ASTAc/02 Biostatistika
5. cvičení

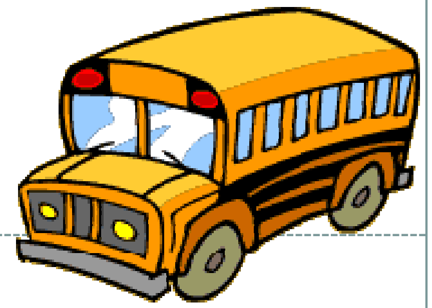
Opakování
Shrnutí statistických testů
Neparametrické testy

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

Co byste měli umět z minula:

1. Vybrat typ parametrického testu – jednovýběrový, párový nebo dvouvýběrový?
2. Ověřit předpoklady parametrických testů (normalitu, shodu rozptylů; graficky i pomocí testů).
3. Provést testování v softwaru Statistica.
4. Interpretovat výsledky testování.

Jednovýběrový t-test: doplnění



- Určitá linka autobusové městské dopravy má v době dopravní špičky průměrnou rychlost 8 km/hod. Uvažovalo se o tom, zda by změna trasy vedla ke změně průměrné rychlosti. Nová trasa byla proto projeta v deseti náhodně vybraných dnech a byly zjištěny tyto průměrné rychlosti: 7,8; 7,9; 9,0; 7,8; 8,0; 7,8; 8,5; 8,2; 8,2; 9,3. Rozhodněte, zda změna trasy vede ke změně průměrné rychlosti. Předpokládáme normální rozdělení a α=0,05.
- **Postup:**

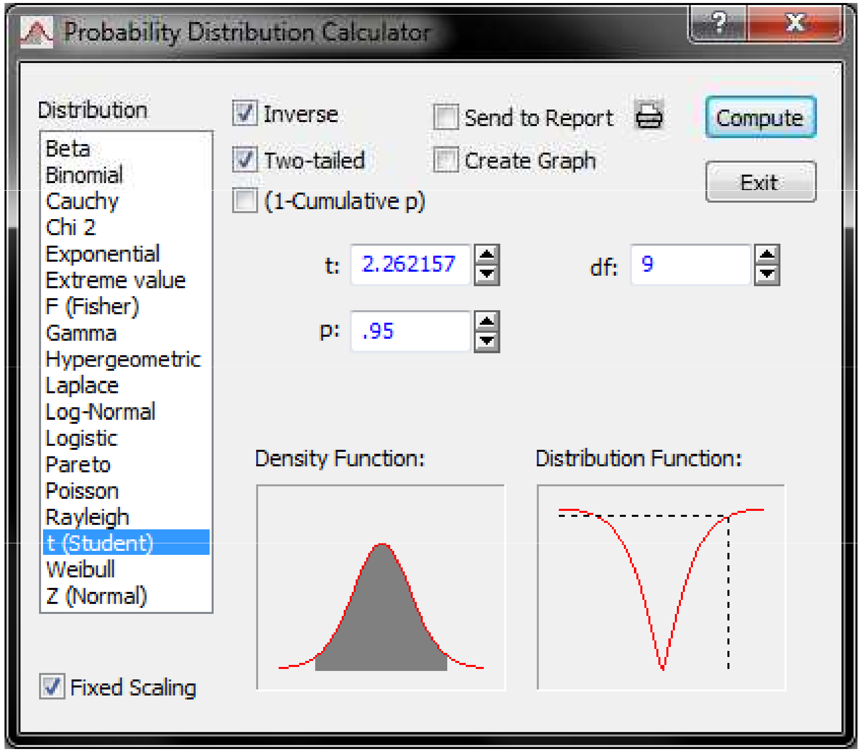
1. Na hladině významnosti 0,05 testujeme **hypotézu H_0:** $\mu = 8$ **, proti H_A :** $\mu \neq 8$
2. Vypočteme **aritmetický průměr** a **rozptyl výběrového souboru.**
3. Vypočteme **testovou statistiku t:**
$$t = \frac{\bar{x} - \mu}{s}\sqrt{n} = \frac{8{,}25 - 8}{0{,}530}\sqrt{10} = 1{,}492$$
4. Vypočtené **t porovnáme s kritickou hodnotou $t_{1-\alpha/2(n-1)}$:** $\boxed{t_{0,975}(9) = 2{,}262}$

Jak zjistíme hodnotu kvantilu pomocí softwaru Statistica, abychom nemuseli hledat v matematických tabulkách? Jak bez složitých výpočtů zjistit p-hodnotu?

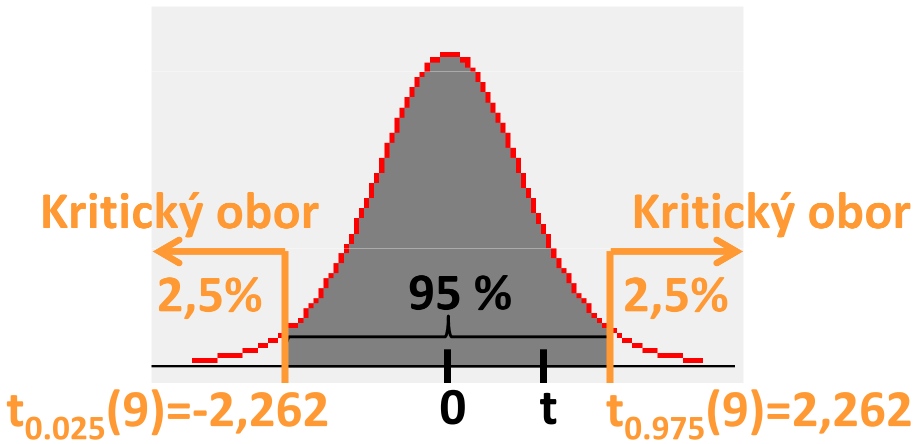
Výpočet kritického oboru ve Statistice

4. Určení **kritického oboru** v softwaru Statistica

   \- Postup: *Statistics* → *Basic Statistics* → *Probability calculator* → vybereme typ rozdělení *t (Studentovo)* → zadáme stupně volnosti (*df*=9) → do *p* zadáme hodnotu 1-hladina významnosti → zaškrtneme Two-tailed (v případě oboustranné nulové hypotézy) → *Compute* → spočítá se hodnota *t*, která definuje kritický obor.



Rozložení testové statistiky za předpokladu platnosti nulové hypotézy

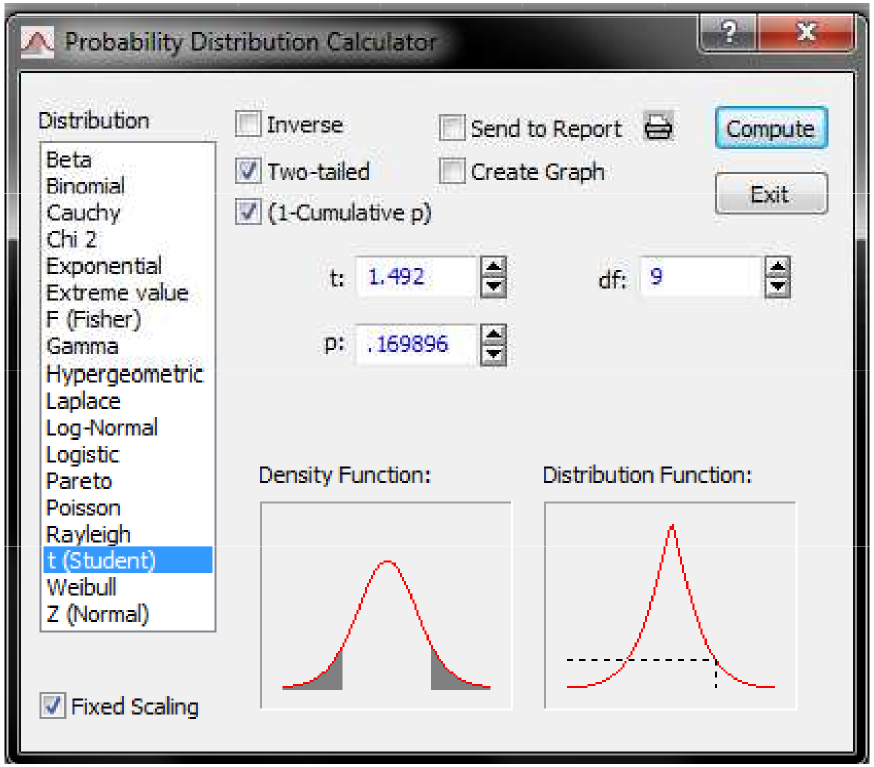


Kritický obor = obor zamítnutí nulové hypotézy. Pokud testová statistika padne do této oblasti, zamítáme nulovou hypotézu.

