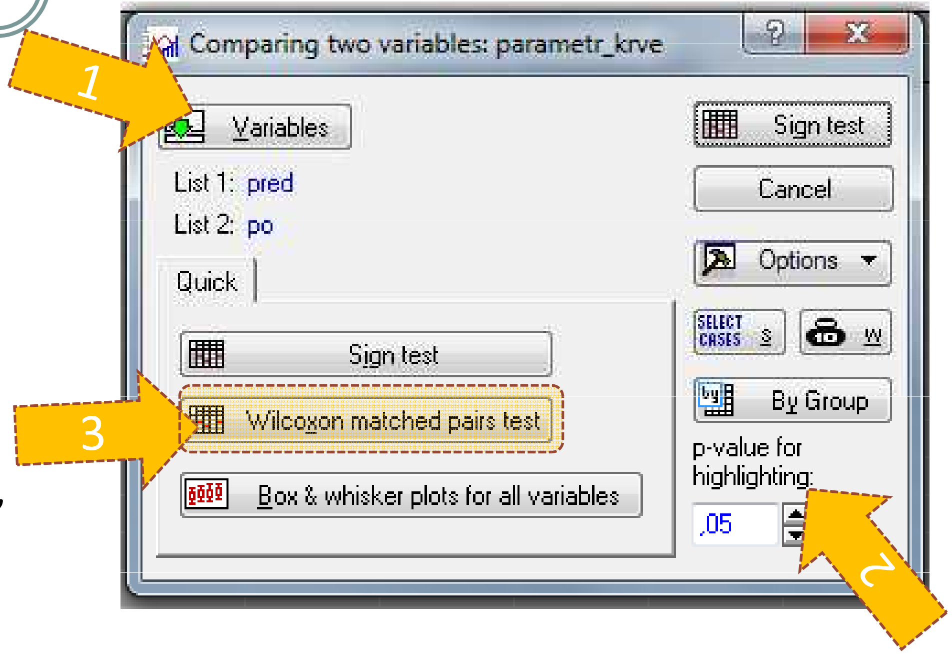
	1 pred	2 po
1	142	138
2	140	136
3	144	147
4	144	139
5	142	143
6	146	141
7	149	143
8	150	
9	142	
10	148	146

Příklad 2: Řešení v softwaru Statistica II

- Vybereme proměnné, které chceme testovat

- ***p-value for highlighting-*** Úroveň p lze změnit

- Kliknutím na ***Wilcoxon matched pairs test,*** získáme výstupy:



Rozsah výběru

	Wilcoxon Matched Pairs Test (parametr_krve) Marked tests are significant at p <,05000			
Pair of Variables	Valid N	T	Z	p-value
pred & po	10	4,000000	2,395342	0,016605

POZOR: podmínka pro použití asymptotické p-hodnoty je: n≥ 30

Hodnota testovací statistiky

<u>Asymptotická p-hodnota</u>

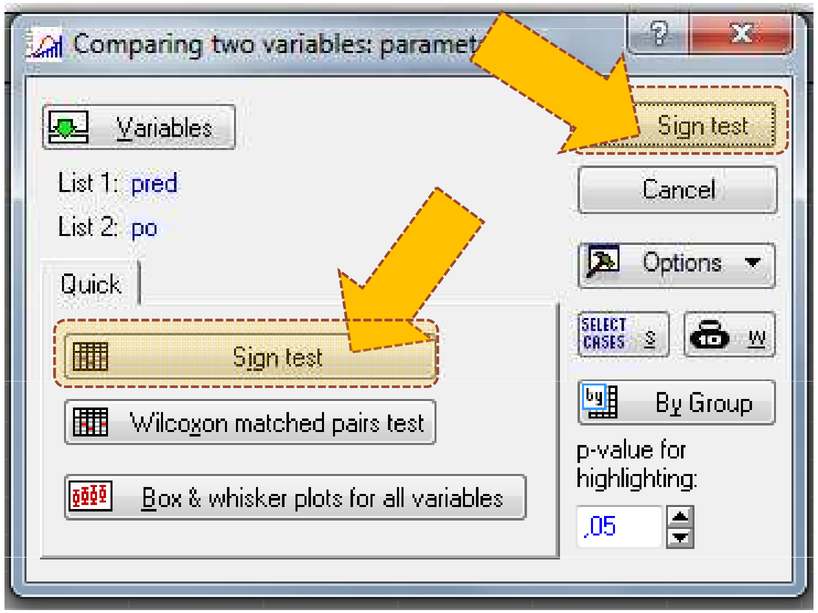
Hodnota asymptotické testové statistiky

Párový znaménkový test

• Vybereme proměnné, které chceme testovat

• ***p-value for highlighting-*** Úroveň p lze změnit

• Kliknutím na ***Sign test (párový znaménkový test)*** získáme výstupy:



Počet nenulových hodnot, z nich záporných je 20%.

