Cvičení 9: Neparametrické úlohy o mediánech

Úkol 1.: Párový znaménkový test a párový Wilcoxonův test

Při zjišťování kvality jedné složky půdy se používají dvě metody označené A a B. Výsledky:

Vzorek	1	2	3	4	5	6	7	8	9	10	11	12
A	0,275	0,312	0,284	0,3	0,365	0,298	0,312	0,315	0,242	0,321	0,335	0,307
B	0,28	0,312	0,288	0,298	0,361	0,307	0,319	0,315	0,242	0,323	0,341	0,315

Na hladině významnosti 0,05 testujte pomocí párového znaménkového testu a poté pomocí párového Wilcoxonova testu hypotézu, že metody A a B dávají stejné výsledky.

Návod:

Načteme datový soubor kvalita_pudy.sta. Proměnná A obsahuje výsledky metody A, proměnná B výsledky metody B.

Nejprve budeme testovat nulovou hypotézu pomocí párového znaménkového testu.

Statistiky –Neparametrická statistika – Porovnání dvou závislých vzorků (proměnné) – OK – 1. seznam proměnných A, 2. seznam proměnných B – OK – Znaménkový test.

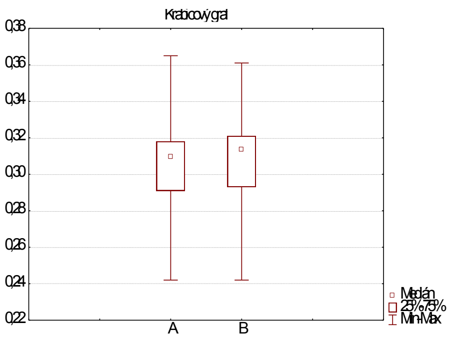
Znaménkový test (kvalita_pudy.s
Označené testy jsou významné n

Dvojice prom	Počet různý	proce v < V	Z	Úroveň
A & B	9	77,77	1,333	0,182

Komentář: Vidíme, že nenulových hodnot n = 9. Z nich záporných je 77,7%, tj.7. Kladných je tedy 9 – 7 = 2, což je hodnota testové statistiky S_Z^+. Asymptotická testová statistika U_0 (zde označená jako Z) se realizuje hodnotou 1,3. Odpovídající asymptotická p-hodnota je 0,18422, tedy na asymptotické hladině významnosti 0,05 nezamítáme hypotézu, že obě metody dávají stejné výsledky.

Upozornění: V tomto případě není splněna podmínka pro využití asymptotické normality statistiky S_Z^+, tj. n > 20. Je tedy vhodnější najít v tabulkách kritické hodnoty pro znaménkový test (viz skripta Základní statistické metody, tabulka na straně 156). Pro n = 9 a α = 0,05 jsou kritické hodnoty $k_1 = 1$, $k_2 = 8$. Protože kritický obor $W = \{1\} \cup \{\ \}$ neobsahuje hodnotu 2, nezamítáme H_0 na hladině významnosti 0,05. Dostáváme stejný výsledek při použití asymptotického testu

Nyní graficky znázorníme výsledky obou metod: Návrat do Porovnání 2 proměnných - Krabicový graf všech proměnných – OK – A, B – OK.



Komentář: Z krabicových diagramů je vidět, že obě metody se poněkud liší v úrovni, ale neliší se ve variabilitě.

Dále provedeme Wilcoxonův párový test.
Návrat do Porovnání 2 proměnných – Wilcoxonův párový test.