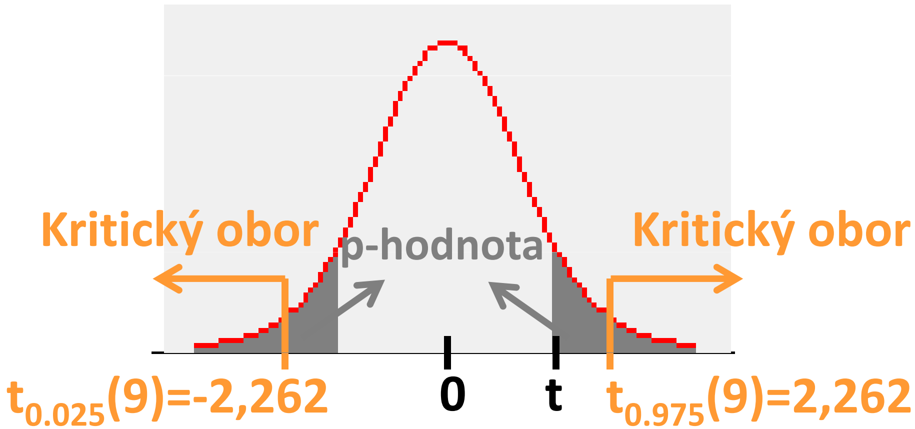
Výpočet p-hodnoty ve Statistice

5. Určení **p-hodnoty** z hodnoty testové statistiky v softwaru Statistica

   - Postup: *Statistics* → *Basic Statistics* → *Probability calculator* → vybereme typ rozdělení *t (Studentovo)* → zadáme stupně volnosti (df = 9) → zadáme hodnotu testové statistiky (t=1,492) → zaškrtneme *Two-tailed* a *(1-Cumulative p)* v případě oboustranné nulové hypotézy)→ *Compute* → spočítá se p-hodnota.

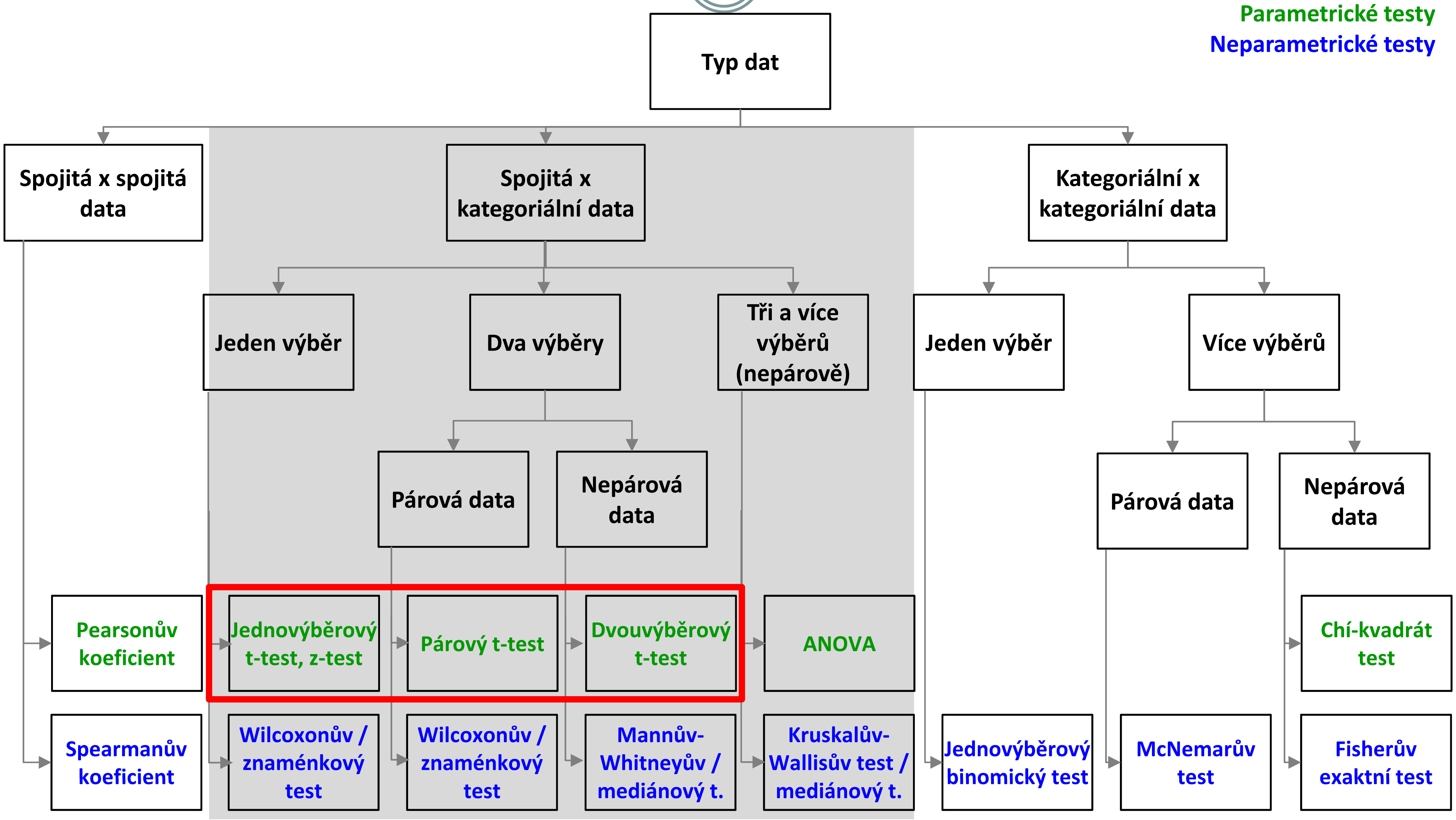


Rozložení testové statistiky za předpokladu platnosti nulové hypotézy

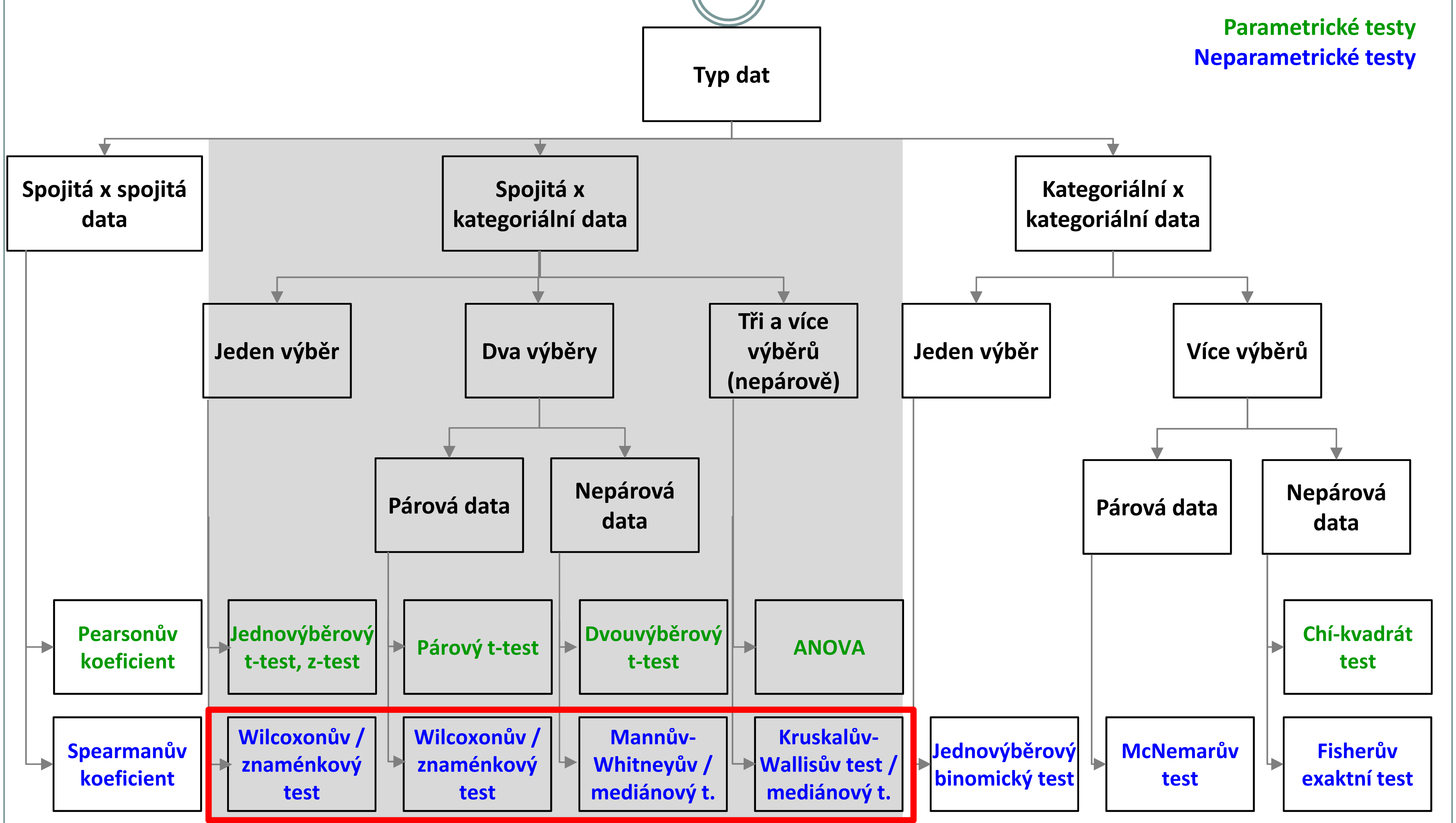


P-hodnota vyjadřuje pravděpodobnost, že testová statistika nabyde stejné nebo extrémnější hodnoty za předpokladu, že nulová hypotéza platí.

