MASARYKOVA UNIVERZITA • PŘÍRODOVĚDECKÁ FAKULTA

Bakalářská práce

Neparametrické metody v systému STATISTICA

DAGMAR LAJDOVÁ

VEDOUCÍ BAKALÁŘSKÉ PRÁCE

RNDr. MARIE BUDÍKOVÁ, Dr.

Brno 2009

Čestné prohlášení

Čestně prohlašuji, že jsem bakalářskou práci vypracovala samostatně pod vedením RNDr. Marie Budíkové, Dr. a s použitím uvedené literatury.

Poděkování

Chtěla bych poděkovat RNDr. Marii Budíkové, Dr. za odborné vedení bakalářské práce, cenné rady a především čas, který mi věnovala.

Obsah

Úvod

V této práci se čtenář seznámí s nejdůležitějšími neparametrickými testy aplikovanými na jedno, dvou, nebo vícerozměrné náhodné výběry a s provedením jejich výpočtu ve statistickém systému STATISTICA. Neparametrické testy se využívají, není-li známo rozložení základního souboru, z něhož byl proveden výběr, nebo když nepochází z normálního rozložení a zejména při malých výběrech z neznámého rozložení.

Práce je rozdělena do pěti kapitol. První kapitola se věnuje vysvětlení základních pojmů potřebných k porozumění následujícího textu a obsahuje seznam použitých značení, která se v textu vyskytují. Druhá kapitola se věnuje testování normality, neboť neparametrické testy se používají, pokud náhodný výběr nepochází z normálního rozložení. Konkrétně je zde vysvětlen princip Shapirova-Wilkova testu a jak se dá ověřovat normalita pomocí grafů. Třetí kapitola se zaobírá jednovýběrovými neparametrickými testy, přesněji znaménkovým a jednovýběrovým Wilcoxonovým testem. Čtvrtá kapitola vysvětluje dvouvýběrový Wilcoxonův test. A pátá kapitola se věnuje vícevýběrovým neparametrickým testům, přesněji Kruskalovu-Wallisovu, mediánovému a Fridmannovu testu. Využití všech testů je ukázáno na příkladě.

V příloze jsou části statistických tabulek, které byly využity při výpočtech.

Celá práce je vysázená systémem TEXLive.

Kapitola 1

Základní pojmy a použité značení

1.1 Základní pojmy

Tato práce předpokládá základní znalosti z oblasti statistiky, proto si vysvětlíme jen ty pojmy, které se daného témata bezprostředně týkají, neboli testování statistických hypotéz.

Testováním statistických hypotéz rozumíme postup, pomocí něhož na základě výsledků zjištěných náhodným výběrem ze základního souboru ověřujeme (testujeme) určitý předpoklad (hypotézu) o parametrech nebo rozdělení základního souboru.

Při testování hypotéz vždy klademe proti sobě dvě hypotézy (tvrzení), z nichž jedna něco tvrdí (předpokládá), druhá to popírá. Předpoklad, který ověřujeme, se nazývá **nulová hypotéza** (někdy také testovaná hypotéza) a označujeme ho H_0. Tvrzení, které tento předpoklad popírá, se nazývá **alternativní hypotéza** a označujeme ho H_1. Volba alternativní hypotézy není libovolná, ale vyplývá z konkrétní situace a z možných rozhodnutí, k nimž test vede. Může být buď oboustranná (tvrdí, že parametr základního souboru je jiný, než tvrdí nulová hypotéza), nebo jednostranná (tvrdí, že parametr je větší nebo že je menší než hodnota, kterou uvádí nulová hypotéza).

Statistický test je rozhodovací proces s pravidly, pomocí nichž na základě hodnoty testové statistiky vypočtené z výběrových dat rozhodneme buď o nezamítnutí nulové hypotézy H_0, nebo o jejím zamítnutí a tím o přijetí altenativní hypotéry H_1.

Testová statistika (kriterium) má pro testování různých hypotéz různý tvar, přičemž může existovat více testových statistik pro test určité hypotézy. Každá testová statistika má známé rozdělení (například $\mathrm{N}, \mathrm{t}, \mathrm{F}, \dots$). Hodnota testové statistiky je vypočtena z dat náhodného výběru, má tedy charakter náhodné veličiny a může nabývat různých hodnot. Přitom obor hodnot testové statistiky tvoří dva neslučitelné obory - obor nezamítnutí nulové hypotézy a kritický obor.

Kritický obor tvoří hodnoty testové statistiky, které svědčí ve prospěch alternativní hypotézy H_1. Jsou to tedy hodnoty, které při platnosti nulové hypotézy mohou nastat jen zřídka (v závislosti na volbě hladiny významnosti α).

Rozhodnutí o nulové hypotéze děláme na základě hodnoty testové statistiky. Realizuje-li se hodnota testového kritéria v kritickém oboru, zamítneme nulovou hypotézu H_0 a přijmeme tvrzení alternativní hypotézy H_1. Naopak realizuje-li se hodnota testové statistiky v oboru přijetí, nulovou hypotézu H_0 nezamítáme.

Při samotném testování hypotéz mohou být závěry i nesprávné, můžeme se dopustit dvou různých druhů chyb.

Chyba 1.druhu spočívá v chybném zamítnutí pravdivé (správné) nulové hypotézy H_0.

Chyba 2.druhu spočívá v nezamítnutí nulové hypotézy, která je ve skutečnosti nesprávná (je to tedy chybné přijetí nulové hypotézy H_0).

Pravděpodobnost chyby 1.druhu nazýváme **hladina významnosti** a označujeme ji α. Pravděpodobnost chyby 2.druhu označujeme β. Pravděpodobnost $1 - \beta$ nazýváme **síla testu**. Udává pravděpodobnost zamítnutí nesprávné hypotézy.

Postup testování:

Stanovíme nulovou hypotézu H_0 a alternativní hypotézu H_1.
Stanovíme hladinu významnosti α.
Vybereme vhodný statistický test a vhodnou testovou statistiku s ohledem na nulovou hypotézu H_0, velikost výběru n a informace o základním souboru.
Z výběrových dat vypočítáme hodnotu příslušné testové statistiky.
Určíme kritický obor.
Provedeme rozhodnutí o nulové hypotéze na základě hodnoty testové statistiky.

V jednotlivých testech v systému STATISTICA je k hodnotě testové statistiky připojována vypočítaná **p-hodnota** (significance level), která představuje minimální hladinu významnosti, na které je možno zamítnout nulovou hypotézu. Nulovou hypotézu H_0 zamítneme, pokud je p-hodnota $\leq \alpha$, a nezamítneme, pokud je p-hodnota $> \alpha$.

Systém STATISTICA zároveň vždy používá asymptotickou variantu testu, bez ohledu na počet pozorování. Proto u příkladů s menším počtem pozorování, než umožňuje využití asymptotického testu, je lepší, když použijeme ruční výpočet.

1.2 Použité značení

W	kritický obor
α	hladina významnosti (zde vždy 0,05)
$x_{0,50}$	medián náhodného výběru $X_1, \ldots, X_n$
u_α	alfa kvantil standardizovaného normálního rozložení
$\chi^2_\alpha(k)$	alfa kvantil Perasonova rozložení s k stupni volnosti
SW	testová statistika pro Shapirův-Wilkův test
a_i	koeficienty u Shapirova-Wilkova testu
$sw_\alpha(n)$	kritická hodnota Shapirova-Wilkova testu pro hladinu významnosti α a rozsah náhodného výběru n
U_Z	testová statistika pro znaménkový test
S_Z^+	počet kladných rozdílů $Y_i = X_i - c$ u znaménkového testu
S_Z^-	počet záporných rozdílů $Y_i = X_i - c$ u znaménkového testu
Z	testová statistika pro kvantilový test
U_W	testová statistika pro jednovýběrový Wilcoxonův test
R_i,R_{ij}	pořadí veličiny $\lvert Y_i \rvert$ u jednovýběrového Wilcoxonova testu, resp. Y_i u dvouvýběrového Wilcoxonova testu, resp. Y_{ij} u Friedmannova testu
S_W^+	součet pořadí přes kladné rozdíly $Y_i = X_i - c$ u jednovýběrového Wilcoxonova testu
S_W^-	součet pořadí přes záporné rozdíly $Y_i = X_i - c$ u jednovýběrového Wilcoxonova testu
$w_\alpha(n)$	kritická hodnota jednovýběrového Wilcoxonoa testu pro hladinu významnosti α a rozsah náhodného výběru n
U_W2	testová statistika pro dvouvýběrový Wilcoxonův test
T_i	součet pořadí hodnot, které patří do i-tého výběru u dvouvýběrového Wilcoxonova testu, respektive u Kruskalova-Wallisova testu
Q_K	testová statistika pro Kruskalův-Wallisův test
Q_M	testová statistika pro mediánový test
Q_F	testová statistika pro Friedmannův test

Kapitola 2

Testování normality

Neparametrické testy používáme, neznáme-li rozložení základního souboru, z něhož byl proveden výběr, nebo když víme, že není normální, nebo pokud mají data pouze ordinální charakter a zejména při malých výběrech z neznámého rozložení.

2.1 Shapirův-Wilkův test

Pomocí Shapirova-Wilkova testu testujeme hypotézu, která tvrdí že náhodný výběr $X_1, \ldots, X_n$ pochází z rozložení $\mathrm{N}(\mu, \sigma^2)$. Testová statistika má pro rozsah výběru n od 3 do 50 tvar

$$SW = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}.$$

Výpočet hodnoty testového kritéria SW vyžaduje koeficienty a_i, jež byly odvozeny speciálně pro potřeby tohoto testu a nalezneme je v tabulce č.1. Přičemž vzhledem k symetrii kvantilů normálního rozložení platí $a_i = -a_{n-i+1}$. Testová statistika se tedy dá přepsat do tvaru

$$SW = \frac{\left(\sum_{i=1}^{m} a_i (x_{(n-i+1)} - x_{(i)})\right)^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2},$$

kde $m = n/2$, je-li n sudé číslo, a $m = (n-1)/2$, je-li n liché.

