S 8 Početní postupy s procenty, Kruskal-Wallisův test, Mann-Whitneyův test, Wilcoxonův test

1. Početní postupy s procenty

TEORIE

Předpokladem je, že n je větší než 20 (je zřejmé, že procentní počet získaný z šetření méně než 20ti osob je nespolehlivým údajem)

$$\% = \frac{b}{n}n$$

b= část souboru, kterou chceme vyjádřit v procentech

Interval spolehlivosti pro procentový údaj:

Výpočet provádíme z hodnot výběrového procenta, který chceme zevšeobecnit a z rozsahu výběru. V úvahu bereme pravděpodobnost, se kterou budeme šíři intervalu posuzovat. Interval spolehlivosti je dán vztahem:

$$IS(\%) = p_v \pm t_p \sqrt{\frac{p_v(100 - p_v)}{n}}$$

p_v= výběrové procento t_p= pravděpodobnostní veličina při 99% = 2,58 a 95% = 1,96

PŘÍKLAD

Příslušnicí vězeňské služby (n=40) splnili výkonnostní limit ve vytrvalostním běhu v počtu 30 osob. Zajímá nás kolik je to procent.

$$\% = \frac{30}{40}100 = 75\%$$

Vypočítali jsme tedy, že výkonnostní limit ve vytrvalostním běhu splnilo 75 % příslušníků vězeňské služby. Chceme zjistit interval, ve kterém se nalézá neznámé procento všech příslušníků vězeňské služby v ČR (základního souboru).

$$IS(75\%) = 75 \pm 1{,}96 \sqrt{\frac{75(100 - 75)}{40}} = 75 \pm 13{,}419$$

Interval spolehlivosti pro 75% je s pravděpodobností 95%v rozsah 61,6-88,4%

2. Neparametrické varianty T testu

- Mann – Whitney U test,
- Kruskal – Wallisův test
- Wilcoxonův test

2.1 Mann – Whitney U test

Tento test je obdobou parametrického T testu pro nezávislé výběry (seminář 3), na rozdíl od něj však pracuje s neparametrickými daty, popřípadě pokud soubory s parametrickými daty nevykazují normalitu rozdělení četností. Testovány jsou hypotézy o mediánu.

Testované hypotézy jsou následující:
H_0 : Mediány obou souborů se rovnají.
H_1 : Mediány obou souborů jsou odlišné.

Testové kritérium se počítá ze vztahů:

$$U_1 = S_1 - \frac{n_1(n_1+1)}{2} \qquad U_2 = S_1 - \frac{n_2(n_2+1)}{2}$$

$U_{1,2}$ …součty pořadí v jednotlivých skupinách

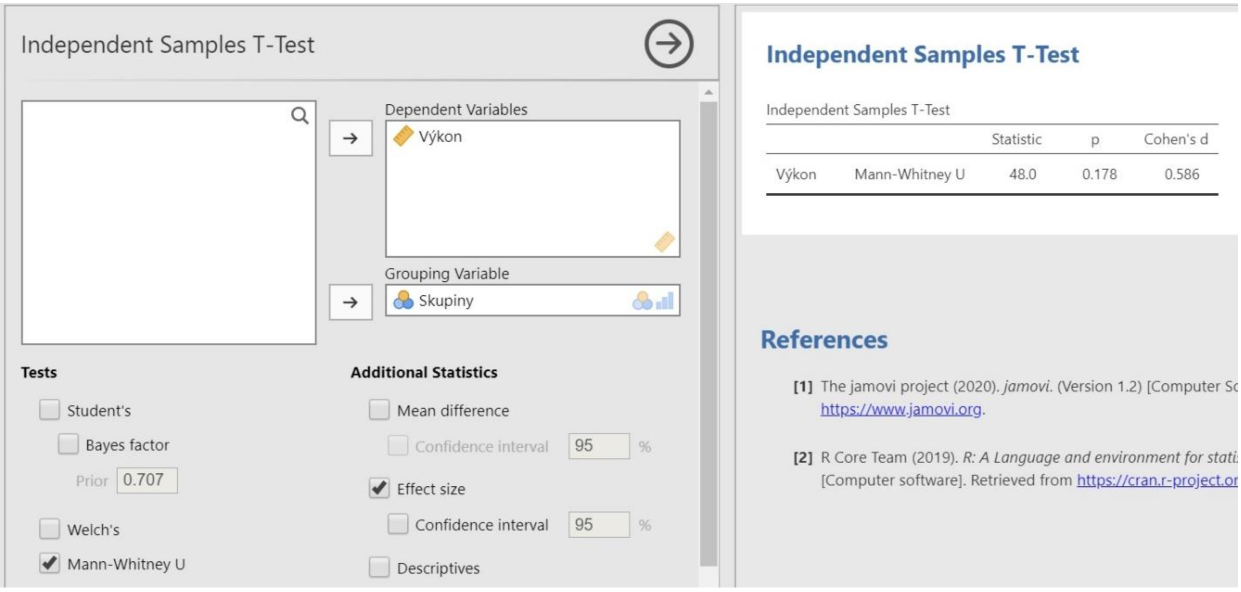
PŘÍKLAD

V tabulce jsou uvedeny výkony randomizací vybraných členů dvou skupin ve vrhu koulí. Zjistěte, zda je mezi skupinami statisticky významný rozdíl.

Skupina1 (cm)	Skupina 2 (cm)
886	776
992	547
997	887
857	993
654	569
534	449
765	943
458	659
991	499
667	668
994	865
995	599

Po testu normality rozdělení četností (viz seminář 3) jsme zjistili že soubor A vykazuje porušení normality rozdělení. (Shapiro- Wilk p=0,042). Proto je k testování hypotézy o rozdílu mezi nezávislými soubory potřeba použít Mann-Whitney test.
Po zadání dat pokračujeme shodně jako při volbě T testu (seminář 3), tedy:
Analyses → T-Tests – Independent samples T- Test

Při následné volbě zaškrtneme pole *Man-Whitney test.*



$p > 0{,}05$, nulovou hypotézu nemůžeme zamítnout, mezi skupinami není statisticky významný rozdíl.

Pozn. Protože soubory byly náhodně vybrány ze základního souboru, věcnou významnost v tomto případě nepočítáme. Pokud by toto bylo potřeba, zaškrtneme příslušnou volbu pro výpočet Cohenova d. Viz výsledky.

2.2 Kruskal – Wallisův test

Tento test je rozšířením předchozího Mann- Whitney testu na více jak 2 skupiny. Měrná stupnice musí být přinejmenším ordinální, všechny hodnoty jsou zjištěny u náhodných výběrů.

Testovým kritériem je hodnota H, která se vypočítá podle vzorce

$$H = \left[\frac{12}{N(N+1)}\sum \frac{R_i^2}{n_i}\right] - 3(N+1)$$

N = celková četnost všech hodnot
R_i = součet pořadí v jednotlivých skupinách
n_i = četnosti hodnot v jednotlivých skupinách

PŘÍKLAD

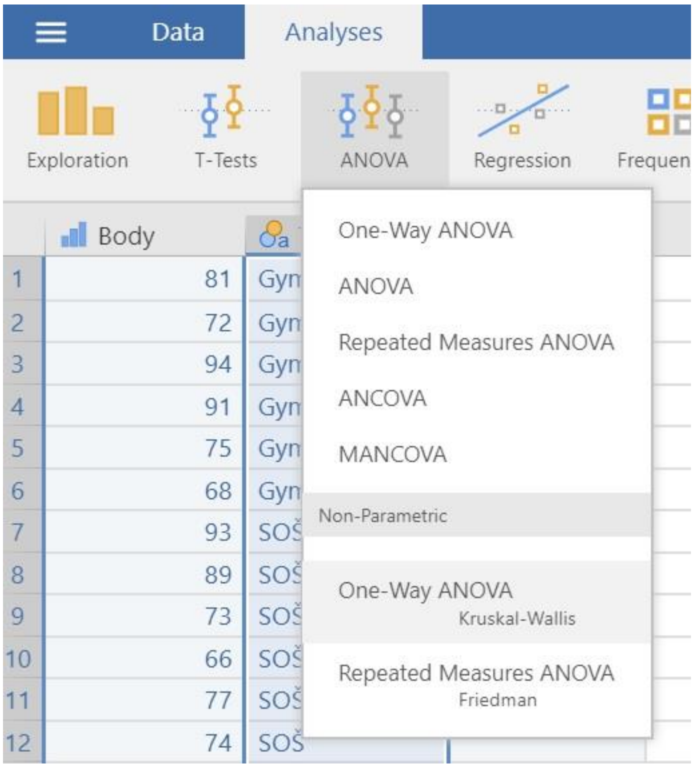
Pro přijímací řízení uchazečů bakalářského studijního programu, oboru TVS, je zařazen písemný test z problematiky všeobecného přehledu v oblasti tělesné kultury a sportu. Chceme posoudit, zda se výsledky testu významně liší podle typu škol, ze kterých se uchazeč na obor hlásí. Náhodně vybereme z jednotlivých typů škol (Gymnázia, SOŠ, SOU) 6 uchazečů. Hladinu významnosti jsme stanovili na 0,05%