Základní rozhodování o výběru statistických testů - co jsme probírali minule



Základní rozhodování o výběru statistických testů - co budeme probírat dnes



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

- Vyžadují méně předpokladů 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 neparametrické testy nevyužívají původní hodnoty, ale nejčastěji pouze jejich pořadí
- Souvisí s malou velikostí souboru (nejsme schopni normalitu dat ověřit)

Proč nemusí parametrický a neparametrický test vyjít stejně?

1. Statistické testy o parametrech jednoho výběru

Jednovýběrový Wilcoxonův test
Znaménkový test

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

Jednovýběrový Wilcoxonův test

- Předpokladem je symetrické rozdělení dat kolem mediánu.
- Testuje, zda je medián jednoho výběru roven hodnotě *c* (v případě párového designu je $x_{0.5}$ reprezentováno mediánem rozdílu hodnot)

$$H_0: x_{0.5}=c \quad \text{proti } H_1: x_{0.5} \neq c.$$

Postup:

1. Spočítáme rozdíly hodnot výběru s testovanou hodnotou mediánu.
2. Absolutní hodnoty rozdílů uspořádáme vzestupně a přiřadíme jim pořadí.
3. Spočítáme statistiky S_w^+ a S_w^-, které odpovídají součtu pořadí kladných (S_w^+) a záporných rozdílů (S_w^-). Jako finální hodnotu testové statistiky bereme minimum z S_w^+ a S_w^-. Nulovou hypotézu zamítáme, pokud hodnota testové statistiky menší nebo rovna tabelované kritické hodnotě (při dané hladině významnosti a počtu nenulových rozdílů).

nebo

3. Pro N > 30 lze využít asymptotické normality statistiky S_w^+

$$E(S_w+) = \frac{n(n+1)}{4} \qquad D(S_w+) = \frac{n(n+1)(2n+1)}{24} \qquad Z = \frac{S_w + -E(S_w+)}{\sqrt{D(S_w+)}} \approx N(0,1)$$

Pokud $|Z| \geq u_{1-\alpha/2}$ zamítáme nulovou hypotézu, že medián výběru je roven hodnotě *c*.

Znaménkový test

- Lze použít v situaci, kdy není splněn předpoklad symetrie rozdělení kolem mediánu.
- Testuje, zda je medián jednoho výběru roven hodnotě *c* (v případě párového designu je $x_{0.5}$ reprezentováno mediánem rozdílu hodnot)

$$H_0: x_{0.5}=c \quad \text{proti } H_1: x_{0.5} \neq c.$$

Postup:

1. Spočítáme rozdíly hodnot výběru s testovanou hodnotou mediánu.
2. Spočítáme statistiku S_z^+, která odpovídá počtu kladných rozdílů → test nevyužívá hodnot pořadí původních dat ale pouze informaci, zda se hodnota realizuje nad nebo pod mediánem → dochází ke snížení síly testu
3. Nulovou hypotézu zamítáme, pokud statistika S_z^+ realizuje v kritickém oboru hodnot $W=(0,k_1)U(k_2,n)$, kde *n* odpovídá počtu nenulový rozdílů a hodnoty k_1 a k_2 lze dohledat v matematických tabulkách.

nebo

3. Pro N > 20 lze využít asymptotické normality statistiky S_z^+.

$$E(S_z+) = \frac{n}{2} \qquad D(S_z+) = \frac{n}{4} \qquad Z = \frac{S_z + - E(S_z+)}{\sqrt{D(S_z+)}} \approx N(0,1)$$

Pokud $|Z| \geq u_{1-\alpha/2}$ zamítáme nulovou hypotézu, že medián výběru je roven hodnotě *c*.

Příklad 1: jednovýběrový test

- U 15 náhodně vybraných pacientů byla vyhodnocena doba, kterou museli strávit v čekárně, než byli sestrou pozváni do ordinace. Na 5% hladině významnosti testujte nulovou hypotézu, že medián čekací doby je roven půl hodině.

Příklad 1: jednovýběrový test – Wilcoxonův test

- U 15 náhodně vybraných pacientů byla vyhodnocena doba, kterou museli strávit v čekárně, než byli sestrou pozváni do ordinace. Na 5% hladině významnosti testujte nulovou hypotézu, že medián čekací doby je roven půl hodině.

Pacient č.	čekací doba (min)	medián	rozdíl	|rozdíl|	pořadí
1	1	30	-29	29	15
2	45	30	15	15	10
3	25	30	-5	5	3.5
4	15	30	-15	15	10
5	34	30	4	4	2
6	19	30	-11	11	8
7	31	30	1	1	1
8	25	30	-5	5	3.5
9	8	30	-22	22	14
10	12	30	-18	18	12
11	20	30	-10	10	6
12	15	30	-15	15	10
13	40	30	10	10	6
14	20	30	-10	10	6
15	10	30	-20	20	13

$S_w^+=19$
$S_w^-=101$

min $(S_w^+,S_w^-)=19$
Kritická hodnota $w_{15}(0,05)=25$
Hodnota testové statiky je menší než kritická hodnota → **zamítáme H_0**

Příklad 1: jednovýběrový test – Znaménkový test

- U 15 náhodně vybraných pacientů byla vyhodnocena doba, kterou museli strávit v čekárně, než byli sestrou pozváni do ordinace. Na 5% hladině významnosti testujte nulovou hypotézu, že medián čekací doby je roven půl hodině.

Pacient č.	čekací doba (min)	medián	rozdíl	Větší než medián?
1	1	30	-29	Ne
2	45	30	15	Ano
3	25	30	-5	Ne
4	15	30	-15	Ne
5	34	30	4	Ano
6	19	30	-11	Ne
7	31	30	1	Ano
8	25	30	-5	Ne
9	8	30	-22	Ne
10	12	30	-18	Ne
11	20	30	-10	Ne
12	15	30	-15	Ne
13	40	30	10	Ano
14	20	30	-10	Ne
15	10	30	-20	Ne