H_0 zamítáme na hladině významnosti α, pokud je hodnota testové statistiky menší nebo rovna kritické hodnotě pro Shapirův-Wilkův test normality $sw_\alpha(n)$, která je uvedena také v tabulce č.1.

2.2 Příklad

Jsou dány hodnoty $1, 2, 3, 4, 4, 2, 3, 9$. Pomocí Shapirova-Wilkova testu zjistěte na hladině významnosti 0,05, zda pocházejí z normálního rozložení.

Řešení. Využijeme statistiku SW, přičemž uspořádaný náhodný výběr má tvar $(1, 2, 2, 3, 3, 4, 4, 9)$ a průměr daného výběru $\overline{x}$ je roven $3, 5$.

$$SW = \frac{\Big(0,6052(9-1)+0,3164(4-2)+0,1743(4-2)+0,0561(3-3)\Big)^2}{42} \doteq$$
$$\doteq \frac{33,91}{42} \doteq 0,807$$

H_0 zamítáme na hladině významnosti α, pokud je hodnota testové statistiky menší nebo rovna kritické hodnotě $sw_\alpha(n)$. Podle tabulky kritických hodnot pro SW test by měla být hodnota testového kritéria pro $n = 8$ a $\alpha = 0,05$ rovna $0,818$. Protože $0,807 < 0818$, zamítáme na hladině významnosti 0,05 nulovou hypotézu, že data pochází z normálního rozložení.

Postup v programu STATISTICA:
Vytvoříme si datový soubor o jedné proměnné a 8 případech a zapíšeme do něho dané hodnoty. V menu zvolíme *Statistics - Basic Statistic/Tables - Frequency tables*, kde v záložce *Normality* zaškrtneme možnost *Shapiro-Wilk's W test* a spustíme test pomocí tlačítka *Tests for normality.*

	Tests of Normality (Spreadsheet00		
Variable	N	W	p
Data	8	0.807404	0.034337

V tabulce máme uvedené 3 hodnoty. První nám určuje počet hodnot ve výběru. Druhá nám určuje hodnotu testové statistiky W a třetí p-hodnotu.

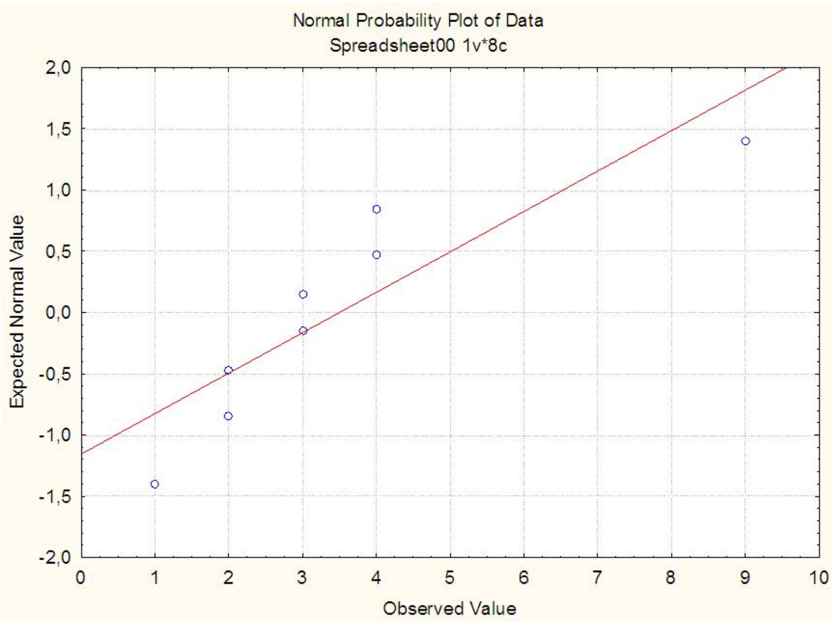
Protože p-hodnota je menší než hladina významnosti α $(0,034337 < 0,05)$, zamítáme na této hladině významnosti předpoklad, že daný výběr pochází z normálního rozložení.

2.3 Grafické ověřování normality

Mezi grafy, pomocí kterých můžeme posuzovat, zda data pochází z normálního rozložení, patří například N-P plot, Q-Q plot, P-P plot, krabicový diagram (box plot) nebo histogram.

Normal probability plot (N-P plot)

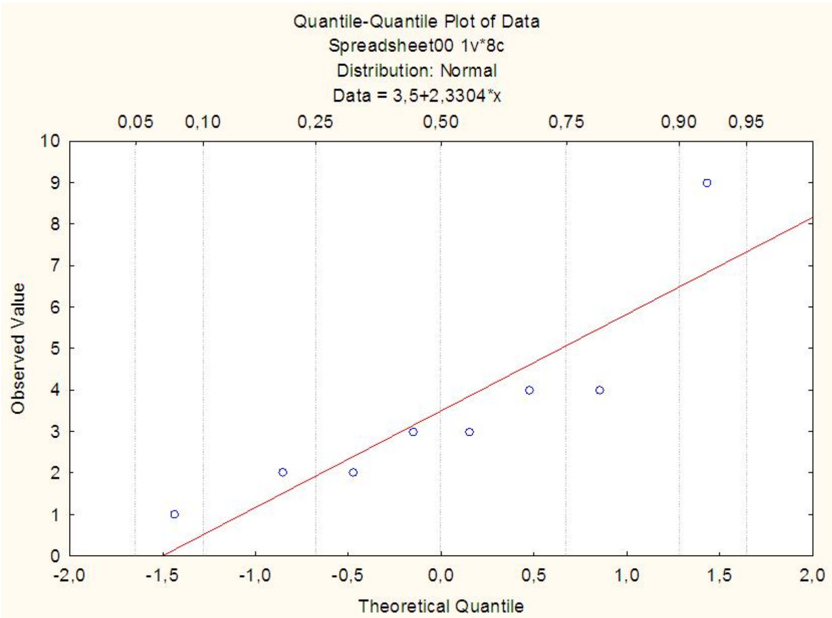
Pomocí tohoto grafu se dá posoudit, zda data pochází z normálního rozložení. V případě normálního rozložení dat budou všechny znázorněné body ležet na přímce (pro data s kladnou šikmostí se body řadí do konvexní křivky a u záporné šikmosti do konkávní křivky). V programu STATISTICA zvolíme v menu *Graphs - 2D Graphs - Normal Probability Plots* a stiskneme tlačítko *OK*.



Vidíme, že body neleží na znázorněné přímce, data tedy nepochází z normálního rozložení, jak nám potvrdil předchozí příklad.

Quantile quantile plot (Q-Q plot)

Pomocí tohoto grafu se dá posoudit, zda data pochází z nějakého známého rozložení (třeba právě z normálního). V programu STATISTICA zvolíme v menu *Graphs - 2D Graphs - Quantile-Quantile Plots* a stiskneme tlačítko *OK*.



Vidíme, že body neleží na znázorněné přímce, data tedy ani podle Q-Q plotu nepochází z normálního rozložení.

Kapitola 3

Jednovýběrové neparametrické testy

Jednovýběrové neparametrické testy jsou neparametrické obdoby jednovýběrových a párových t-testů. V této kapitole se zaměříme na znaménkový test a jednovýběrový Wilcoxonův test. Vysvětlíme si jejich princip a nakonec předvedeme jejich aplikaci na příkladě.

3.1 Znaménkový test

Nechť $X_1, \ldots, X_n$ je náhodný výběr ze spojitého rozložení s mediánem $x_{0,50}$. Platí tedy

$$P(X_i < x_{0,50}) = P(X_i > x_{0,50}) = \frac{1}{2} \quad , \quad i = 1, \ldots, n.$$

U znaménkového testu testujeme hypotézu $H_0 : x_{0,50} = c$ proti oboustranné alternativě $H_1 : x_{0,50} \neq c$ (resp. levostranné alternativě $H_1 : x_{0,50} < c$, resp. proti pravostranné alternativě $H_1 : x_{0,50} > c$), kde c je reálná konstanta.

Položme $Y_i = X_i - c$, $i = 1, \ldots, n$. Je-li některá z veličin X_i rovna c, zpravidla se toto pozorování vynechává. Pak n o počet těchto pozorování snížíme.

Zaveďme statistiku S_Z^+, která udává počet těch rozdílů Y_i, které jsou kladné. Platí-li H_0, má S_Z^+ binomické rozložení $\mathrm{Bi}(n, 1/2)$, tedy $\mathrm{E}(S_Z^+) = n/2$, $\mathrm{D}(S_Z^+) = n/4$.

Hypotézu H_0 zamítneme, bude-li S_Z^+ blízké nule nebo blízké číslu n. Je-li n malé (v praxi $n \leq 20$),používají se tabulky kritických hodnot k_1 a k_2 s vlastnostmi

$$P(S_Z^+ \leq k_1) \leq \frac{\alpha}{2} \quad , \quad P(S_Z^+ \geq k_2) \geq \frac{\alpha}{2}. \tag{3.1}$$

Přitom k_1 je nejmenší a k_2 je největší z čísel, pro která (3.1) platí. Vzhledem k tomu, že binomické rozložení je pro $\vartheta = 1/2$ symetrické, je $k_2 = n - k_1$.

Kritický obor W je tedy $\langle 0; k_1 \rangle \cup \langle k_2; n \rangle$ a hypotézu H_0 zamítneme na hladině významnosti α, když $S_Z^+ \in W$.

Je-li $n \geq 20$, používá se v praxi následující testová statistika.

Platí-li H_0, má veličina

$$U_Z = \frac{S_Z^+ - \frac{n}{2}}{\sqrt{\frac{n}{4}}} = \frac{S_Z^+ - \mathrm{E}(S_Z^+)}{\sqrt{\mathrm{D}(S_Z^+)}}$$

podle centrální limitní věty asymptoticky rozložení $\mathrm{N}(0,1)$.

Poznámka 3.1. Aproximace normálním rozložením $\mathrm{N}(0,1)$ se zlepší použitím tzv. korekce na nespojitost. Testová statistika U_Z má pak tvar

$$U_Z = \frac{S_Z^+ - \frac{n}{2} \pm \frac{1}{2}}{\sqrt{\frac{n}{4}}},$$

přičemž $\frac{1}{2}$ přičítáme, pokud $S_Z^+ \leq \frac{n}{2}$, a odečítáme v opačném případě.