	Wilcoxonův párový test (kvalita_ Označené testy jsou významné n			
Dvojice prom	Poče platný	T	Z	p-hod
A & B	9	5,000	2,073	0,038

Komentář: Výstupní tabulka poskytne hodnotu testové statistiky S_W^+ (zde označena T), hodnotu asymptotické testové statistiky U_0 a p-hodnotu pro U_0. V tomto případě je p-hodnota 0,038153, tedy nulová hypotéza se zamítá na asymptotické hladině významnosti 0,05. Ze srovnání asymptotických p-hodnot pro znaménkový test a pro Wilcoxonův test plyne, že Wilcoxonův test je silnější.
Upozornění: V tomto případě není splněna podmínka pro využití asymptotické normality statistiky S_W^+, tj. $n \geq 30$. Je tedy vhodnější najít v tabulkách kritickou hodnotu pro Wilcoxonův párový test (viz skripta Základní statistické metody, tabulka na straně 157). Pro $n = 9$ a $\alpha = 0,05$ je kritická hodnota rovna 5. Protože kritický obor $W \leq 5$ obsahuje hodnotu 5, zamítáme H_0 na hladině významnosti 0,05. To souhlasí s výsledkem asymptotického testu.

Úkol 2.: **Znaménkový test a jednovýběrový Wilcoxonův test**
Vyráběné ocelové tyče mají kolísavou délku s předpokládanou hodnotou mediánu 10 m. Náhodný výběr 10 tyčí poskytl tyto výsledky:
9,83 10,10 9,72 9,91 10,04 9,95 9,82 9,73 9,81 9,90
Na hladině významnosti 0,05 testujte hypotézu, že předpoklad o mediánu délky tyčí je oprávněný.

Návod:
Načteme datový soubor ocelove_tyce.sta. Proměnná X obsahuje naměřené hodnoty a proměnná Y obsahuje konstantu 10. Provedení znaménkového a Wilcoxonova testu je nyní stejné jako v předešlém případě.

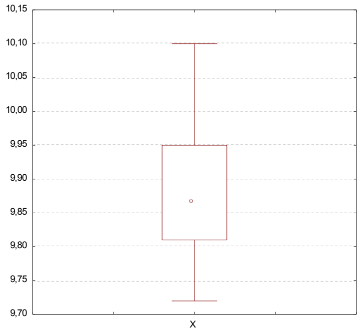
	Znaménkový test (ocelove_tyce.s Označené testy jsou významné n			
Dvojice prom	Poče různý	proce v < V	Z	Úrove
X & Y	1	80,00	1,581	0,113

	Wilcoxonův párový test (ocelove Označené testy jsou významné n			
Dvojice prom	Poče platný	T	Z	Úrove
X & Y	1	5,500	2,242	0,024

Komentář: Znaménkový test poskytl asymptotickou p-hodnotu 0,113846, tedy nulová hypotéza se nezamítá na hladině významnosti 0,05. Wilcoxonův test dává asymptotickou p-hodnotu 0,024933, tedy nulová hypotéza se zamítá na asymptotické hladině významnosti 0,05. Podobně jak v úkolu 1 by bylo vhodnější najít kritické hodnoty v tabulkách. V případě

znaménkového testu jsou kritické hodnoty pro n = 10 a α = 0,05 rovny 1 a 9, testová statistika $S_Z^+ = 2$. Protože S_Z^+ nepatří do kritického oboru $W = \langle 1 \rangle \cup \langle 9 \rangle$, nelze nulovou hypotézu zamítnout na hladině významnosti 0,05, což je v souladu s výsledkem asymptotického testu.
V případě Wilcoxonova testu je kritická hodnota pro n = 10 α = 0,05 rovna 8. Protože kritický obor $W = \langle 8 \rangle$ obsahuje hodnotu 5,5, zamítáme H_0 na hladině významnosti 0,05. I zde tedy existuje soulad mezi výsledkem přesného a asymptotického testu.