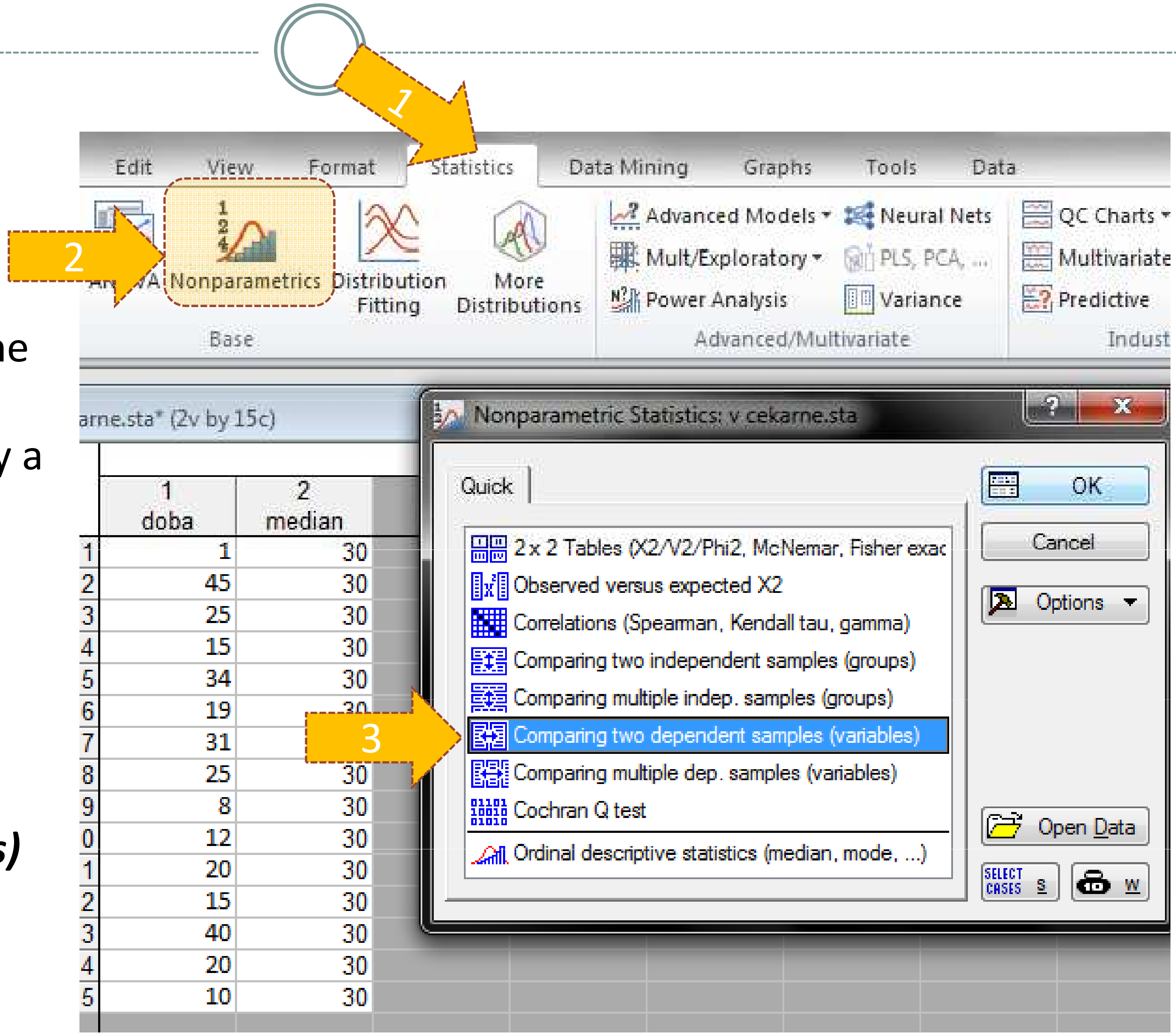
$S_z^+=4$

Kritický obor: W=(0,3)U(12,15)
Hodnota statistiky se realizuje mimo kritický obor hodnot → **nezamítáme H_0**

Příklad 1: Řešení v softwaru Statistica I

• Datový soubor si připravíme tak, že první proměnná obsahuje testované hodnoty a druhá proměnná medián, který chceme testovat

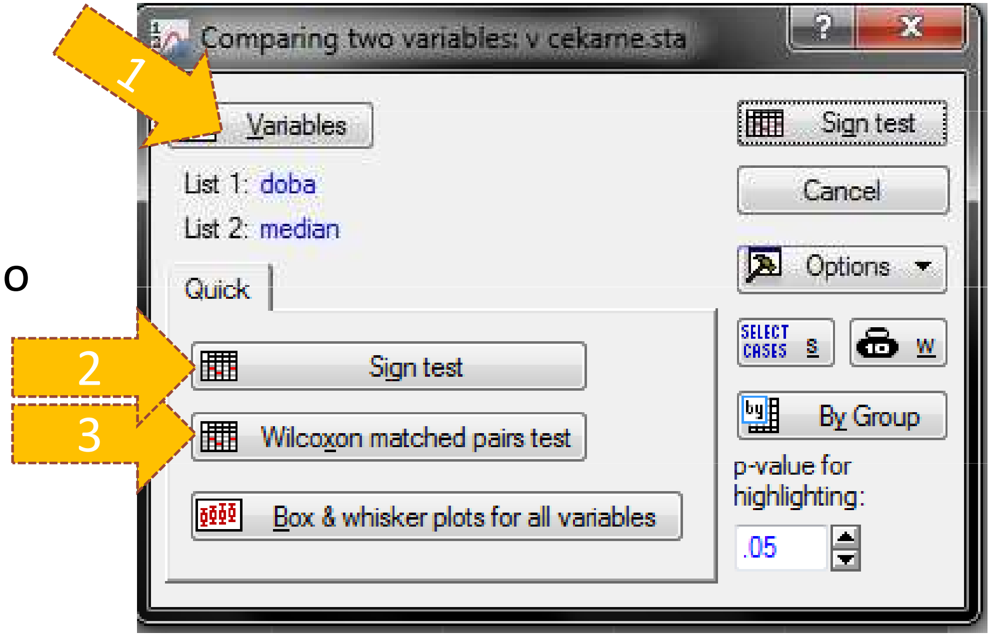
• V menu ***Statistics*** zvolíme ***Nonparametrics,*** vybereme ***Comparing two dependent samples (groups)***



Příklad 1: Řešení v softwaru Statistica II

• Vybereme proměnné, které chceme testovat (testovaný parametr, medián)

• Kliknutím na ***Sign test*** a následně ***Wilcoxon matched pair test*** získáme výsledky znaménkového a jednovýběrového Wilcoxonova testu



Příklad 1: Řešení v softwaru Statistica III

1) Výstup Wilcoxonova testu

Testová statistika: min (S_w^+,S_w^-)

Pair of Variables	Wilcoxon Matched Pairs Test (v cekarne.sta) Marked tests are significant at p <.05000 Valid N	T	Z	p-value
doba & median	15	19.00000	2.328644	0.019879

Statistika a p-hodnota pro asymptotickou variantu testu (používat pouze pro N > 30)

Počet nenulových rozdílů

2) Výstup znaménkového testu

Podíl hodnot menších než testovaný medián

Pair of Variables	Sign Test (v cekarne.sta) Marked tests are significant at p <.05000 No. of Non-ties	Percent v < V	Z	p-value
doba & median	15	73.33333	1.549193	0.121335

Statistika a p-hodnota pro asymptotickou variantu testu (používat pouze pro N > 20)

Počet nenulových rozdílů

2. Statistické testy o parametrech dvou výběrů

Nepárový Mannův-Whitneyův test
Párový Wilcoxonův a znaménkový test

Mannův-Whitneyův U test

- Neparametrická alternativa dvouvýběrového t-testu.
- Počítá s pořadím dat v souborech namísto s originálními daty.
- Předpoklad: rozdělení pravděpodobnosti veličiny ve skupinách se může lišit pouze posunutím.

Postup:

1. Stanovíme nulovou a alternativní hypotézu (F(x)=distribuční funkce):

   H_0: $F(x_1)=F(x_2)$

   H_1: $F(x_1)\neq F(x_2)$.

2. Čísla obou souborů jsou sloučena a je určeno jejich pořadí v tomto sloučeném souboru.
3. Pro oba výběry zvlášť je spočítán součet pořadí (T_1 a T_2).
4. Ze součtů pořadí ve skupinách je určena finální hodnota testové statistiky U.

$$U_1 = n_1 n_2 + \frac{n_1 - (n_1 + 1)}{2} - T_1 \qquad U_2 = n_1 n_2 + \frac{n_2 - (n_2 + 1)}{2} - T_2$$

$$U = \min(U_1, U_2)$$

5. Hodnotu testové statistiky U porovnáme s kritickou hodnotou testu, pokud je tato hodnota menší než kritická hodnota testu, zamítáme nulovou hypotézu shody distribučních funkcí obou skupin.

Mannův-Whitneyův U test – asymptotická varianta

5. Pro velká n_1 a n_2 (>30) lze využít asymptotické normality statistiky U.

$$E(U) = \frac{n_1 n_2}{2} \qquad D(U) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

6. Pro testování lze využít Z-statistiky:

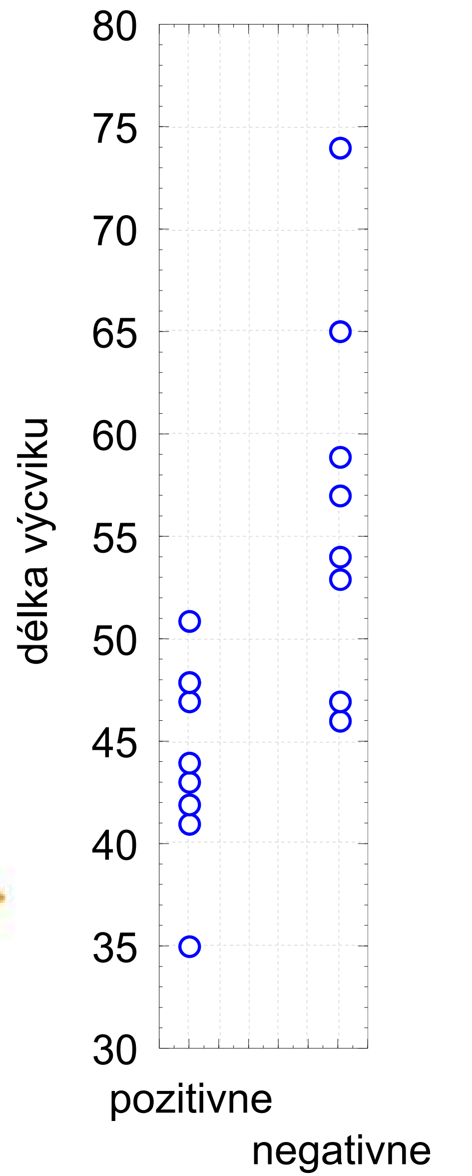
$$Z = \frac{U - E(U)}{\sqrt{D(U)}} \approx N(0,1)$$