Důsledek 3.1. *Kritický obor pro oboustrannou alternativu má tvar*

$$W = (-\infty, -u_{1-\alpha/2}\rangle \cup \langle u_{1-\alpha/2}, \infty).$$

(Analogicky pro jednostranné alternativy.) Pokud se statistika U_Z realizuje v kritickém oboru W, zamítáme H_0 na asymptotické hladině významnosti α.

Poznámka 3.2. Znaménkový test lze modifikovat na párový test.

Věta 3.2. *Nechť $(X_1, Y_1), \dots, (X_n, Y_n)$ je náhodný výběr ze spojitého dvourozměrného rozložení. Testujeme hypotézu $H_0 : x_{0,50} - y_{0,50} \neq c$, kde c je reálná konstanta. Vytvoříme rozdíly $Z_i = X_i - Y_i$, $i = 1, \dots, n$ a testujeme podle předchozího postupu hypotézu o mediánu $z_{0,50}$, tj. $H_0 : z_{0,50} = c$ proti $H_1 : z_{0,50} \neq c$. (Jednostranné alternativy analogicky.)*

Poznámka 3.3. Jelikož tento test má poměrně malou sílu (pravděpodobnost chyby druhého druhu, kdy nezamítáme H_0, i když je nepravdivá, je ve srovnání s jinými testy značně velká), je žádoucí mít k dispozici větší počet pozorování n.

Poznámka 3.4. Znaménkový test je speciálním příkladem kvantilového testu, kdy testujeme hypotézu o mediánu. Pomocí kvantilového testu testujeme hypotézu $H_0 : x_q = c$ proti oboustranné alternativě $H_1 : x_q \neq c$ (resp. levostranné alternativě $H_1 : x_q < c$, resp. proti pravostranné alternativě $H_1 : x_q > c$), kde $q \in (0,1; 0,9)$ a c je reálná konstanta.
K rozhodnutí o platnosti nulové hypotézy se využívá testová statistika

$$Z = \frac{S_Z^+ - nq}{\sqrt{nq(1-q)}},$$

která má za platnosti H_0 standardní normální rozložení $\mathrm{N}(0,1)$.
Vidíme tedy, že pro $q = 0,5$ dostáváme $Z = U_Z$.

3.2 Příklad

Jednou za semestr studenti pomocí bodů hodnotí přednášející. K dispozici jsou průměrné výsledky na jednoho studenta za dva roky pro deset náhodně vybraných přednašejících. Úkolem je na hladině významnosti 0,05 posoudit, zda přednášející získali v letošním roce od studentů stejný počet bodů jako v minulém roce. K rozhodnutí využijeme následující údaje a znaménkový test.

Číslo učitele	1	2	3	4	5	6	7	8	9	10
Body vloni	932	906	943	907	893	870	889	902	866	787
Body letos	933	923	942	909	908	893	890	900	888	803

Řešení. Máme určit, zda přednášející získali v letošním roce (hodnoty Y_i) stejný počet bodů jako v loňském roce (hodnoty X_i). Jedná se o párový test, kde testujeme hypotézu $H_0 : x_{0,50} - y_{0,50} = 0$ proti oboustranné alternativě $H_0 : x_{0,50} - y_{0,50} \neq 0$

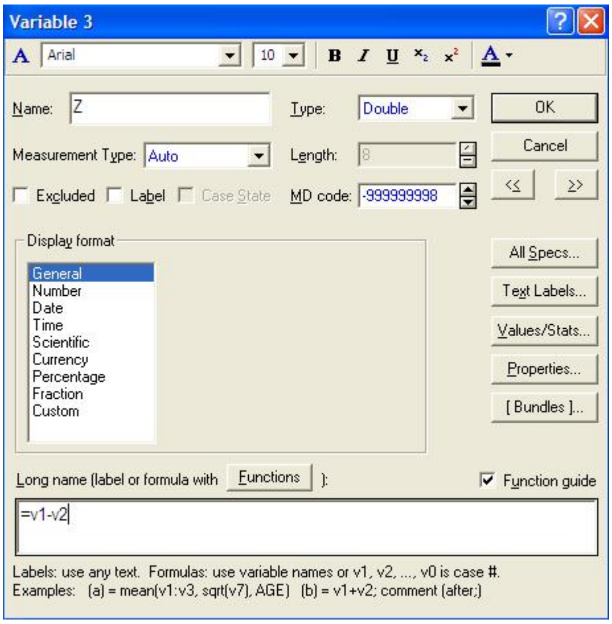
Nejdříve spočítáme rozdíly hodnot X_i a Y_i.

Číslo učitele	1	2	3	4	5	6	7	8	9	10
$X_i - Y_i$	-1	-17	1	-2	-15	-23	-1	2	-22	-16

Tímto jsme převedli párový test na jednovýběrovou variantu. Zároveň vidíme, že testová statistika S_Z^+, která udává počet rozdílů, které jsou kladné, nabývá hodnoty 2 (učitelé číslo 3 a 8). Ve statistických tabulkách najdeme pro $n = 10$ a $\alpha = 0,05$ kritické hodnoty $k_1 = 1$, $k_2 = 9$. Protože se jedná o oboustrannou alternativu, kritický obor je $W = \langle 0, k_1 \rangle \cup \langle k_2, n \rangle = \langle 0, 1 \rangle \cup \langle 9, 10 \rangle$. Protože kritický obor neobsahuje hodnotu 2, nemůžeme H_0 zamítnout na hladině významnosti $0,05$.

Přestože tedy podle výsledků získalo osm učitelů lepší hodnocení než v minulém roce, nemůžeme zamítnout tvrzení, které říká, že přednášející získali letos od studentů v průměru stejný počet bodů jako vloni.

Postup v programu STATISTICA:
Vytvoříme datový soubor o 4 proměnných a 10 případech. Do prvního sloupce, který znázorňuje první proměnnou, zapíšeme hodnoty X_i a do druhého sloupce, který znázorňuje druhou proměnnou, zapíšeme hodnoty Y_i. Do třetí proměnné zapíšeme rozdíly $Z_i = X_i - Y_i$ a do posledního sloupce zapíšeme do všech řádků hodnoty konstanty c, což je v našem případě nula. Hodnoty Z_i a c nemusíme vypisovat ručně. Dvakrát poklepeme na hlavičku příslušné proměnné a tím se nám otevře okno, ve kterém v dolní části (*Long name*) nastavíme vzorec pro výpočet dané proměnné. V případě Z_i napíšeme *=v1-v2*, což značí, že hodnoty třetí proměnné jsou vypočítány jako rozdíly první a druhé proměnné. V případě c napíšeme *=0*, což odpovídá hodnotě konstanty c, která je pro všechna i rovna nule.



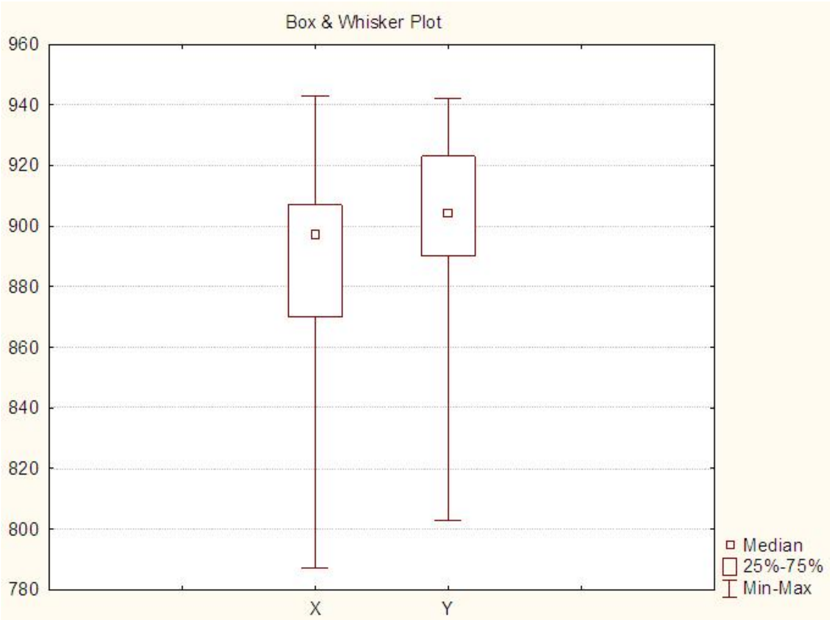
Dále zvolíme v menu *Statistics - Nonparametrics - Comparing two dependent samples (variables)*. Ve *Variables* zvolíme proměnné, které nás zajímají (v tomto případě proměnné 3 a 4 - hodnoty Z_i a c) a spustíme test pomocí tlačítka *Sign test*.

	Sign Test (Spreadsheet01) Marked tests are significant at p <,05000			
Pair of Variables	No. of Non-ties	Percent v < V	Z	p-level
Z & c	10	80,00000	1,581139	0,113846

V tabulce máme uvedené 4 hodnoty. První nám určuje počet nenulových rozdílů, tedy rozsah výběru, se kterým se výpočet provádí. Druhá udává procento záporných rozdílů (80% z 10 = 8). Třetí nám určuje hodnotu testové statistiky U_Z a čtvrtá p-hodnotu.

Protože p-hodnota je větší než hladina významnosti α $(0,113846 > 0,05)$, nemůžeme H_0 na této hladině významnosti zamítnout.

Mohli bychom také hodnotu testové statistiky U_Z porovnat s příslušným α-kvantilem, kdy pro $|U_Z| \geq u_{1-\alpha/2}$ H_0 zamítáme. Vidíme, že $1,581139 \ngeq 1,95996$, takže ani tímto způsobem nemůžeme H_0 zamítnout. V tomto případě je však vhodnější využít kritické hodnoty z tabulek, protože rozsah výběru nesplňuje podmínku asymptotické normality statistiky S_Z^+, tj. $n > 20$.



Z krabicového grafu lze usoudit, že předpoklad o rovnosti mediánů obou výběrů je oprávněný.