Podobně jako v úkolu 1. znázorníme výsledky měření pomocí krabicového diagramu:



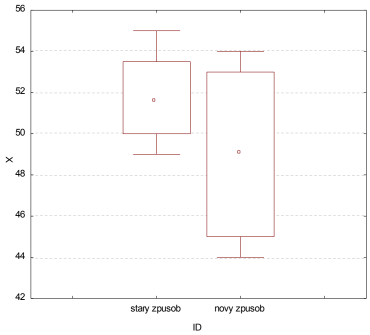
Úkol 3.: Dvouvýběrový Wilcoxonův test, dvouvýběrový Kolmogorovův - Smirnovův test
Bylo vybráno 10 polí stejné kvality. Na čtyřech z nich se zkoušel nový způsob hnojení, zbylých šest bylo ošetřeno starým způsobem. Pole byla oseta pšenicí a sledoval se její hektarový výnos. Je třeba testovat na hladině významnosti 0,05, zda nový způsob hnojení má týž vliv na průměrné hektarové výnosy pšenice jako starý způsob hnojení.
hektarové výnosy při novém způsobu: 51 52 49 55
hektarové výnosy při starém způsobu: 45 54 48 44 53 50

Návod:
Načteme datový soubor hnojeni_poli.sta. Proměnná X udává výnosy pšenice při obou způsobech hnojení a proměnná ID nabývá hodnoty 1 pro starý způsob hnojení, hodnoty 2 pro nový způsob hnojení.
Dvouvýběrový Wilcoxonův test:
Statistiky – Neparametrická statistika – Porovnání dvou nezávislých vzorků (skupiny) – OK - Seznam závislých proměnných X, Grupovací proměnná ID - OK - Mann – Whitneyův U test.

Mann-Whitneyův U test (hnojeni_poli)
Dle proměn. ID
Označené testy jsou významné na hladině p <,05000

Proměn	Sčt p st. zp	Sčt p n. zp	U	Z	p-hod	Z upra	p-hod	N pla st. zp	N pla n. zp	2*1st přesn
X	27,00	28,00	7,000	0,959	0,337	0,959	0,337	4	6	0,352

Komentář: Ve výstupní tabulce jsou součty pořadí T_1, T_2, hodnota testové statistiky $\min(U_1, U_2)$ označená U, hodnota asymptotické testové statistiky U_0 (označená Z), asymptotická p-hodnota pro U_0 a přesná p-hodnota (ozn. 2*1 str. přesné p – ta se používá pro rozsahy výběrů pod 30). V našem případě přesná p-hodnota = 0,352381, tedy H_0 nezamítáme na hladině významnosti 0,05.
Výpočet je vhodné doplnit krabicovým diagramem.



Komentář: Je zřejmé, že výnosy při novém způsobu hnojení jsou vesměs nižší než při starém způsobu a také vykazují mnohem větší variabilitu.

Dvouvýběrový Kolmogorovův – Smirnovův test (viz skripta Základní statistické metody, věta 9.4.3.1.) poskytne tabulku:

	Kolmogorov-Smirnovův test (hnojeni_poli.sta) Dle proměn. ID Označené testy jsou významné na hladině p <,05000							
Proměi	Max zá rozdí	Max kl rozdí	p-hod	Průmě stary zpu	Průmě novy zpu	Sm.odc stary zpu	Sm.odc novy zpu	N platn stary zpu
X	-0,083	0,500	p > .	51,75	49,00	2,500	4,098	4

Komentář: Dostaneme maximální záporný a maximální kladný rozdíl mezi hodnotami obou výběrových distribučních funkcí, dolní omezení pro p-hodnotu (p > 0,1), průměry, směrodatné odchylky a rozsahy obou výběrů.
Protože p > 0,05, nulovou hypotézu nezamítáme na hladině významnosti 0,05.

Úkol 4.: **Kruskalův – Wallisův test a mediánový test**
Voda po holení jisté značky se prodává ve čtyřech různých lahvičkách stejného obsahu. Údaje o počtu prodaných lahviček za týden v různých obchodech:
1. typ: 50 35 43 30 62 52 43 57 33 70 64 58 53 65 39
2. typ: 31 37 59 67 44 49 54 62 34 42 40
3. typ: 27 19 32 20 18 23
4. typ: 35 39 37 38 28 33.
Posuďte na 5% hladině významnosti, zda typ lahvičky ovlivňuje úroveň prodeje.