7. Pokud $|Z| \geq u_{1-\alpha/2}$ zamítáme nulovou hypotézu o shodnosti distribučních funkcí.

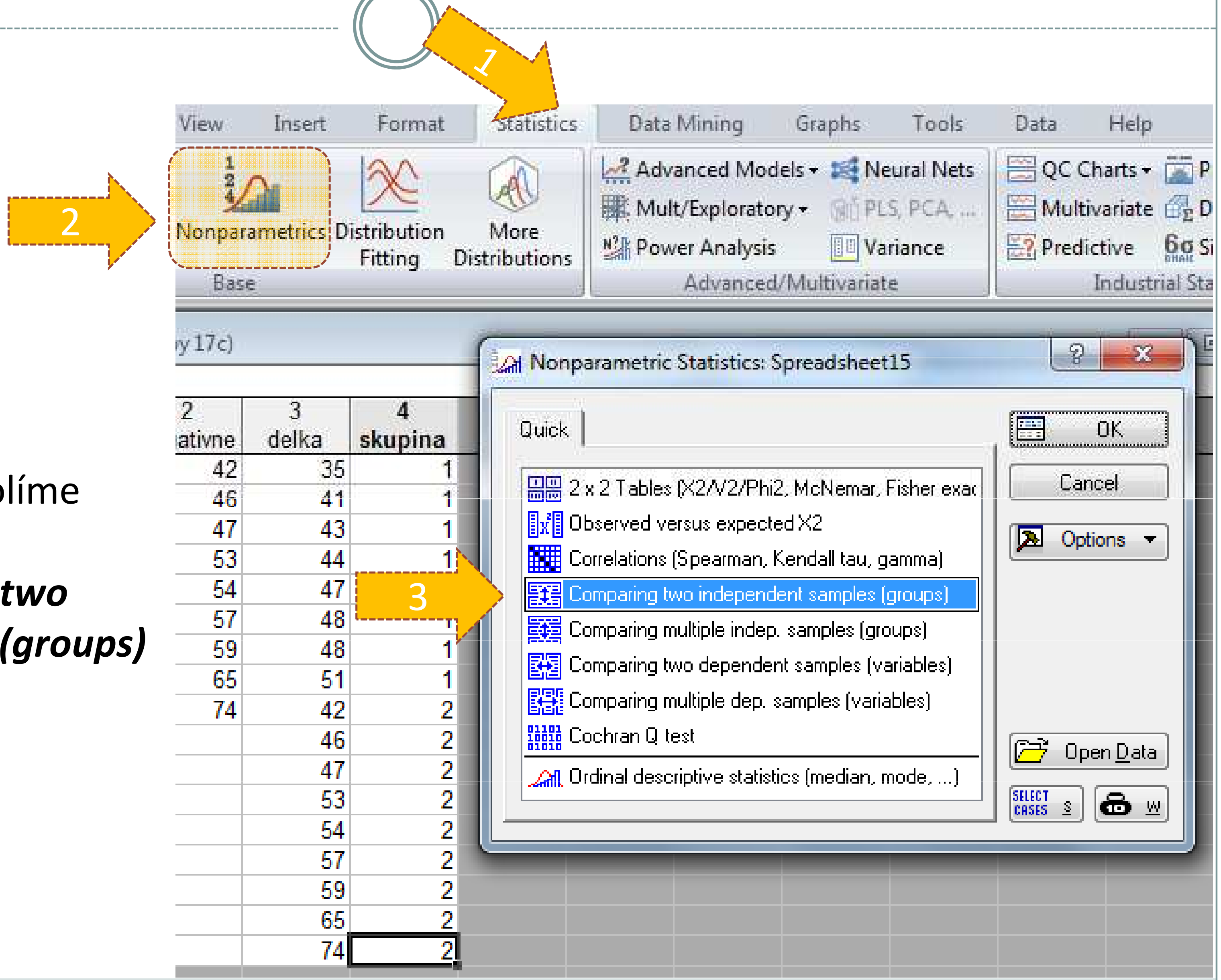
Příklad 2: Mannův – Whitneyův U test

- 17 štěňat bylo trénováno v chození na záchod metodou pozitivní motivace (pochvala, když jde na záchod venku) nebo negativní motivace (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 + kvůli nízkému počtu hodnot je vhodné použít neparametrický test.
- Je vytvořeno pořadí hodnot v kompletním souboru.
- Hodnota testové statistiky je určena ze součtu pořadí hodnot v jednotlivých skupinách.

- **Jak dopadne testování?**



Příklad 2: Řešení v softwaru Statistica I



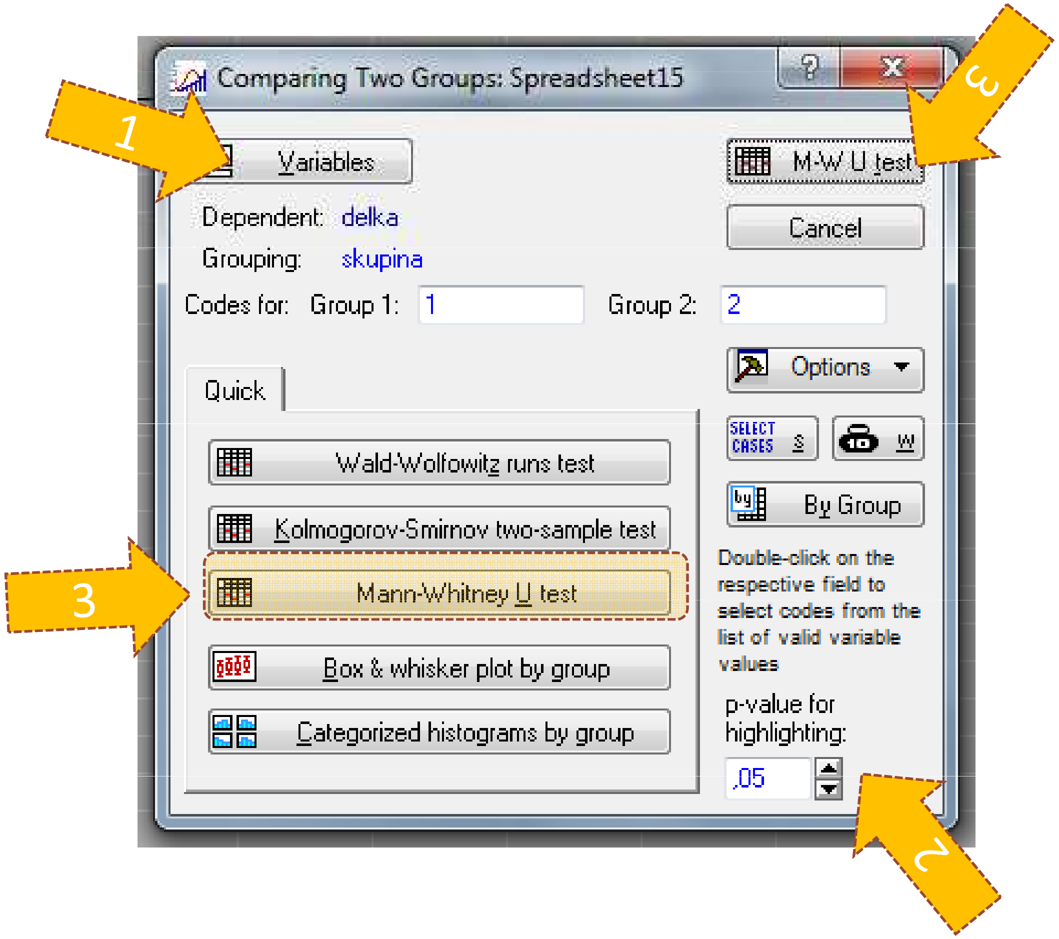
- V menu ***Statistics*** zvolíme ***Nonparametrics ,*** vybereme ***Comparing two independent samples (groups)***

Příklad 2: 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



Příklad 2: Řešení v softwaru Statistica III

Součet pořadí T_1

Součet pořadí T_2

Hodnota Z 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
(použít, jestliže rozsah výběru je menší než 30)

Párový Wilcoxonův a znaménkový test

- Vycházíme z rozdílů párových hodnot a přecházíme na design jednovýběrových testů
- Testuje, zda je medián diferencí (D) párových hodnot roven hodnotě *c*

$$H_0: D_{0.5}=c \quad \text{proti } H_1: D_{0.5}\neq c.$$

Wilcoxonův párový test

1. Spočítáme rozdíly **diferencí** výběru s testovanou hodnotou mediánu = c.
2. Absolutní hodnoty rozdílů uspořádáme vzestupně a přiřadíme jim pořadí.
3. Spočítáme statistiky S_W^+ a S_W^-, které odpovídají součtu pořadí kladných (S_W^+) a záporných rozdílů (S_W^-). Jako finální hodnotu testové statistiky bereme minimum z S_W^+ a S_W^-. Nulovou hypotézu zamítáme, pokud hodnota testové statistiky menší nebo rovna tabelované kritické hodnotě (při dané hladině významnosti a počtu nenulových rozdílů).