Poznámka 3.5. Dalo by se postupovat jednodušeji a vytvořit datový soubor o pouze 2 proměnných, do kterých bychom zapsali hodnoty X_i a c (v případě jednovýběrové varianty), případně X_i a Y_i (u párové varianty). Program STATISTICA totiž řeší zadaný problém tak, že porovnává dvě proměnné a určuje, zda je jejich rozdíl nulový.

3.3 Wilcoxonův test

Wilcoxonův test je obdobou předešlého znaménkového testu s tím rozdílem, že se nejen zkoumá, zda je daná hodnota větší nebo menší než medián, ale že se v úvahu berou i velikosti odchylek. Proti znaménkovému testu má Wilcoxonův test větší sílu.

Nechť $X_1, \dots, X_n$ je náhodný výběr ze spojitého rozložení s hustotou $f(x)$, která je symetrická kolem mediánu $x_{0,50}$ a kladná v jeho okolí. Platí tedy $f(x_{0,50} - x) = f(x_{0,50} + x)$. Testujeme hypotézu $H_0 : x_{0,50} = c$ proti oboustranné alternativě $H_1 : x_{0,50} \neq c$ (resp. levostranné alternativě $H_1 : x_{0,50} < c$, resp. proti pravostranné alternativě $H_1 : x_{0,50} > c$), kde c je reálná konstanta.

Položme $Y_i = X_i - c$, $i = 1, \dots, n$. Je-li některá z veličin X_i rovna c, zpravidla se toto pozorování vynechává. Pak n o počet těchto pozorování snížíme.

Veličiny Y_i seřadíme do neklesající posloupnosti podle jejich absolutní hodnoty. Označme R_i pořadí veličiny $|Y_i|$ a zaveďme statistiky

$$S_W^+ = \sum_{Y_i > 0} R_i \quad , \quad S_W^- = \sum_{Y_i < 0} R_i,$$

kde S_W^+ je součet pořadí přes kladné hodnoty Y_i a S_W^- je součet pořadí přes záporné hodnoty Y_i. Přitom platí $S_W^+ + S_W^- = n\,(n+1)\,/2$.

Občas se stane, že některé hodnoty nejsou různé. Těm potom přiřazujeme průměrné pořadí.

H_0 zamítáme na hladině významnosti α, pokud je hodnota testové statistiky menší nebo rovna tabelované kritické hodnotě $w_\alpha(n)$. Testová statistika pro oboustrannou alternativu je $\min(S_W^+, S_W^-)$ (resp. S_W^+ pro levostrannou alternativu, resp. S_W^- pro pravostranou alternativu).

Z učiněných předpokladů vyplývá, že při platnosti H_0 jsou $Y_1, \ldots, Y_n$ nezávislé stejně rozdělené náhodné veličiny, jejichž rozložení je symetrické kolem nuly.

Věta 3.3. *Platí-li H_0, pak vektory $(signY_1, \ldots, signY_n)'$ a $(|Y|_{(1)}, \ldots, |Y|_{(n)})$ jsou nezávislé.*

Důkaz. Jelikož veličiny Y_i jsou nezávislé, vektory $(signY_i, |Y_i|)'$ jsou také nezávislé. Ze spojitosti a ze symetrie rozložení vyplývá, že

$$\mathrm{P}(signY_i = 1) = \mathrm{P}(signY_i = -1) = \frac{1}{2}$$

Dále máme pro libovolné $y > 0$

$$\begin{aligned}
\mathrm{P}(signY_i = 1, |Y_i| < y) &= \mathrm{P}(0 < Y_i < y) = \frac{1}{2}\mathrm{P}(-y < Y_i < y) = \\
&= \frac{1}{2}\mathrm{P}(|Y_i| < y) = \frac{1}{2}\mathrm{P}(signY_i = 1)\mathrm{P}(|Y_i| < y).
\end{aligned}$$

Proto veličiny $signY_i$ a $|Y_i|$ jsou pro každé i nezávislé. Celkově dostáváme, že vektory $(signY_1, \ldots, signY_n)'$ a $(|Y_1|, \ldots, |Y_n|)'$ jsou nezávislé. Protože vektor $(|Y|_{(1)}, \ldots, |Y|_{(n)})'$ je funkcí vektoru $(|Y_1|, \ldots, |Y_n|)'$, je tím věta dokázána. □

Věta 3.4. *Označme $S = \sum_{i=1}^n R_i signY_i = S_W^+ - S_W^-$. Pak*

$$S_W^+ = \frac{1}{2}S + \frac{1}{4}n(n+1).$$

Důkaz. Platí $S = S_W^+ - S_W^-$ a $S_W^+ + S_W^- = n(n+1)/2$. Odtud vypočteme S_W^+ □

Věta 3.5. *Platí-li H_0, pak*

$$\mathrm{E}(S_W^+) = \frac{1}{4}n(n+1) \quad , \quad \mathrm{D}(S_W^+) = \frac{1}{24}n(n+1)(2n+1).$$

Důkaz. Nejprve si všimněme, že $\mathrm{E}(signY_i) = 0$ pro každé i. Z věty 3.3 dostaneme, že $\mathrm{E}(R_i signY_i) = \mathrm{E}(R_i)\mathrm{E}(signY_i)$, a tak

$$\mathrm{E}(R_i signY_i) = 0. \tag{3.2}$$

Odtud

$$\mathrm{E}(S) = \sum_{i=1}^n \mathrm{E}(R_i signY_i) = 0.$$

Vzhledem k (3.2) platí

$$\begin{aligned}\mathrm{D}(R_i signY_i) &= \mathrm{E}[(R_i signY_i)^2] = \mathrm{E}[(R_i)^2]\mathrm{E}[(signY_i)^2] = \\ &= \mathrm{E}[(R_i)^2] = 1^2\frac{1}{n} + 2^2\frac{1}{n} + \cdots + n^2\frac{1}{n} = \frac{1}{6}(n+1)(2n+1).\end{aligned}$$

Obdobně se dokáže, že platí

$$\mathrm{C}(R_i signY_i, R_j signY_j) = 0 \quad \text{pro } i \neq j.$$

Nyní společně s využitím věty 3.4 se vypočítá

$$\mathrm{E}(S_W^+) = \mathrm{E}[\frac{1}{2}S + \frac{1}{4}n(n+1)] = \frac{1}{2}\mathrm{E}(S) + \mathrm{E}[\frac{1}{4}n(n+1)] = \frac{1}{4}n(n+1)$$

$$\begin{aligned}\mathrm{D}(S_W^+) &= \mathrm{D}[\frac{1}{2}S + \frac{1}{4}n(n+1)] = \frac{1}{4}\mathrm{D}(S) = \frac{1}{4}\mathrm{D}(\sum_{i=1}^{n} R_i signY_i) = \\ &= \frac{1}{4}n\frac{1}{6}(n+1)(2n+1) = \frac{1}{24}n(n+1)(2n+1).\end{aligned}$$

□

Za platnosti H_0 má S_W^+ asymptoticky normální rozložení. Pro $n > 30$ lze tedy test hypotézy H_0 založit na veličině

$$U_W = \frac{S_W^+ - \mathrm{E}(S_W^+)}{\sqrt{\mathrm{D}(S_W^+)}} = \frac{S_W^+ - \frac{1}{4}n(n+1)}{\sqrt{\frac{1}{24}n(n+1)(2n+1)}},$$

která má za platnosti H_0 standardní normální rozložení $\mathrm{N}(0,1)$.

Důsledek 3.6. *Kritický obor pro oboustrannou alternativu má tvar*

$$W = (-\infty, -u_{1-\alpha/2}\rangle \cup \langle u_{1-\alpha/2}, \infty).$$

(Analogicky pro jednostranné alternativy.) Pokud se statistika U_W realizuje v kritickém oboru W, zamítáme H_0 na asymptotické hladině významnosti α.

Poznámka 3.6. Jednovýběrový Wilcoxonův test lze modifikovat na párový test.

Věta 3.7. *Nechť $(X_1, Y_1), \ldots, (X_n, Y_n)$ je náhodný výběr ze spojitého dvourozměrného rozložení. Testujeme hypotézu $H_0 : x_{0,50} - y_{0,50} \neq c$, kde c je reálná konstanta. Vytvoříme rozdíly $Z_i = X_i - Y_i$, $i = 1, \ldots, n$ a testujeme podle předchozího postupu hypotézu o mediánu $z_{0,50}$, tj. $H_0 : z_{0,50} = c$ proti $H_1 : z_{0,50} \neq c$. (Jednostranné alternativy analogicky.)*

3.4 Příklad

Spočítejme si stejný příklad jako u znaménkového testu. Máme tedy určit, zda přednášející získali v průměru stejná honocení v loňském i letošním roce. Připomeňme si, jaká byla:

Číslo učitele	1	2	3	4	5	6	7	8	9	10
Body vloni	932	906	943	907	893	870	889	902	866	787
Body letos	933	923	942	909	908	893	890	900	888	803

Hypotézu budeme testovat na hladině významnosti 0,05 a tentokrát použijeme jednovýběrový Wilcoxonův test.

Řešení. Testujeme hypotézu $H_0 : x_{0,50} - y_{0,50} = 0$ proti oboustranné alternativě $H_0 : x_{0,50} - y_{0,50} \neq 0$. Opět si párový test převedeme na jednovýběrovou variantu. Rozdíly hodnot X_i a Y_i jsou následující:

Číslo učitele	1	2	3	4	5	6	7	8	9	10
$Z_i = X_i - Y_i$	-1	-17	1	-2	-15	-23	-1	2	-22	-16

Veličiny Z_i seřadíme do neklesající posloupnosti podle jejich absolutní hodnoty a přiřadíme jim příslušná pořadí.

$Z_i = X_i - Y_i$	-1	-1	1	-2	2	-15	-16	-17	-22	-23
Pořadí R_i	2	2	**2**	4,5	**4,5**	6	7	8	9	10

Spočítáme si hodnoty statistik S_W^+ a S_W^- udávající hodnotu součtu pořadí přes kladné, respektive záporné hodnoty.

$$S_W^+ = 6,5 \quad , \quad S_W^- = 48,5.$$

Testová statistika pro oboustrannou alternativu je $\min(6,5; 48,5)$. Hodnota testové statistiky je tedy, jak vidíme, 6,5. H_0 zamítáme na hladině významnosti α, pokud je hodnota testové statistiky menší nebo rovna tabelované kritické hodnotě $w_\alpha(n)$. Pro $n = 10$ a $\alpha = 0,05$ odpovídá kritická hodnota číslu 8. Protože $6,5 < 8$, zamítáme nulovou hypotézu H_0, že hodnocení byla v obou letech v průměru stejná, na hladině významnosti $0,05$.