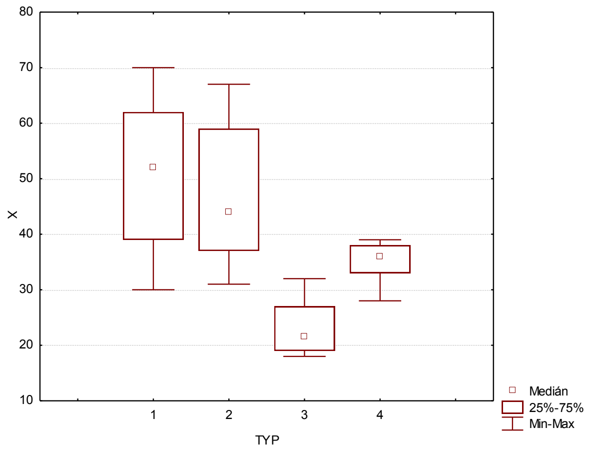
Návod:
Načteme datový soubor voda_po_holeni.sta. Proměnná X udává počet prodaných lahviček a proměnná TYP udává typ lahvičky. Úloha vede na Kruskalův – Wallisův test nebo mediánový test (viz skripta Základní statistické metody, věta 9.5.1.1. resp. věta 9.5.3.1.): Statistiky – Neparametrická statistika – Porovnání více nezávislých vzorků (skupiny) – OK – Proměnné – Seznam závislých proměnných X, Grupovací proměnná TYP – OK – Shrnutí: Kruskal-Wallis ANOVA & Mediánový test. Ve dvou výstupních tabulkách se objeví výsledky K-W testu a mediánového testu.

Závis X	Kruskal-Wallisova ANOVA X založ. na pořadí Nezávislá (grupovací) proměnná : TYP Kruskal-Wallisův test: H (3, N= 38) =18		
	Ko	Poče platný	Souč pořad
1	1	1	379,0
2	2	1	257,0
3	3	6	24,00
4	4	6	81,00

Závislá: X	Mediánový test, celk. medián = 39,500 X Nezávislá (grupovací) proměnná : TYP Chi-Kvadr. = 17,53939 sv = 3 p = ,0005				
	1	2	3	4	Celke
<= Medián: po	4,000	3,000	6,000	6,000	19,00
oček	7,500	5,500	3,000	3,000	
poz.-	-3,50	-2,50	3,000	3,000	
> Medián: po	11,00	8,000	0,000	0,000	19,00
oček	7,500	5,500	3,000	3,000	
poz.-	3,500	2,500	-3,00	-3,00	
Celkem:	15,00	11,00	6,000	6,000	38,00

Komentář: Oba testy zamítají hypotézu o shodě mediánů v daných čtyřech skupinách. Testová statistika K-W testu je 18,802, počet stupňů volnosti 3, odpovídající asymptotická p-hodnota 0,0003. Testová statistika mediánového testu je 17,539, počet stupňů volnosti 3, odpovídající asymptotická p-hodnota 0,0005. K-W test je poněkud silnější (p-hodnota = 0,0003, zatímco p-hodnota pro mediánový test je 0,0005).

Grafické znázornění výsledků: návrat do Shrnutí: Kruskal-Wallis ANOVA & Mediánový test – Krabicový graf – Vyberte proměnnou: X – OK – Typ krabicového grafu: Medián/Kvartily/Rozpětí – OK.



Komentář: Je vidět, že úroveň prodeje pro 1. typ je nevyšší, zatímco pro 3. typ nejnižší. Pro 1. a 2. typ je variabilita prodeje značná, pro 3. a 4. typ naopak malá.