Sign Test (parametr_krve)
Marked tests are significant at p <,05000

Pair of Variables	No. of Non-ties	Percent v < V	Z	p-value
pred & po	10	20,00000	1,581139	0,113846

POZOR: podmínka pro použití asymptotické p-hodnoty je: n > 20

Hodnota asymptotické testové statistiky

Asymptotická p-hodnota

Znaménkový test – příklad I

Párově uspořádaný experiment pro nominální data

I. Dva preparáty, každý na ½ listu

- sledovaná veličina: počet skvrn (hodnoceno pouze jako rozdíl)

	Počet skvrn									
A	V	V	M	V	V	M	M	V	V	V
B	M	M	V	M	M	V	V	M	M	M

V – větší; M – menší

n = 10 listů s rozdílnými výsledky

jev

A je větší: + $n_+ = 7$

B je menší: - $n_- = 3$

$\min(n_+; n_-) = 3$

II. dvě protilátky z různých zdrojů (A;B)

– aplikované na vzorek s antigenem

n = 10

A	+	+	-	+	-	+	-	+	+	-
B	-	-	+	-	+	+	-	-	+	-

n – nenulových rozdílů: 6

A: $n_+ = 4$

A: $n_- = 2$

$\min(n_+; n_-) = 2$

Znaménkový test – příklady II

- Na konferenci veterinářů bylo předneseno,že průměrný čas konzultace je 12 minut. Následovala debata, zda je lepší použít medián nebo průměr. Jeden z nich se rozhodl ověřit teorii, že průměrná konzultace trvá 12 minut na vlastní praxi a zaznamenal si trvání svých 43 konzultací. K otestování hypotézy, že podíl konzultací kratších a delších než 12 minut použil znaménkový test.

Délka konzultace	Počet
<12	22
12	6
>12	15
Celkem	43

Další výpočet probíhá obdobně jako v případě klasického znaménkového testu na diferencích dvou skupin dat.

Neparametrická obdoba analýzy rozptylu

Kruskalův – Wallisův test

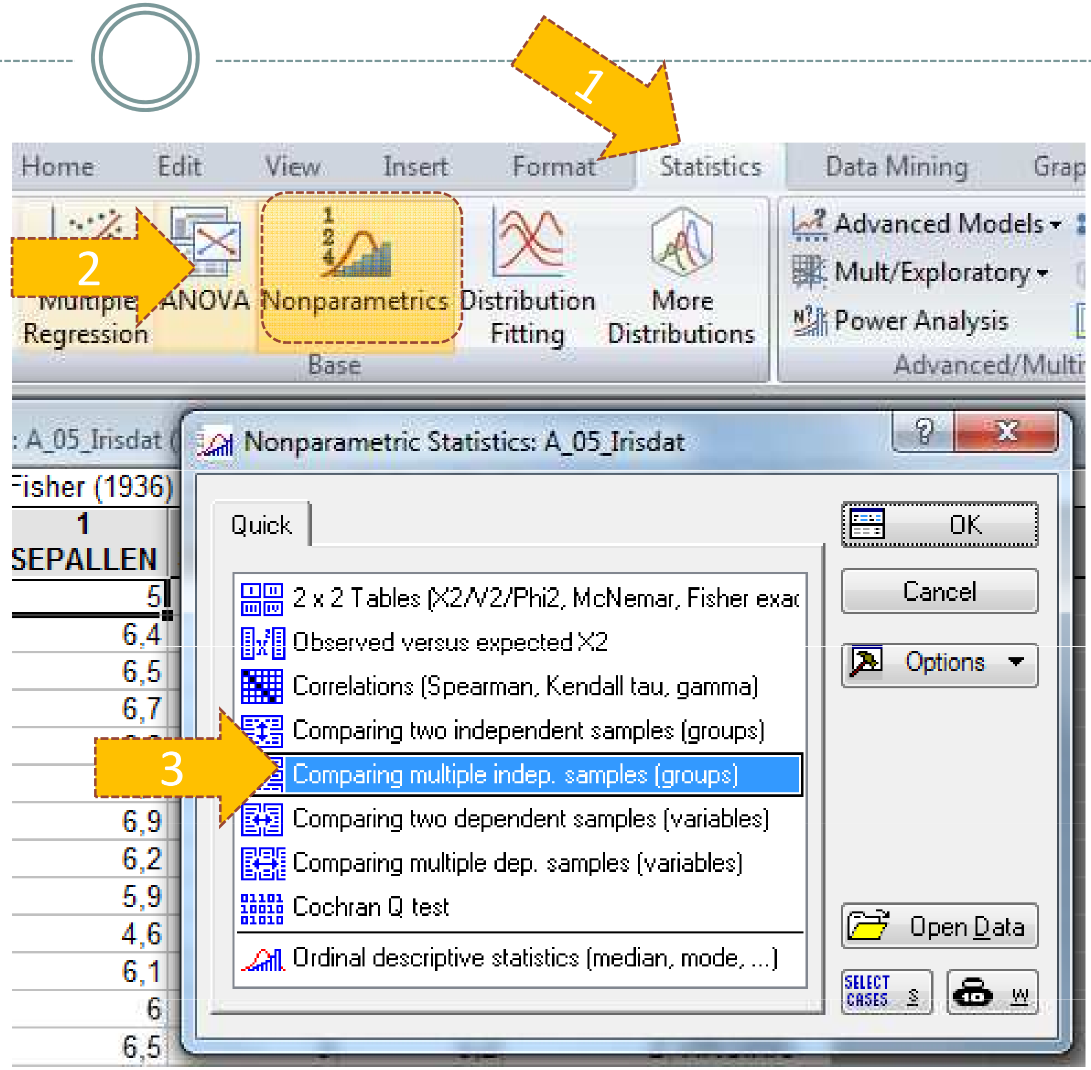
- K dispozici jsou alespoň 3 nezávislé náhodné výběry
- Nulová hypotéza tvrdí, že všechny tyto výběry pocházejí z téhož rozložení
- Nejprve všechny hodnoty uspořádáme a určíme pořadí každé hodnoty, poté pro každý výběr sečteme pořadí hodnot (T_j), které do něj patří . Testová statistika má tvar: $$Q = \frac{12}{n(n+1)} \sum_{j=1}^{r} \frac{T_j^2}{n_j} - 3(n+1)$$
- V případě zamítnutí nulové hypotézy, se ptáme, které dvojice náhodných výběrů se liší, k tomuto účelu je vhodné použít metody mnohonásobného porovnávání