Jak vidíme, na rozdíl od znaménkového testu Wilcoxonův test hypotézu zamítá. To odpovídá předpokladu, že u znaménkového testu je pravděpodobnost chyby druhého druhu vyšší.

Postup v programu STATISTICA:

Mohli bychom postupovat stejně jako v případě znaménkového testu a vytvořit si datový soubor o 4 proměnných (X_i, Y_i, Z_i, c), ale i tentokrát stačí postupovat jednodušeji a vytvořit si datový soubor o pouze 2 proměnných a 10 případech. Do příslušných proměnných zadáme hodnoty pro X_i a Y_i.

Dále v menu zvolíme *Statistics - Nonparametrics - Comparing two dependent samples (variables)* (úplně stejné jako v případě znaménkového testu). Ve *Variables* zvolíme proměnné a spustíme test pomocí tlačítka *Wilcoxon matched pairs test.*

Pair of Variables	Wilcoxon Matched Pairs Test (Spreadsheet01) Marked tests are significant at p <,05000 Valid N	T	Z	p-level
X & Y	10	6,500000	2,140518	0,032314

V tabulce vidíme 4 hodnoty. První udává počet hodnot, se kterými se pracuje ($n = 10$), druhá udává hodnotu testové statistiky $\min(S_W^+, S_W^-)$, třetí hodnotu statistiky $|U_W|$ a čtvrtá p-hodnotu.

Protože p-hodnota je menší než hladina významnosti α ($0,032314 < 0,05$), zamítáme H_0 na této hladině významnosti.

Kapitola 4

Dvouvýběrové neparametrické testy

Dvouvýběrové neparametrické testy jsou neparametrickou obdobou dvouvýběrového t-testu. V této kapitole se zaměříme na vysvětlení principu dvouvýběrového Wilcoxonova testu a ukážeme si jeho výpočet na příkladě.

4.1 Dvouvýběrový Wilcoxonův test

Nechť $X_1, \ldots, X_m$ je náhodný výběr ze spojitého rozložení s distribuční funkcí F a nechť $Y_1, \ldots, Y_n$ je na něm nezávislý náhodný výběr ze spojitého rozložení s distribuční funkcí G. Testujeme hypotézu $H_0 : F = G$ proti alternativě $H_1 : F \neq G$.

Všech $m + n$ veličin $X_1, \ldots, X_m, Y_1, \ldots, Y_n$ (sdružený výběr) uspořádáme vzestupně podle velikosti. Přitom se označení výběrů volí tak, aby platilo $m \geq n$. V tomto sdruženém výběru mají veličiny $X_1, \ldots, X_m$ pořadí $R_1, \ldots, R_m$ a veličiny $Y_1, \ldots, Y_n$ pořadí $R_m + 1, \ldots, R_m + n$. Označme T_1 součet pořadí veličin $X_1, \ldots, X_m$ a T_2 součet pořadí veličin $Y_1, \ldots, Y_n$. Tedy musí platit $T_1 + T_2 = (m + n)(m + n + 1)/2$.

Věta 4.1. *Platí-li H_0, pak*

$$\mathrm{E}(T_1) = \frac{1}{2}m(m + n + 1), \quad \mathrm{D}(T_1) = \frac{1}{12}mn(m + n + 1).$$

Důkaz. Viz [1], str.236-237. □

Místo veličin T_1, T_2 se zpravidla používají veličiny

$$U_1 = mn + m(m + 1)/2 - T_1, \quad U_2 = mn + n(n + 1)/2 - T_2.$$

Přitom platí $U_1 + U_2 = mn$. Testu založeném na U_1, U_2 se někdy také říká **Mannův-Whitneyův test**. Pokud $\min(U_1, U_2)$ je menší nebo rovno tabelované kritické hodnotě (pro dané rozsahy výběrů m, n a dané α), pak nulovou hypotézu o totožnosti obou distribučních funkcí zamítáme na hladině významnosti α.

Pro velká m a n (prakticky $n, m > 30$, ale dá se aplikovat už při $m, n > 10$) se použije následující postup.
Z věty 4.1 plyne, že

$$\mathrm{E}(U_1) = \frac{1}{2}mn, \quad \mathrm{D}(U_1) = \frac{1}{12}mn(m+n+1).$$

Jelikož $U_2 = mn - U_1$, platí $\mathrm{E}(U_2) = \mathrm{E}(U_1), \mathrm{D}(U_2) = \mathrm{D}(U_1)$. Zároveň pro $m \to \infty$ a $n \to \infty$ má veličina U_1 (i veličina T_1) asymptoticky normální rozložení. Vypočte se tedy

$$U_{W2} = \frac{U_1 - \mathrm{E}(U_1)}{\sqrt{\mathrm{D}(U_1)}}.$$

U_{W2} má za platnosti H_0 standardní normální rozložení $\mathrm{N}(0,1)$.

Poznámka 4.1. Aproximace normálním rozložením $\mathrm{N}(0,1)$ se zlepší použitím tzv. korekce na nespojitost. Testová statistika U_{W2} má pak tvar

$$\frac{U_1 - \mathrm{E}(U_1) \pm \frac{1}{2}}{\sqrt{\mathrm{D}(U_1)}},$$

přičemž $\frac{1}{2}$ přičítáme, pokud $U_1 - \mathrm{E}(U_1) \leq \frac{n}{2}$, a odečítáme v opačném případě.

Důsledek 4.2. *Kritický obor pro oboustrannou alternativu má tvar*

$$W = (-\infty, -u_{1-\alpha/2}\rangle \cup \langle u_{1-\alpha/2}, \infty).$$

Pokud se statistika U_{W2} realizuje v kritickém oboru W, zamítáme H_0 na asymptotické hladině významnosti α.

4.2 Příklad

Trenér běžeckého družstva se rozhodl vyzkoušet novou metodu tréninku. Osm běžců využívalo starou metodu a čtyři se rozhodli vyzkoušet nový způsob. Po určité době se změřily výsledky, za kolik sekund uběhne každý běžec vzdálenost 100 metrů. Ty byly následující:

Metoda A	13,05	12,58	14,26	15,08	12,37	10,45	14,64	12,73
Metoda B	11,98	12,68	13,04	12,49				

Naším úkolem je na hladině významnosti 0,05 posoudit, zda je nová metoda tréninku stejně efektivní jako metoda stará.

Řešení. Testujeme hypotézu $H_0 : x_{0,50} = y_{0,50}$ proti oboustranné alternativě $H_0 : x_{0,50} \neq y_{0,50}$.

Všechny hodnoty seřadíme do neklesající posloupnosti a přiřadíme jim příslušná pořadí.

Seřazené hodnoty	10,45	11,98	12,37	12,49	12,58	12,68
Pořadí A	1	-	3	-	5	-
Pořadí B	-	2	-	4	-	6

Seřazené hodnoty	12,73	13,04	13,05	14,26	14,64	15,08
Pořadí A	7	-	9	10	11	12
Pořadí B	-	8	-	-	-	-

Nyní si vypočítáme hodnoty veličin T_1 a T_2 udávající součet pořadí prvního, respektive druhého výběru, a z nich poté určíme hodnoty veličin U_1 a U_2:

$$T_1 = 1 + 3 + 5 + 7 + 9 + 10 + 11 + 12 = 58, \quad T_2 = 2 + 4 + 6 + 8 = 20$$

$$U_1 = 8 \cdot 4 + 8 \cdot 9/2 - 58 = 10, \quad U_2 = 8 \cdot 4 + 4 \cdot 5/2 - 20 = 22$$

H_0 zamítáme na hladině významnosti α, pokud je $\min(U_1, U_2)$ menší nebo rovno tabelované kritické hodnotě. Pro $m = 8$, $n = 4$ a $\alpha = 0,05$ odpovídá kritická hodnota číslu 4 a $min(10, 22) = 10$. Protože $10 \nleq 4$, nemůžeme na dané hladině významnosti zamítnout nulovou hypotézu H_0, že jsou obě metody tréninku stejně efektivní.

Postup v programu STATISTICA:
Vytvoříme si datový soubor o 2 proměnných a 12 případech. Do první proměnné napíšeme naměřené časy a do druhé čísla 1 nebo 2 podle toho, kterou metodu tréninku běžec využíval.