Vzhledem k tomu, že jsme zamítli nulovou hypotézu o shodě mediánů na asymptotické hladině významnosti 0,05, provedeme metoda mnohonásobného porovnávání:Vícenás. porovnání průměrného pořadí pro vš. sk.

Závis X	Vícenásobné porovnání p h X (hodnota (obous Nezávislá (grupovací) proměnná : TYP Kruskal-Wallisův test: H (3, N= 38) =18,			
	1 R:25,:	2 R:23,:	3 R:4,0(	4 R:13,!
1		1,000	0,000	0,170
2	1,000		0,003	0,481
3	0,000	0,003		0,832
4	0,170	0,481	0,832	

Komentář: Tabulka obsahuje p-hodnoty pro test hypotézy, že l-tý a k-tý výběr pocházejí z téhož rozložení. Vidíme, že na hladině významnosti 0,05 zamítáme nulovou hypotézu pro 1. a 3. typ lahvičky a pro 2. a 3. typ lahvičky.

Příklady k samostatnému řešení

Příklad 1.: U osmi osob byl změřen systolický krevní tlak před pokusem a po něm.

č. osoby	1	2	3	4	5	6	7	8
tlak před	130	185	162	136	147	181	128	139
tlak po	139	190	175	135	155	175	158	149

Na hladině významnosti 0,05 testujte hypotézu, že pokus neovlivní systolický krevní tlak.

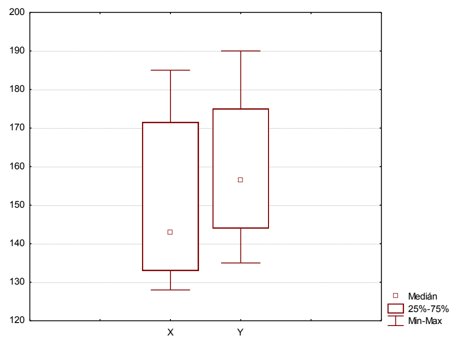
Řešení:

Stejně jako v úkolu 1 provedeme párový znaménkový a párový Wilcoxonův test. Načteme soubor tlak.sta. Proměnná X obsahuje hodnoty tlaku před pokusem, proměnná Y po pokusu. Výstupní tabulka:

Dvojice prom	Znaménkový test (tlak.sta) Označené testy jsou významné n			
	Poče různý	proce v < V	Z	Úroveň
X&Y	8	75,00	1,060	0,288

Komentář: Jelikož p-hodnota = 0,288844 > 0,05, nelze nulovou hypotézu zamítnout na hladině významnosti 0,05. Vzhledem k malému rozsahu výběru je však vhodnější najít v tabulkách kritické hodnoty pro znaménkový test (viz skripta Základní statistické metody, tabulka na straně 156). Pro n = 8 a $\alpha = 0{,}05$ jsou kritické hodnoty $k_1 = 0$, $k_2 = 8$. Hodnotu testové statistiky S_Z^+ získáme jako 75% z 8, což je 6. Protože kritický obor neobsahuje hodnotu 6, nezamítáme H_0 na hladině významnosti 0,05. Dostáváme stejný výsledek při použití asymptotického testu

Grafické znázornění výsledků pomocí krabicového diagramu:



Komentář: Úroveň tlaku před pokusem byla poněkud nižší než po pokusu, variabilita je jen nepatrně odlišná.

Výstupní tabulka Wilcoxonova testu:

	Wilcoxonův párový test (tlak.sta) Označené testy jsou významné n			
Dvojice prom	Poče platný	T	Z	Úroveň
X&Y	8	4,000	1,960	0,049

Vidíme, že asymptotická p-hodnota = 0,049951, nulová hypotéza se tedy zamítá na asymptotické hladině významnosti 0,05. Rozsah souboru je pouze 8, není splněna podmínka dobré aproximace standardizovaným normálním rozložením (n > 30). Proto zjistíme ve skriptech Základní statistické netody v tabulce na str. 157 kritickou hodnotu pro n = 8 a α = 0,05. Kritická hodnota je rovna 3, hodnota testové statistiky (ve výstupní tabulce označena T) je 4 > 3, tedy nulovou hypotézu nezamítáme na hladině významnosti 0,05, což je v souladu s výsledkem znaménkového testu.