Znaménkový párový test

1. Spočítáme rozdíly **diferencí** výběru s testovanou hodnotou mediánu = c.
2. Spočítáme statistiku S_z^+, která odpovídá počtu kladných rozdílů → test nevyužívá hodnot pořadí původních dat ale pouze informaci, zda se hodnota realizuje nad nebo pod mediánem → dochází ke snížení síly testu
3. Nulovou hypotézu zamítáme, pokud statistika S_z^+ realizuje v kritickém oboru hodnot $W=(0,k_1)U(k_2,n)$, kde *n* odpovídá počtu nenulový rozdílů a hodnoty k_1 a k_2 lze dohledat v matematických tabulkách.

Příklad 3: Wilcoxonův párový test

- Na 5% hladině významnosti testujte, zda se liší hladina krevního parametru před a po podání léku. $H_0: D_{0.5}=0$ proti $H_1: D_{0.5} \neq 0$.

pacient	Před podáním léku	Po podání léku	Diference (D)	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

S_w^+ …..součet pořadí přes kladné hodnoty rozdílů = 51

S_w^- …..součet pořadí přes záporné hodnoty rozdílů = 4

$\min(S_w^+;S_w^-) = 4$

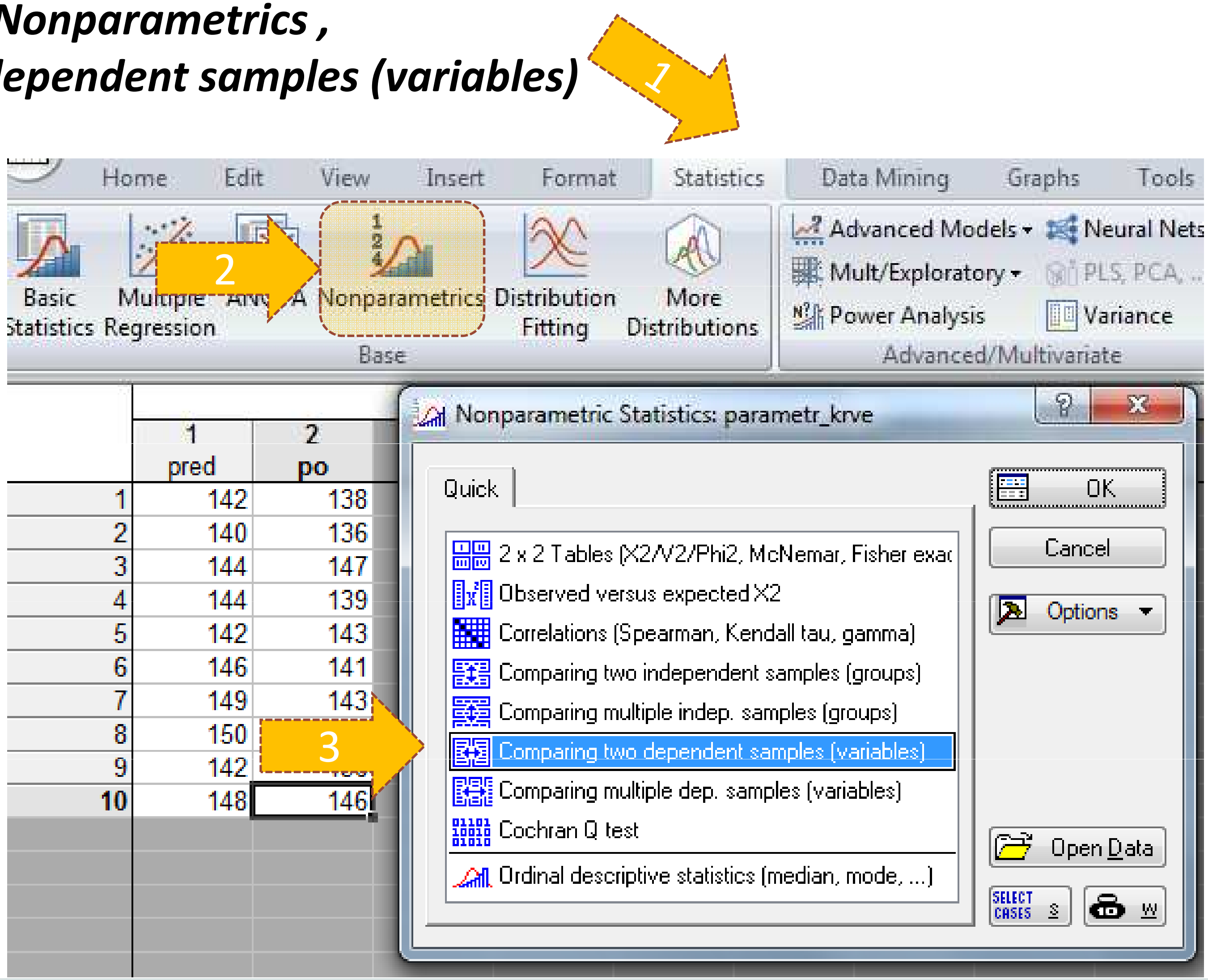
počet párů = n = 10

$w_n(\alpha) = w_{10}(0,05) = 8$

Hodnota testové statiky je menší než kritická hodnota → **zamítáme H_0**

Příklad 3: 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	[illegible]
9	142	[illegible]
10	148	146

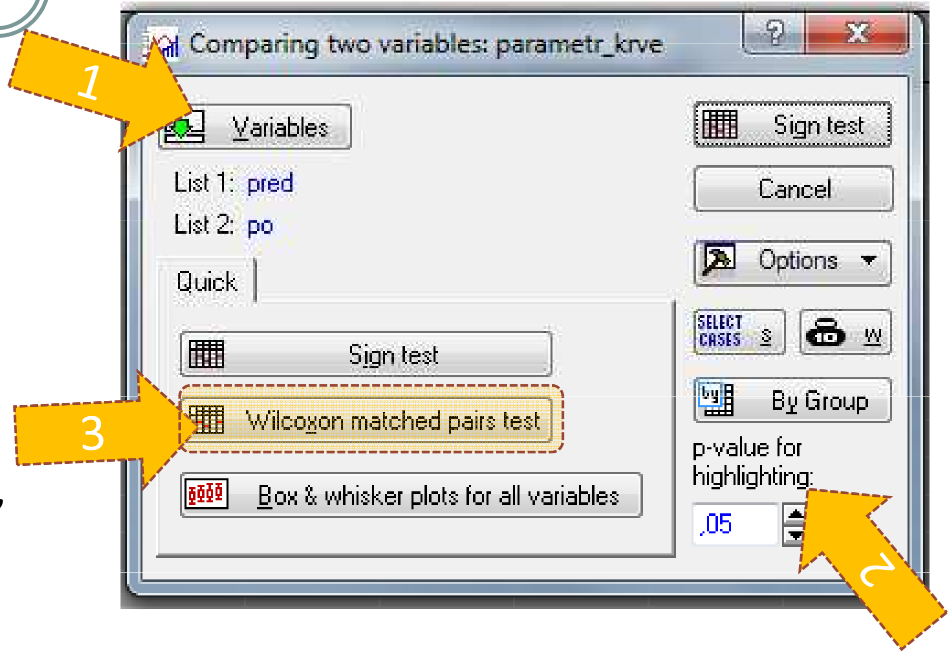
Pozn.: Pokud bychom chtěli testovat c různé od 0, musíme vstupní data uspořádat tak, že první proměnná bude obsahovat diference párových hodnot a druhá proměnná testovanou hodnotu mediánu c.