	1 Hodnoty	2 Metoda
1	13,05	1
2	12,58	1
3	14,26	1
4	15,08	1
5	12,37	1
6	10,45	1
7	14,64	1
8	12,73	1
9	11,98	2
10	12,68	2
11	13,04	2
12	12,49	2

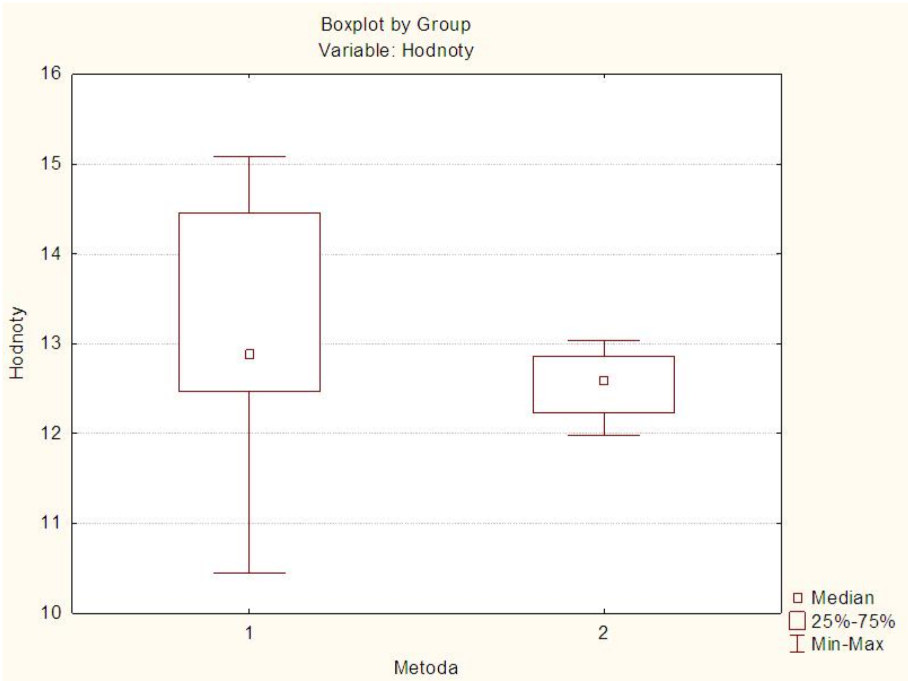
Dále v menu zvolíme *Statistics - Nonparametrics - Comparing two independent samples (groups)*. Ve *Variables* zvolíme za závislou proměnnou první proměnnou (Hodnoty) a za nezávislou (grupovací) proměnnou zvolíme druhou proměnnou (Metoda). Test spustíme pomocí tlačítka *Mann-Whitney U test.*

Mann-Whitney U Test (Spreadsheet02)
By variable Metoda
Marked tests are significant at p <,05000

variable	Rank Sum Group 1	Rank Sum Group 2	U	Z	p-level	Z adjusted	p-level	Valid N Group 1	Valid N Group 2	2*1sided exact p
Hodnoty	58,00000	20,00000	10,00000	0,934129	0,350238	0,934129	0,350238	8	4	0,367677

V tabulce máme postupně hodnoty T_1, T_2, $\min(U_1, U_2)$, U_{W2}, p-hodnota, upravené U_{W2}, p-hodnota, m, n a přesná p-hodnota, která nás v tomto případě nejvíce zajímá, protože se používá pro rozsahy výběrů pod 30.

Jelikož $0,367677 \nleq 0,05$, nemůžeme na dané hladině významnosti zamítnout nulovou hypotézu H_0, že jsou obě metody tréninku stejně efektivní.



Z krabicového grafu lze usoudit, že předpoklad o rovnosti mediánů obou výběrů je oprávněný.

Kapitola 5

Vícevýběrové neparametrické testy

Vícevýběrové neparametrické testy jsou neparametrickou obdobou parametrických testů založených na analýze rozptylu jednoduchého třídění, respektive dvojného třídění. V této kapitole se zaměříme na vysvětlení principu mediánového, Kruskalova-Wallisova a Friedmannova testu a ukážeme si jejich výpočet na příkladě.

5.1 Kruskalův-Wallisův test

Kruskalův-Wallisův test je neparametrickou obdobou analýzy rozptylu jednoduchého třídění.

Nechť $Y_{i1}, \dots, Y_{in_i}, i = 1, \dots, r$ je r náhodných výběrů o rozsazích $n_1, \dots, n_r$. Nechť všechny tyto výběry jsou na sobě nezávislé a pocházejí ze spojitého rozložení, přičemž i-tý výběr pochází z rozložení s distribuční funkcí $F_i, i = 1, \dots, r$. Testujeme hypotézu H_0, že všechny tyto výběry pocházejí ze stejného rozložení, neboli

$$H_0 : F_1(x) = \dots = F_r(x) \quad \text{pro } \forall x$$

Veličiny Y_{ij} vytvoří dohromady sdružený výběr o rozsahu $N = n_1 + \dots + n_r$. Uspořádáme je do rostoucí posloupnosti, ve které mají veličiny $Y_{11}, \dots, Y_{1n_1}$ pořadí $R_1, \dots, R_{n_1}$, veličiny $Y_{21}, \dots, Y_{2n_2}$ pořadí $R_{n_1+1}, \dots, R_{n_1+n_2}$ a tak dále. Označmě T_i součet pořadí těch hodnot, které patří do i-tého výběru, $i = 1, \dots, r$. Tedy musí platit $T_1 + \dots + T_r = N(N+1)/2$.

Testová statistika má tvar

$$Q_K = \frac{12}{N(N+1)} \sum_{i=1}^{r} \frac{T_i^2}{n_i} - 3(N+1).$$

Platí-li H_0, má Q_K asymptoticky rozložení $\chi^2(r-1)$, rostou-li rozsahy výběrů nade všechny meze. Tedy pro $Q_K \geq \chi^2_{1-\alpha}(r-1)$ zamítneme H_0 na hladině významnosti, která se s rostoucími n_i blíží číslu α.

Poznámka 5.1. Zamítneme-li H_0, je třeba rozhodnout, které dvojice výběrů se od sebe významně liší. K tomu se využívají **metody mnohonásobného porovnávání**.

Pokud mají výběry stejný rozsah n, můžeme použít **Neményiho metodu**: Je-li $|T_l - T_k| \geq$ tabelovaná kritická hodnota, zamítáme na hladině významnosti α hypotézu, že l-tý a k-tý výběr pocházejí ze stejného rozložení.

Pokud mají výběry různé rozsahy, použijeme **obecnou metodu mnohonásobného porovnávání**: Je-li

$$|T_l - T_k| \geq \sqrt{\frac{1}{12}\Big(\frac{1}{n_l} + \frac{1}{n_k}\Big)N(N+1)h_\alpha},$$

zamítáme na hladině významnosti α hypotézu, že l-tý a k-tý výběr pocházejí ze stejného rozložení. Při větších rozsazích výběrů je možno nahradit kritickou hodnotu h_α kvantilem $\chi_{1-\alpha}(r-1)$.

5.2 Příklad

Máme 4 druhy výrobků a 6 firem, které je prodávají. Na základě údajů o počtu prodaných kusů jednotlivých druhů výrobků každou firmou za jeden měsíc máme určit, zda se od sebe mediány počtu prodaných kusů jednotlivých výrobků liší. Hypotézu budeme testovat na hladině významnosi 0,05.

	Firmy					
	1.	2.	3.	4.	5.	6.
1.výrobek	642	725	680	732	801	643
2.výrobek	558	632	721	823	801	561
3.výrobek	645	698	769	767	806	603
4.výrobek	679	785	568	621	689	777

Řešení. Testujeme hypotézu H_0, že mediány počtu prodaných kusů jednotlivých výrobků jsou stejné, proti alternativě, že alespoň jeden medián se liší. Nejprve všechny hodnoty uspořádáme vzestupně podle velikosti a stanovíme jejich pořadí.

Seřazené hodnoty	558	561	568	603	621	632	642	643
1.výrobek	-	-	-	-	-	-	7	8
2.výrobek	1	2	-	-	-	6	-	-
3.výrobek	-	-	-	4	-	-	-	-
4.výrobek	-	-	3	-	5	-	-	-

Seřazené hodnoty	645	679	680	689	698	721	725	732
1.výrobek	-	-	11	-	-	-	15	16
2.výrobek	-	-	-	-	-	14	-	-
3.výrobek	9	-	-	-	13	-	-	-
4.výrobek	-	10	-	12	-	-	-	-

Seřazené hodnoty	767	769	777	785	801	801	806	823
1.výrobek	-	-	-	-	21,5	-	-	-
2.výrobek	-	-	-	-	-	21,5	-	24
3.výrobek	17	18	-	-	-	-	23	-
4.výrobek	-	-	19	20	-	-	-	-

Vypočteme součty pořadí jednotlivých výrobků:

$$T_1 = 78,5, \quad T_2 = 68,5, \quad T_3 = 84, \quad T_4 = 69$$

Nyní určíme hodnotu testové statistiky Q_K:

$$Q_K = \frac{12}{24 \cdot 25}\Big(\frac{78,5^2}{6} + \frac{68,5^2}{6} + \frac{84^2}{6} + \frac{69^2}{6}\Big) - 3 \cdot 25 \doteq 0,572$$

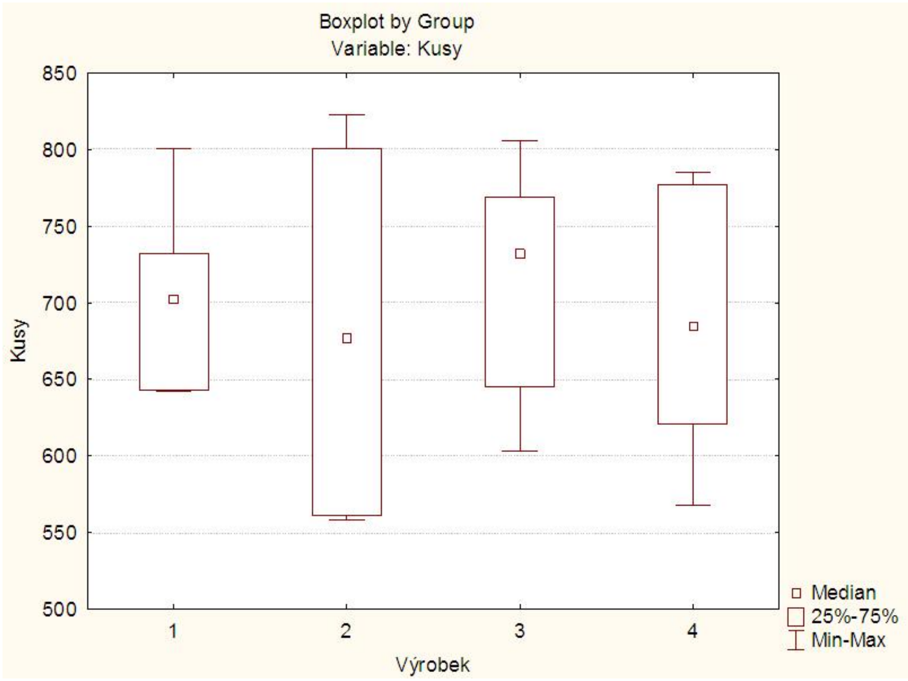
H_0 zamítáme na hladině významnosti α, pokud $Q_K \geq \chi^2_{1-\alpha}(r-1)$. Hodnota $\chi^2_{0,95}(3)$ je rovna číslu $7,81$, nemůžeme tedy na dané hladině významnosti zamítnout hypotézu o rovnosti mediánů počtu prodaných kusů jednotlivých výrobků.