Příklad 2.: Majitel obchodu chtěl zjistit, zda velikost nákupů (v dolarech) placených kreditními kartami Master/EuroCard a Visa jsou přibližně stejné. Náhodně vybral 7 nákupů placených Master/EuroCard: 42 77 46 73 78 33 37 a 9 placených Visou: 39 10 119 68 76 126 53 79 102. Lze na hladině významnosti 0,05 tvrdit, že nákupů placených těmito dvěma typy karet se shodují?

Řešení:
Stejně jako úkolu 3 použijeme dvouvýběrový Wilcoxonův test a Kolmogorovův - Smirnovův test.
Načteme datový soubor kreditni_karty.sta. Proměnná X obsahuje hodnoty nákupů, proměnná ID má hodnotu 1 pro kartu Master/EuroCard a hodnotu 2 pro kartu Visa.
Výstupní tabulka pro dvouvýběrový Wilcoxonův test:

	Mann-Whitneyův U test (kreditni_karty.sta) Dle proměn. ID Označené testy jsou významné na hladině p <,05000								
Promě	Sčt po M/E Ca	Sčt po Visa	U	Z	Úroveň	Z upravené	Úroveň	N plat M/E Ca	N pla Visa	2*1st přesné
X	48,00	88,00	20,00	-1,21	0,223	-1,21	0,223	7	9	0,252

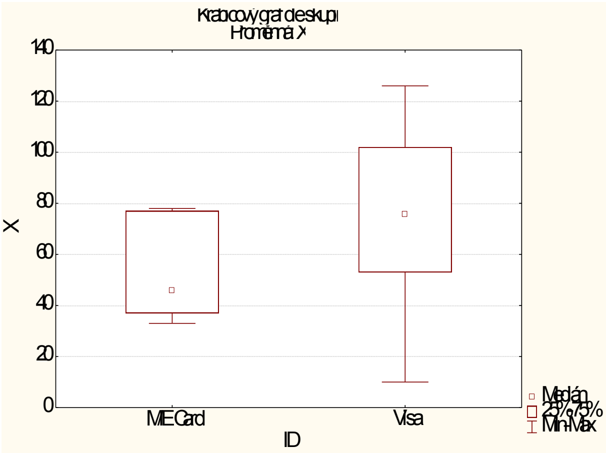
Komentář: Zajímá nás především přesná p-hodnota (ozn. 2*1 sided exact p – ta se používá pro rozsahy výběrů pod 30). V našem případě přesná p-hodnota = 0,252273, tedy H_0 nezamítáme na hladině významnosti 0,05.

Výstupní tabulka pro Kolmogorovův – Smirnovův test:

	Kolmogorov-Smirnovův test (kreditni_karty.sta) Dle proměn. ID Označené testy jsou významné na hladině p <,05000								
Promě	Max zá rozdíl	Max kl rozdíl	Úroveň	Průmě M/E Ca	Prům Visa	Sm.odc M/E Ca	Sm.od Visa	N plat M/E Ca	N pla Visa
X	-0,444	0,111	p > .	55,14	74,66	19,97	37,64	7	9

Komentář: Dostaneme maximální záporný a maximální kladný rozdíl mezi hodnotami obou výběrových distribučních funkcí, dolní omezení pro p-hodnotu ($p > 0,1$), průměry, směrodatné odchylky a rozsahy obou výběrů. Protože $p > 0,05$, nulovou hypotézu nezamítáme na hladině významnosti 0,05.

Výpočet je vhodné doplnit krabicovým diagramem typu Medián/Kvartily/Rozpětí.