Příklad 3: 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

Pair of Variables	Wilcoxon Matched Pairs Test (parametr_krve) Marked tests are significant at p <,05000 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>

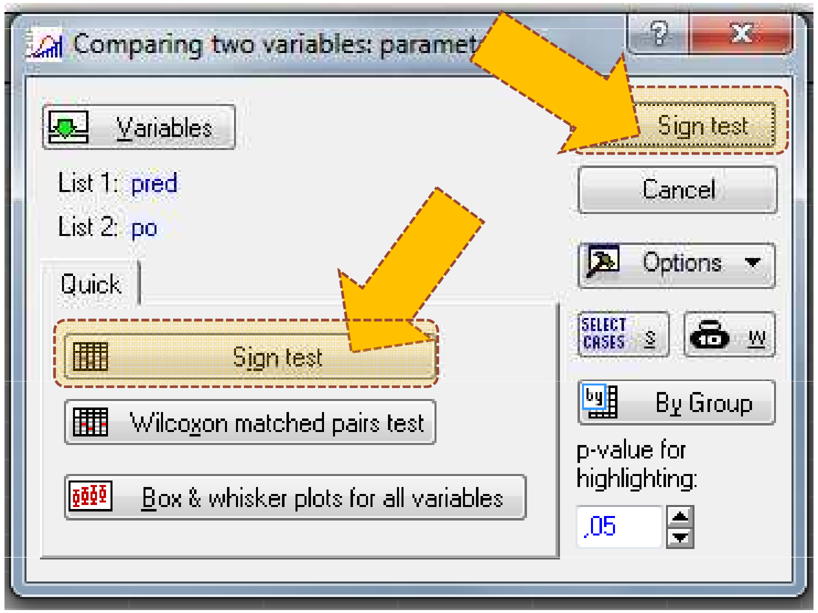
Hodnota asymptotické testové statistiky

Příklad 3: Řešení v softwaru Statistica III

- 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%.

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

3. Statistické testy o parametrech tří a více výběrů

Kruskalův-Wallisův test
Mediánový test

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

Kruskalův-Wallisův test I

- Neparametrická alternativa analýzy rozptylu (ANOVA)
- Zobecnění Mannova-Whitneyova U testu pro **více než dvě** srovnávané skupiny.
- Počítá s pořadím dat v souborech namísto s originálními daty.
- Nulová hypotéza předpokládá stejné rozdělení pravděpodobnosti veličiny ve více skupinách.
- Předpoklad: rozdělení pravděpodobnosti veličiny ve skupinách se může lišit pouze posunutím.

Kruskalův-Wallisův test II

Postup:

1. Stanovíme nulovou a alternativní hypotézu pro *k* skupin (F(x)=distribuční funkce):

   H_0: $F(x_1) = F(x_2) = \ldots = F(x_k)$

   H_1: alespoň jedna $F(x_i)$ se liší od ostatních
2. Čísla obou souborů jsou sloučena a je určeno jejich pořadí v tomto sloučeném souboru.
3. Pro všechny výběry zvlášť je spočítán součet pořadí (T_1, T_2, … T_k).
4. Ze součtů pořadí ve skupinách je určena finální hodnota testové statistiky Q:

$$Q = \frac{12}{n(n+1)} \sum_{j=1}^{k} \frac{T_j^2}{n_j} - 3(n+1)$$

5. Pokud je $Q \geq \chi^2$ (k-1), zamítáme nulovou hypotézu. Pro malé velikosti vzorků určujeme kritický obor z tabulek pro Kruskalův-Wallisův test.
6. V případě zamítnutí nulové hypotézy pomocí ***metod mnohonásobného porovnávání*** určíme, které dvojice skupin se liší.

Mediánový test

- Nepočítá s pořadím hodnot dat v souboru ale pouze z informací, zda je hodnota větší nebo menší než medián celkového souboru → dochází k další redukci informace z dat → má nižší sílu testu než Mannův-Whitneyův nebo Kruskalův-Wallisův test.

Postup:

1. Čísla obou souborů jsou sloučena a je definován medián těchto hodnot.
2. Určíme počet hodnot P_j (pro každou skupinu zvlášť), které jsou větší nebo rovny tomuto mediánu.
3. Testová statistika má tvar:

$$Q_M = 4\sum_{j=1}^{k}\frac{P_j^2}{n_j} - n$$

4. Pokud je $Q_M \geq \chi^2(k-1)$, zamítáme nulovou hypotézu.
5. V případě zamítnutí nulové hypotézy, se ptáme, které dvojice náhodných výběrů se liší, k tomuto účelu je vhodné použít ***metody mnohonásobného porovnávání***