Postup v programu STATISTICA:
Vytvoříme si datový soubor o 2 proměnných a 24 případech. Do první proměnné napíšeme počet prodaných kusů a do druhé čísla 1, 2, 3 nebo 4 podle toho, o který výrobek se jedná.

Dále v menu zvolíme *Statistics - Nonparametrics - Comparing multiple independent samples (groups)*. Ve *Variables* zvolíme za závislou proměnnou první proměnnou (Kusy) a za nezávislou proměnnou zvolíme druhou proměnnou (Výrobky). Test spustíme pomocí tlačítka *Summary: Kruskal-Wallis ANOVA & Median test*.

	Kruskal-Wallis ANOVA by Ranks; Kusy (Spreadsheet03) Independent (grouping) variable: Výrobek Kruskal-Wallis test: H (3, N= 24) =,5719153 p =,9028		
Depend.: Kusy	Code	Valid N	Sum of Ranks
1	1	6	78,50000
2	2	6	68,50000
3	3	6	84,00000
4	4	6	69,00000

První sloupec nám udává, o který druh výrobku se jedná, druhý sloupec udává počet hodnot v každém výběru a třetí sloupec obsahuje hodnoty veličin T_i. Nás ale nejvíce zajímá p-hodnota, kterou můžeme vidět v hlavičce tabulky. Protože $0,9028 \nleq 0,05$, nemůžeme na dané hladině významnosti zamítnout nulovou hypotézu o rovnosti mediánů.



Z krabicového grafu lze usoudit, že předpoklad o rovnosti mediánů všech výběrů je oprávněný.

Poznámka 5.2. Pokud bychom nulovou hypotézu o rovnosti mediánů všech výběrů zamítli, použili bychom metodu mnohonásobného porovnávání k rozhodnutí, které dvojice výběrů se od sebe významně liší. V systému STATISTICA se postupuje následovně: V menu zvolíme *Statistics - Nonparametrics - Comparing multiple independent samples (groups)*. Ve *Variables* zvolíme za závislou proměnnou první proměnnou (Kusy) a za nezávislou proměnnou zvolíme druhou proměnnou (Výrobky). Test spustíme pomocí tlačítka *Multiple comparisons of mean ranks for all groups*. V našem případě by výsledek vypadal takto:

	Multiple Comparisons p values (2-tailed); Kusy (Spreadsheet03) Independent (grouping) variable: Výrobky Kruskal-Wallis test: H (3, N= 24) =,5719153 p =,9028			
Depend.: Kusy	1 R:13,083	2 R:11,417	3 R:14,000	4 R:11,500
1		1,000000	1,000000	1,000000
2	1,000000		1,000000	1,000000
3	1,000000	1,000000		1,000000
4	1,000000	1,000000	1,000000	

Tabulka obsahuje p-hodnoty pro test hypotézy, že l-tý a k-tý výběr pocházejí z téhož rozložení. Vidíme, že žádná z p-hodnot není menší než hladina významnosti 0,05, což odpovídá našemu předchozímu výsledku, že nemůžeme zamítnout nulovou hypotézu o rovnosti mediánů.

5.3 Mediánový test

Nechť $Y_{i1}, \ldots, Y_{in_i}$, $i = 1, \ldots, r$ je r náhodných výběrů o rozsazích $n_1, \ldots, n_r$. Nechť všechny tyto výběry jsou na sobě nezávislé a pocházejí ze spojitého rozložení s distribuční funkcí $F_i, i = 1, \ldots, r$. Testujeme hypotézu H_0, že všechny tyto výběry pocházejí ze stejného rozložení, neboli

$$H_0 : F_1(x) = \cdots = F_r(x) \quad \text{pro } \forall x$$

Veličiny Y_{ij} vytvoří dohromady sdružený výběr o rozsahu $N = n_1 + \cdots + n_r$. Uspořádáme je do rostoucí posloupnosti, ve které mají veličiny $Y_{11}, \ldots, Y_{1n_1}$ pořadí $R_1, \ldots, R_{n_1}$, veličiny $Y_{21}, \ldots, Y_{2n_2}$ pořadí $R_{n_1+1}, \ldots, R_{n_1+n_2}$ a tak dále.

Mediánový test je založen na testové statistice

$$Q_M = 4\sum_{i=1}^{r} \frac{1}{n_i}(A_i - \frac{n_i}{2})^2 = 4\sum_{i=1}^{r} \frac{A_i^2}{n_i} - N,$$

kde

$$A_i = \sum_{j \in g_i} \frac{1}{2}\{sign[R_j - \frac{1}{2}(N+1)] + 1\} \quad \text{pro } i = 1, \ldots, r$$

udává počet veličin i-tého výběru, které jsou větší než medián sdruženého výběru. Příčemž se k tomuto počtu připočte $\frac{1}{2}$, pokud je N liché a medián patří do i-tého výběru.

Platí-li H_0, má Q_M asymptoticky rozložení $\chi^2(r-1)$, rostou-li rozsahy výběrů nade všechny meze. Tedy pro $Q_M \geq \chi^2_{1-\alpha}(r-1)$ zamítneme H_0 na hladině významnosti, která se s rostoucími n_i blíží číslu α.

5.4 Příklad

Spočítejme si stejný příklad jako v případě Kruskalova-Wallisova testu.

Řešení. Medián všech 24 hodnot je 696, 9167. V prvním výběru leží nad mediánem 3 hodnoty, ve druhém výběru 3 hodnoty, ve třetím výběru 4 hodnoty a ve čtvrtém výběru 2 hodnoty. Hodnota testové statistiky Q_M je tedy:

$$Q_M = 4\Big(\frac{3^2}{6} + \frac{3^2}{6} + \frac{4^2}{6} + \frac{2^2}{6}\Big) - 24 = 1,\overline{3}$$

H_0 zamítáme na hladině významnosti α, pokud $Q_M \geq \chi^2_{1-\alpha}(r-1)$. Hodnota $\chi^2_{0,95}(3)$ je rovna číslu $7,81$, nemůžeme tedy na dané hladině významnosti zamítnout hypotézu o rovnosti mediánů počtu prodaných kusů jednotlivých výrobků.

Postup v programu STATISTICA:
Postup je naprosto stejný jako u Kruskalova-Wallisova testu.

Dependent: Kusy	Median Test, Overall Median = 693,500; Kusy (Spreadsheet03) Independent (grouping) variable: Výrobek Chi-Square = 1,333333 df = 3 p = ,7212				
	1	2	3	4	Total
<= Median: observed	3,000000	3,000000	2,00000	4,00000	12,00000
expected	3,000000	3,000000	3,00000	3,00000	
obs.-exp.	0,000000	0,000000	-1,00000	1,00000	
> Median: observed	3,000000	3,000000	4,00000	2,00000	12,00000
expected	3,000000	3,000000	3,00000	3,00000	
obs.-exp.	0,000000	0,000000	1,00000	-1,00000	
Total: observed	6,000000	6,000000	6,00000	6,00000	24,00000

Jelikož p-hodnota není menší než hladina významnosti $(0,7212 \nleq 0,05)$, nemůžeme zamítnout nulovou hypotézu, že se od sebe mediány počtu prodaných kusů jednotlivých výrobků neliší.

5.5 Friedmannův test

Friedmannův test je neparametrickou obdobou analýzy rozptylu dvojného třídění.

Nechť Y_{ij} jsou nezávislé náhodné veličiny se spojitými distribučními funkcemi F_{ij}, $i = 1, \ldots, r$, $j = 1, \ldots, k$. Máme tedy celkem $r \cdot k$ pozorování na r objektech při k opakováních, které se nám rozpadají do r bloků s k prvky. Testujeme hypotézu H_0, že distribuční funkce $F_{i1}, \ldots, F_{ik}$ jsou totožné (F_{ij} nezávisí na j, zatímco na i záviset může).

Friedmannův test se používá k ověření shody úrovně určitého znaku. Důvodem jeho použití je snaha posoudit, zda úroveň sledovaného znaku závisí anebo nezávisí na změně podmínek. Na každém z r objektů (resp. bloků) je aplikováno k ošetření. V závislosti na výsledku určíme, zda ošetření mají na každý objekt stejný efekt. Někdy nejde o ošetření v pravém slova smyslu, ale nějaká veličina je na každém sledovaném objektu zaznamenávána v k časových okamžicích.

Použije se následující postup.
Pro každé i zvlášť se určí pořadí R_{ij} veličiny Y_{ij}. Jde tedy o určení pořadí uvnitř každého bloku, mezi veličinami $Y_{i1}, \ldots, Y_{ik}$.

	j			
i	1	2	...	k
1	Y_{11}	Y_{12}	...	Y_{1k}
2	Y_{21}	Y_{22}	...	Y_{2k}
...	...	...	...	...
r	Y_{r1}	Y_{r2}	...	Y_{rk}

$\Longrightarrow$

	j			
i	1	2	...	k
1	R_{11}	R_{12}	...	R_{1k}
2	R_{21}	R_{22}	...	R_{2k}
...	...	...	...	...
r	R_{r1}	R_{r2}	...	R_{rk}

Testová statistika má tvar

$$Q_F = \frac{12}{rk(k+1)} \sum_{j=1}^{k} (\sum_{i=1}^{r} R_{ij})^2 - 3r(k+1).$$

Platí-li H_0, má Q_F asymptoticky rozložení $\chi^2(k-1)$, rostou-li rozsahy výběrů nade všechny meze. Tedy pro $Q_F \geq \chi^2_{1-\alpha}(k-1)$ zamítneme H_0 na hladině významnosti, která se s rostoucím r blíží číslu α.

5.6 Příklad

Při procesu učení byl studován vliv adaptace na experimentální situaci. Deset pokusných osob vstupovalo do pokusu celkem čtyřikrát s dvoutýdenními odstupy. Úkol byl vždy jiný svým konkrétním obsahem, ale typ úkolu byl zachován. Na základě naměřených výsledků máme na hladině významnosti 0,05 zjistit, zda se učení v průběhu pokusu zlepšuje (učené osoby se „učí učit“) nebo zůstává na stejné úrovni.