Příklad 3: Kruskalův- Wallisův test

- Bylo získáno 150 kosatců pocházejících ze tří základních tříd: iris setosa, iris versicolor, iris virginica. Z botaniky je známo že iris versicolor je hybridem zbývajících dvou druhů. U květů byly měřeny následující údaje: délka a šířka kališních lístků, délka a šířka korunních plátků.
- Na hladině významnosti 0,05 testujte hypotézu, že délka kališních lístků u třech tříd kosatců se neliší. Pokud zamítnete nulovou hypotézu, zjistěte, které dvojice tříd se od sebe liší.

Iris virginica Iris versicolor Iris setosa

Příklad 3: Řešení v softwaru Statistica I



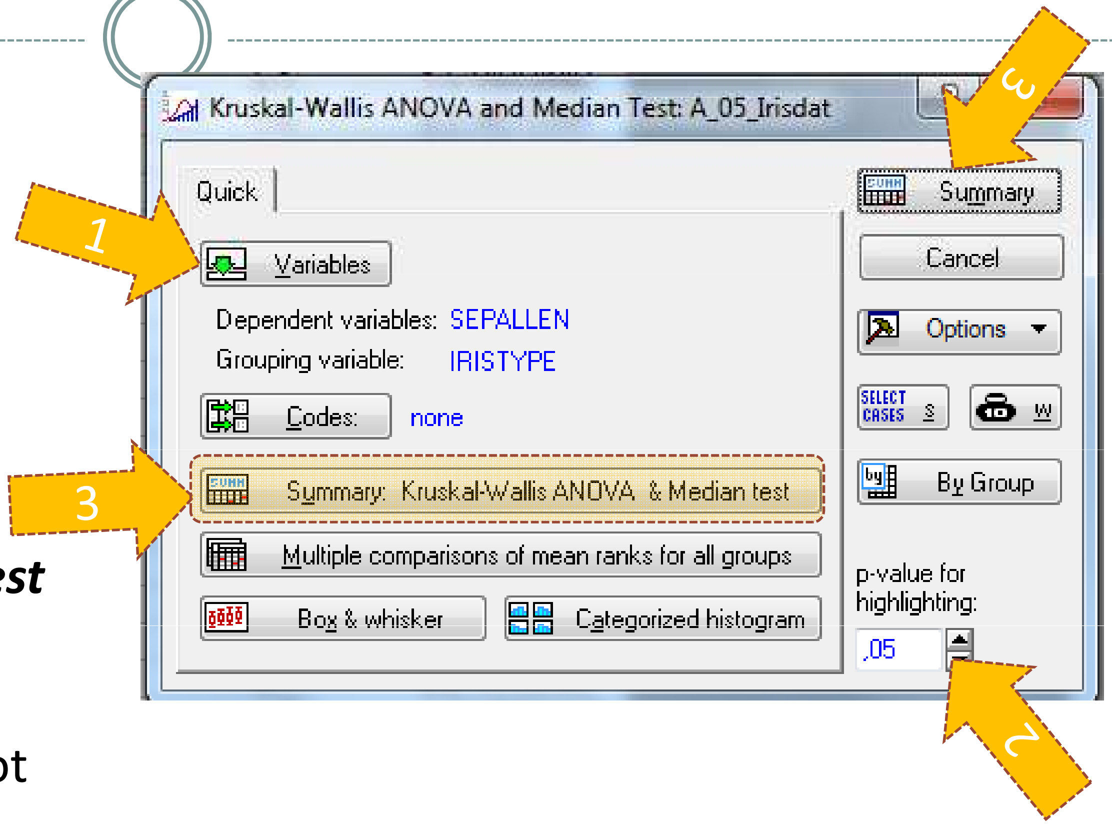
- V menu ***Statistics*** zvolíme ***Nonparametrics ,*** vybereme ***Comparing multiple Indep. samples (groups)***