Komentář: Vidíme, že platby za nákupy kartou Master/EuroCard mají nižší úroveň, ale přibližně stejnou variabilitu jako platby kartou Visa.

Příklad 3.: Z produkce tří podniků vyrábějících televizory bylo vylosováno 10, 8 a 12 kusů. Byly získány následující výsledky zjišťování citlivosti těchto televizorů v mikrovoltech:
1. podnik: 420 560 600 490 550 570 340 480 510 460
2. podnik: 400 420 580 470 470 500 520 530
3. podnik: 450 700 630 590 420 590 610 540 740 690 540 670
Testujte na hladině významnosti 0,05 hypotézu o shodě úrovně citlivosti televizorů v jednotlivých podnicích.

Řešení:
Stejně jako v úkolu 4 provedeme Kruskalův-Wallisův test a mediánový test. Načteme datový soubor televizory.sta. Proměnná X obsahuje hodnoty citlivosti televizorů, proměnná ID udává číslo podniku.
Ve dvou výstupních tabulkách máme výsledky mediánového testu a K-W testu.

	Kruskal-Wallisova ANOVA založ. na Nezávislá (grupovací) proměnná : ID Kruskal-Wallisův test: H (2, N= 30)		
Závislá X	Kó	Počet platný	Součet pořad
1. podi	1	1(	127,0
2. podi	2	9	101,5
3. podi	3	1	236,5

Závislá: X	Mediánový test, celk. medián = 535 Nezávislá (grupovací) proměnná : I[Chi-Kvadr. = 7,632323 sv = 2 p = ,(			
	1. pod	2. pod	3. pod	Celke
<= Medián: pc	6,00(	7,00(	2,00(	15,00
oček	5,00(	4,50(	5,50(	
poz.-	1,00(	2,50(	-3,50(	
> Medián: po	4,00(	2,00(	9,00(	15,00
oček	5,00(	4,50(	5,50(	
poz.-	-1,00(	-2,50(	3,50(	
Celkem:	10,00	9,00(	11,00	30,00

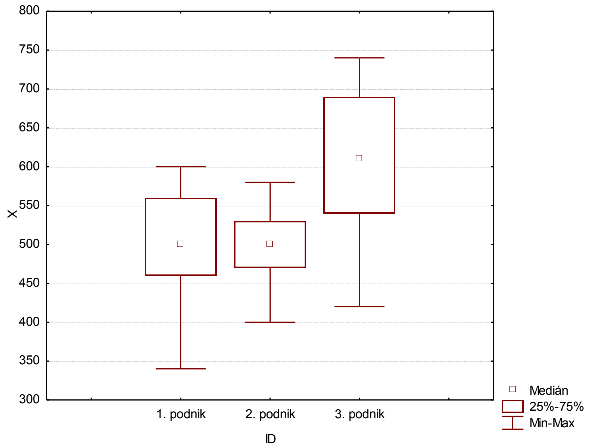
Komentář: Protože zjištěné p-hodnoty jsou menší než zvolená hladina významnosti 0,05, oba testy zamítají hypotézu o shodě mediánů v daných třech skupinách.

Výsledky metody mnohonásobného porovnávání:

Závislá X	Vícenásobné porovnání p hodnot (ob Nezávislá (grupovací) proměnná : ID Kruskal-Wallisův test: H (2, N= 30) =		
	1. pod R:12,7	2. pod R:11,2	3. pod R:21,5
1. podi		1,000	0,066
2. podi	1,000		0,029
3. podi	0,066	0,029	

Komentář: Na hladině významnosti 0,05 se liší citlivost televizorů vyráběných ve 2. a 3. podniku.

Grafické znázornění výsledků:



Komentář: Je vidět, že citlivost televizorů ze 3. podniku je nevyšší, zatímco ze 2.podniku nejnižší. Citlivost výrobků 3. podniku však vykazuje největší variabilitu.