	Opakování			
Osoba	1.	2.	3.	4.
1.	2	3	2	4
2.	4	3	6	6
3.	6	5	7	8
4.	2	1	4	3
5.	0	2	1	4
6.	10	8	9	11
7.	8	9	8	7
8.	5	6	4	7
9.	3	4	6	5
10.	7	6	5	4

Řešení. Testujeme hypotézu H_0, že výsledky pokusů se mění jen náhodně (tzn. nezáleží na tom, kolik pokusů osoba absolvuje).

Nejprve u každé sledované osoby zjištěné výsledky nahradíme jejich pořadím.

	Opakování			
Osoba	1.	2.	3.	4.
1.	1,5	3	1,5	4
2.	2	1	3,5	3,5
3.	2	1	3	4
4.	2	1	4	3
5.	1	3	2	4
6.	3	1	2	4
7.	2,5	4	2,5	1
8.	2	3	1	4
9.	1	2	4	3
10.	4	3	2	1
$\sum$	21	22	25,5	31,5

Pro $r = 10$ a $k = 4$ má testová statistika Q_F tvar:

$$Q_F = \frac{12}{10 \cdot 4 \cdot 5}\left(21^2 + 22^2 + 25,5^2 + 31,5^2\right) - 3 \cdot 10 \cdot 5 = 4,05$$

H_0 zamítáme na hladině významnosti α, pokud $Q_F \geq \chi^2_{1-\alpha}(k-1)$. Hodnota $\chi^2_{0,95}(3)$ je rovna číslu $9,488$, nemůžeme tedy na dané hladině významnosti zamítnout hypotézu, že se výsledky pokusů mění jen náhodně.

Postup v programu STATISTICA:

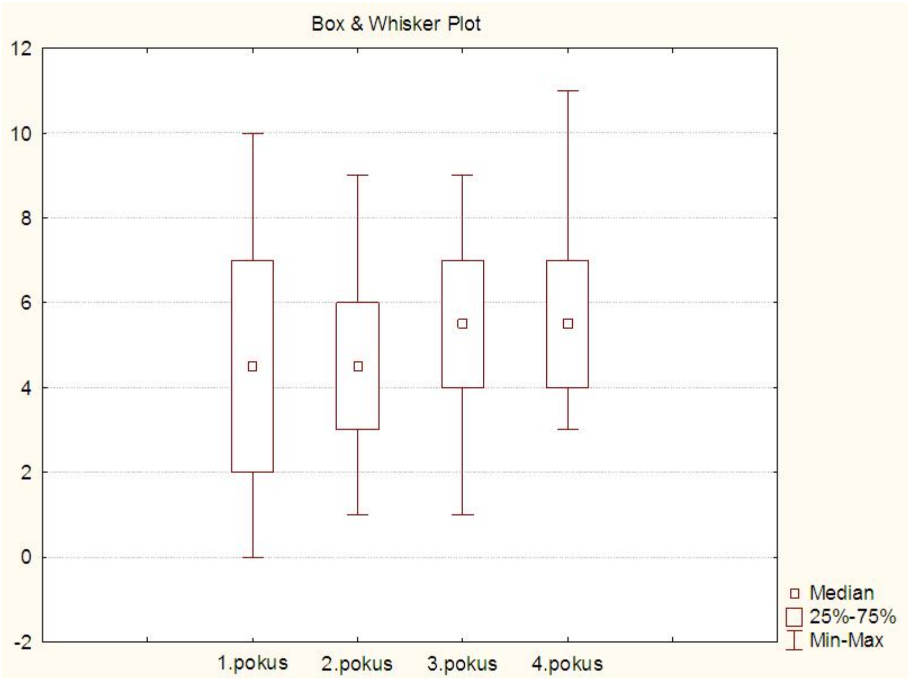
Vytvoříme si datový soubor o 4 proměnných a 10 případech. Každá proměnná odpovídá jednomu pokusu, takže do každého sloupce zapíšeme výsledky každé testované osoby odpovídající pokusům 1 až 4.

Dále v menu zvolíme *Statistics - Nonparametrics - Comparing multiple dependent samples (variables)*. Ve *Variables* zvolíme všechny proměnné (pokusy 1 až 4) a test spustíme pomocí tlačítka *Summary: Friedmann & Kendall's concordance.*

Friedman ANOVA and Kendall Coeff. of Concordance (Spreadsheet04)
ANOVA Chi Sqr. (N = 10, df = 3) = 4,175258 p = ,24315
Coeff. of Concordance = ,13918 Aver. rank r = ,04353

Variable	Average Rank	Sum of Ranks	Mean	Std.Dev.
1.pokus	2,100000	21,00000	4,700000	3,093003
2.pokus	2,200000	22,00000	4,700000	2,584140
3.pokus	2,550000	25,50000	5,200000	2,529822
4.pokus	3,150000	31,50000	5,900000	2,424413

V tabulce máme 4 sloupce. První nám určuje průměrné pořadí naměřených hodnot v každém pokusu ($R_{.j}/r$), druhý součet pořadí v každém pokusu ($R_{.j}$), třetí průměr naměřených hodnot v každém pokusu ($Y_{.j}/r$) a poslední směrodatnou odchylku. Nás ale nejvíce zajímá p-hodnota, kterou najdeme v hlavičce tabulky. Protože $0,24353 \nleq 0,05$, nemůžeme na dané hladině významnosti zamítnout nulovou hypotézu, že se výsledky pokusů mění jen náhodně.



Z krabicového grafu můžeme vyčíst, že se výsledky pokusů v průměru moc nemění, nemůžeme tedy říci, že se osoby „učí učit“)

Literatura

[1] Anděl, Jiří: *Základy matematické statistiky*, Matfyzpress, Praha 2005

[2] Blatná, Dagmar: *Neparametrické metody: testy založené na pořádkových a pořadových statistikách*, Vysoká škola ekonomická v Praze, Praha 1996

[3] Blatná, Dagmar: *Statistika a pravděpodobnost*, CERM, Brno 2003

[4] Budíková, Marie, Lerch, Tomáš, Mikoláš, Štěpán: *Základní statistické metody*, Masarykova univerzita, Brno 2005

[5] Hebák, Petr, Bílková, Diana, Svobodová Alžběta: *Praktikum k výuce matematické statistiky II*, Vysoká škola ekonomická v Praze, Praha 2000

[6] Sprent, Peter: *Applied nonparametric statistical methods*, Chapman & Hall, London 1993

Tabulky

Tabulka 1. - Koeficienty a_i Shaphirova-Wilkova testu a kvantily $sw_\alpha(n)$ pro $\alpha = 0,05$

i	$n = 5$	$n = 6$	$n = 7$	$n = 8$	$n = 9$	$n = 10$
1	0,6646	0,6431	0,6233	0,6052	0,5888	0,5739
2	0,2413	0,2816	0,3031	0,3164	0,3244	0,3291
3	0	0,0875	0,1401	0,1743	0,1976	0,2141
4	-	-	0	0,0561	0,0947	0,1224
$sw_{0,95}(n)$	0,762	0,788	0,803	0,818	0,829	0,842

Zdroj:
http://www.kmt.zcu.cz/person/Kohout/info_soubory/letnisem/tabulky.htm

Tabulka 2. - Kvantily u_α standardizovaného normálního rozložení

α	u_α	α	u_α
0,950	1,64485	0,970	1,88079
0,951	1,65463	0,971	1,89570
0,952	1,66456	0,972	1,91104
0,953	1,67466	0,973	1,92684
0,954	1,68494	0,974	1,94313
0,955	1,69540	0,975	1,95996
0,955	1,70604	0,976	1,97737
0,957	1,71689	0,977	1,99539
0,958	1,72793	0,978	2,01409
0,959	1,73920	0,979	2,03352
0,960	1,75069	0,980	2,05375
0,961	1,76241	0,981	2,07485
0,962	1,77438	0,982	2,09693
0,963	1,78661	0,983	2,12007
0,964	1,79912	0,984	2,14441
0,965	1,81191	0,985	2,17009
0,966	1,82501	0,986	2,19729
0,967	1,83842	0,987	2,22621
0,968	1,85218	0,988	2,25713
0,969	1,86630	0,989	2,29037

Zdroj: [4]

Tabulka 3. - Kritické hodnoty znaménkového testu

	$\alpha = 0,05$	
n	k_1	k_2
6	0	6
7	0	7
8	0	8
9	1	8
10	1	9
11	1	10
12	2	10
13	2	11
14	2	12
15	3	12
16	3	13
17	4	13
18	4	14
19	4	15
20	5	15

Zdroj: [4]

Tabulka 4. - Kritické hodnoty jednovýběrového Wilcoxonova testu

	$\alpha = 0,05$
n	kritická hodnota
6	0
7	2
8	3
9	5
10	8
11	10
12	13
13	17
14	21
15	25
16	29
17	34
18	40
19	46
20	52
21	58
22	65
23	73
24	81
25	89
26	98
27	107
28	116
29	126
30	137

Zdroj: [4]

Tabulka 5. - Kritické hodnoty dvouvýběrového Wilcoxonova testu

	n									
m	1	2	3	4	5	6	7	8	9	10
1	-									
2	-	-								
3	-	-	-							
4	-	-	-	0						
5	-	-	0	1	2					
6	-	-	1	2	3	5				
7	-	-	1	3	5	6	8			
8	-	0	2	4	6	8	10	13		
9	-	0	2	4	7	10	12	15	17	
10	-	0	3	5	8	11	14	17	20	23
11	-	0	3	6	9	13	16	19	23	26
12	-	1	4	7	11	14	18	22	26	29
13	-	1	4	8	12	16	20	24	28	33
14	-	1	5	9	13	17	22	26	31	36
15	-	1	5	10	14	19	24	29	34	39

Zdroj: [4]

Tabulka 6. - Kvantily $\chi^2_\alpha(k)$ Pearsonova rozložení

k	$\alpha = 0,95$
1	3,841
2	5,991
3	7,815
4	9,488
5	11,070
6	12,592
7	14,067
8	15,507
9	16,919
10	18,307
11	19,675
12	21,026
13	22,362
14	23,685
15	24,996

Zdroj: [4]