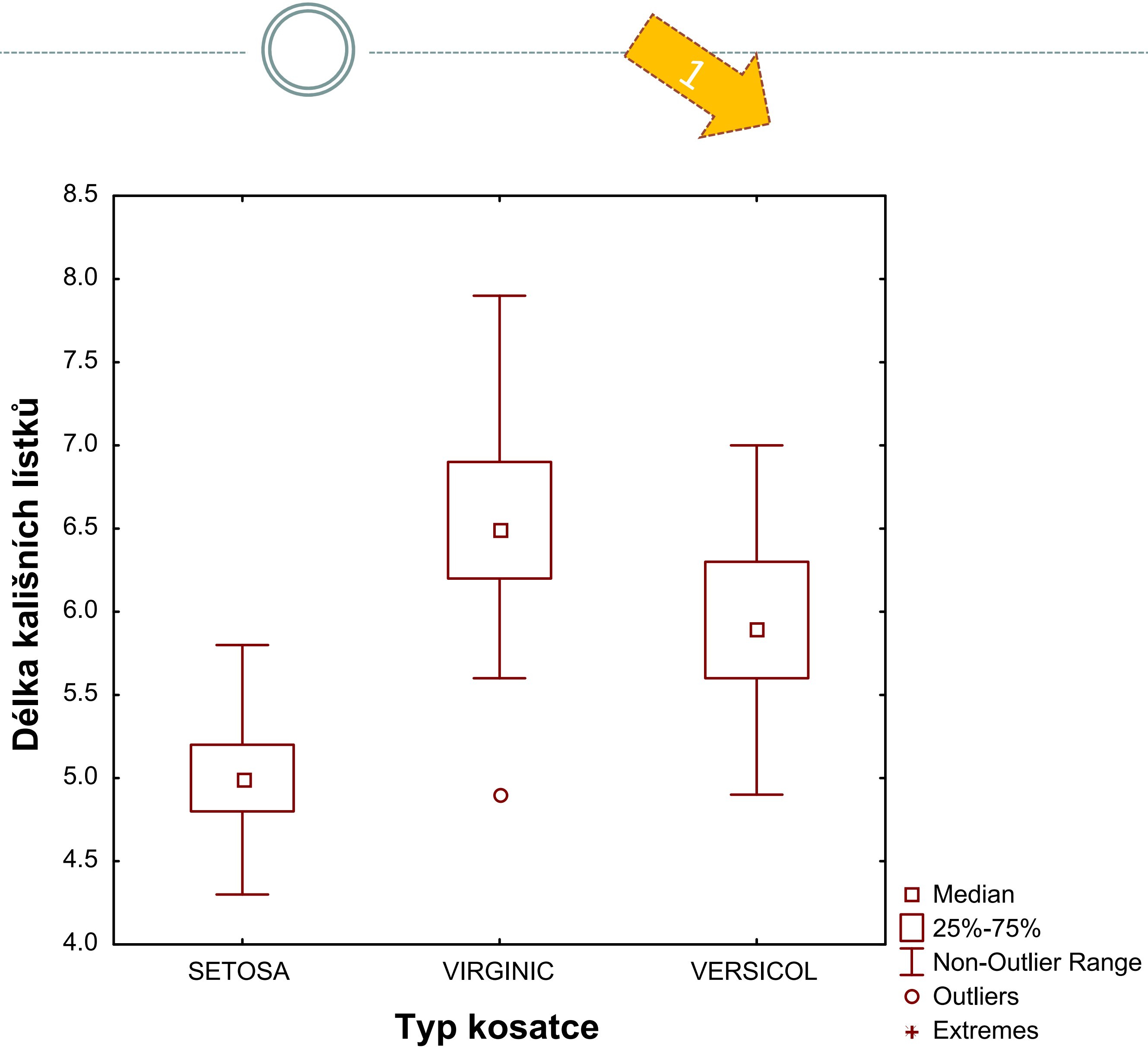
Příklad 4: Kruskalův- Wallisův test

- 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

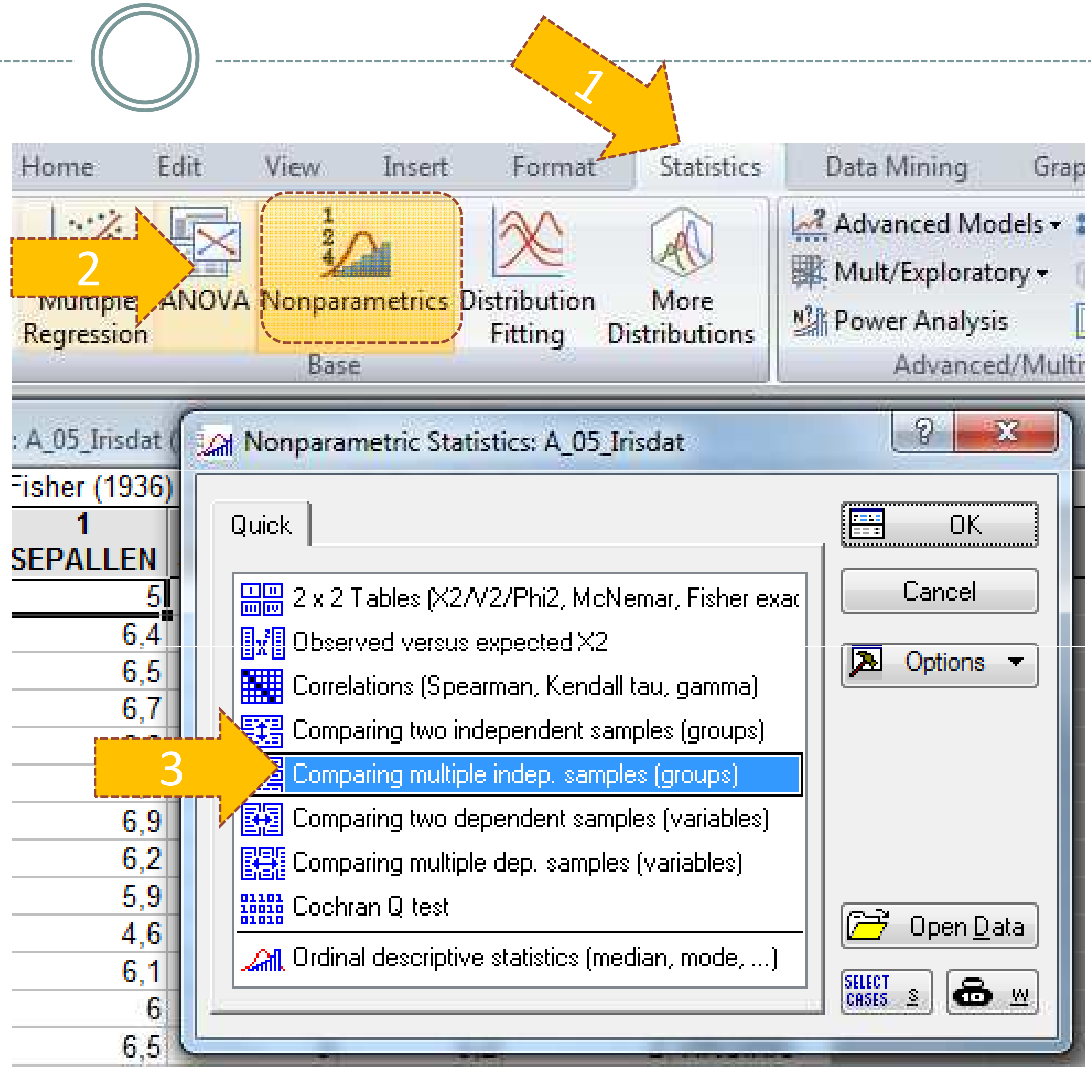
Příklad 4: Řešení v softwaru Statistica I

• nejprve se pomocí grafu podíváme na rozložení dat v rámci srovnávaných skupin



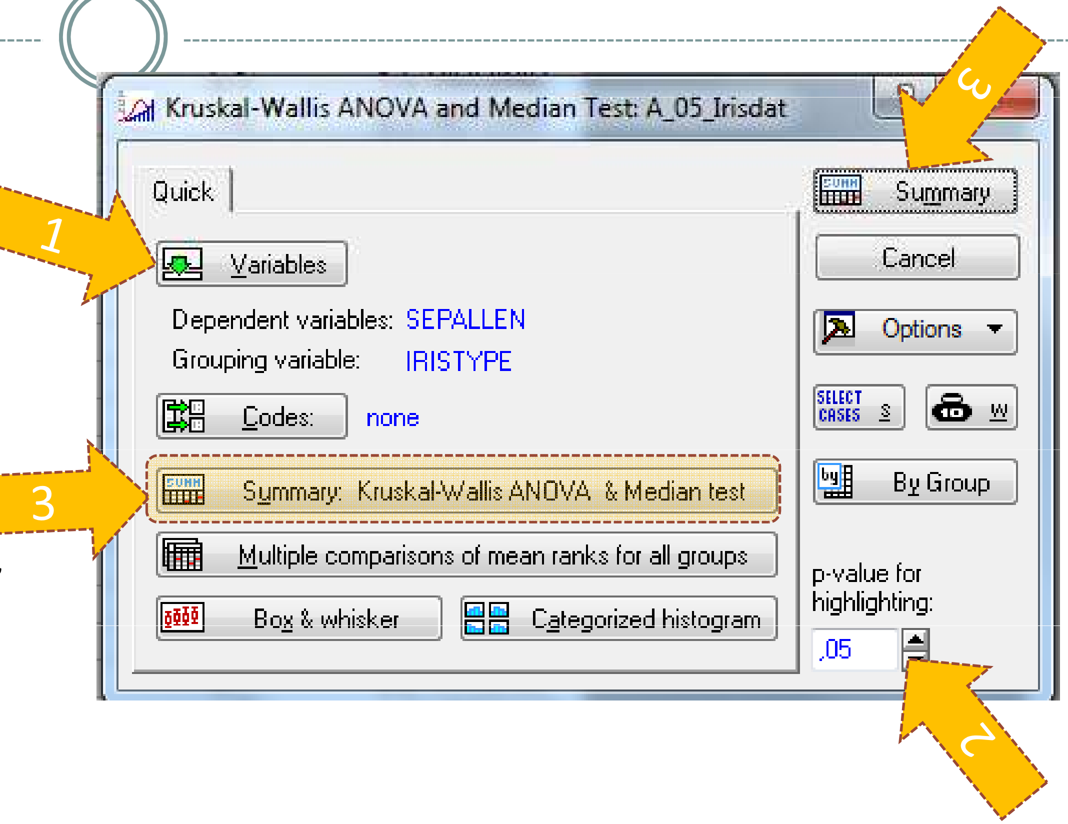
Příklad 4: Řešení v softwaru Statistica II

• V menu ***Statistics*** zvolíme ***Nonparametrics ,*** vybereme ***Comparing multiple indep. samples (groups)***



Příklad 4: 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

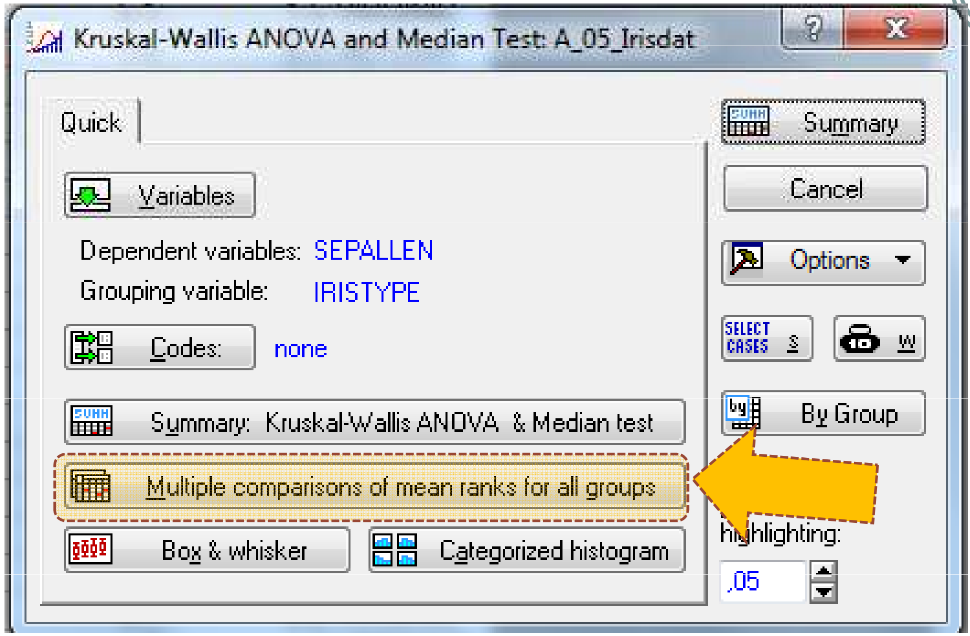
Depend.: SEPALLEN	Code	Valid N	Sum of Ranks	Mean Rank
SETOSA	1	50	1482,000	29,6400
VERSICOL	2	50	4132,500	82,6500
VIRGINIC	3	50	5710,500	114,2100

p-hodnota

Hodnota testové statistiky

Pokud p<0,05, musíme provést **test mnohonásobného porovnání**.

Příklad 4: Řešení v softwaru Statistica III



<u>Testy mnohonásobného porovnávání</u>

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

Depend.: SEPALLEN	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		
	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>