Příklad 3: Řešení v softwaru Statistica II

- Vybereme proměnné, které chceme testovat

- ***p-value for highlighting-*** Úroveň p lze změnit

- Kliknutím na ***Summary: Kruskal-Wallis ANOVA & Median test*** získáme výstupy.



Počet hodnot v každém výběru

Součet pořadí hodnot

Kruskal-Wallis ANOVA by Ranks; SEPALLEN (A_05_Irisdat)
Independent (grouping) variable: IRISTYPE
Kruskal-Wallis test: H (2, N= 150) =96,93744 p =0,000

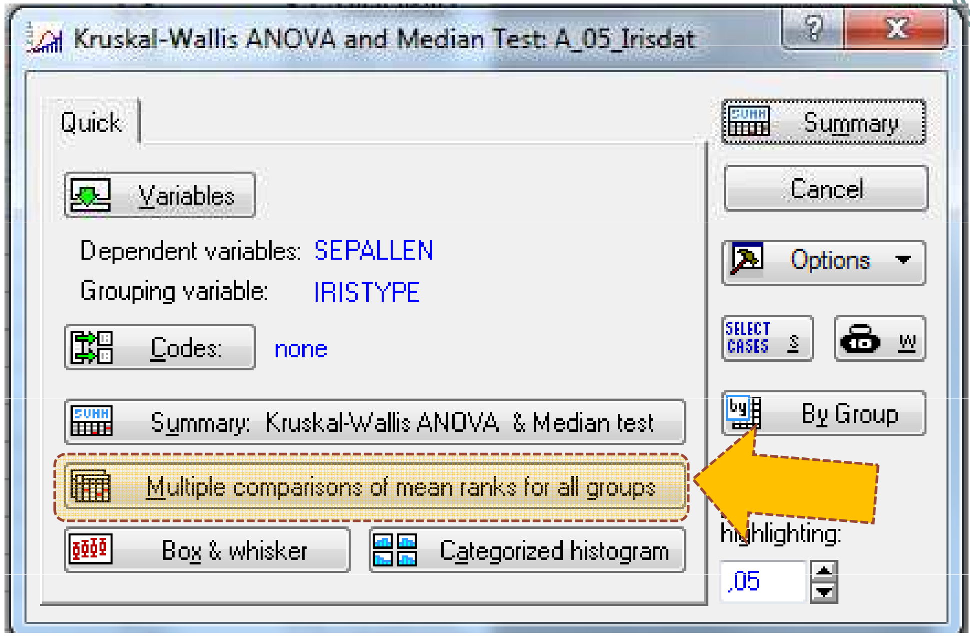
Depend.: SEPALLEN	Code	Valid N	Sum of Ranks	Mean Rank
SETOSA	1	50	1482,000	29,6400
VERSICOL	2	50	4132,500	82,6500
VIRGINIC	3	50	5710,500	114,2100

Hodnota testové statistiky

p-hodnota,

Je– li rozdíl mezi středními hodnotami průkazný (p< 0,05), musíme provést **testy mnohonásobného porovnání**.

Příklad 3: Řešení v softwaru Statistica III



<u>Testy mnohonásobného porovnávání</u>

- Kliknutímna ***Multiple comparisons of mean ranks for all groups***

Multiple Comparisons p values (2-tailed); SEPALLEN (A_05_Irisdat)
Independent (grouping) variable: IRISTYPE
Kruskal-Wallis test: H (2, N= 150) =96,93744 p =0,000

Depend.: SEPALLEN	**SETOSA R:29,640**	VERSICOL R:82,650	VIRGINIC R:114,21
SETOSA		0,000000	0,000000
VERSICOL	0,000000		0,000843
VIRGINIC	0,000000	0,000843	

<u>p-hodnoty</u>