Dosažené výsledky (počet bodů) podle typu škol:

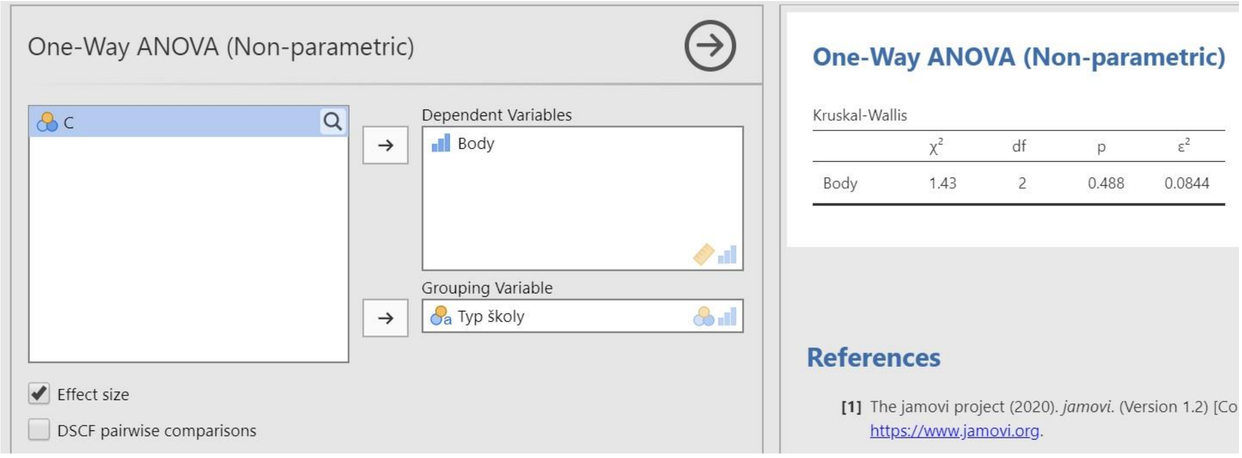
Uchazeč	Gymnázium	SOŠ	SOU
A	81	93	58
B	72	89	66
C	94	73	85
D	91	66	91
E	75	77	71
F	68	74	73

Řešení: Zadáme hodnoty do dvou sloupců a označíme příslušné skupiny a typ dat (body – ordinální a typ školy – nominální). Pokračujeme přes volby:

Analyses → ANOVA →One-Way ANOVA Kruskal-Wallis



Výsledky:



Hodnota $p > 0{,}05$ = mezi skupinami neexistuje statisticky významný rozdíl.
Effect size $\varepsilon^2 < 0{,}1$ = mezi skupinami není věcně významný rozdíl

2.3 Wilcoxonův test

Tento test je neparametrickou obdobou párového T testu (T test pro závislé soubory – viz seminář 4)

Představme si, že testujeme, zda došlo po aplikaci tréninkového plánu ke zlepšení silových schopností testovaných probandů. Provedeme úvodní testování (např. vrh koulí), následně budeme aplikovat tréninkový plán a po jeho aplikaci opětovně provedeme testování. Jak již bylo zmíněno, je důležité, aby byl počet probandů u prvního a druhého měření stejný, tedy ty probandy, kteří se neúčastnili obou měření, je nutné vyloučit. Nyní je důležité zjistit, zda lze výkony ve vrhu koulí posuzovat jako normálně rozdělené (viz seminář 3, Shapiro - Wilkův test).
V případě, že lze výkony posoudit jako normálně rozložené, použili bychom párový T-test (seminář 4). **Wilcoxonův test používáme v případě, kdy nelze usuzovat na normální rozdělení hodnot.**

(Stejný test použijeme i v případě že budeme zpracovávat neparametrická data – obvykle ordinální).

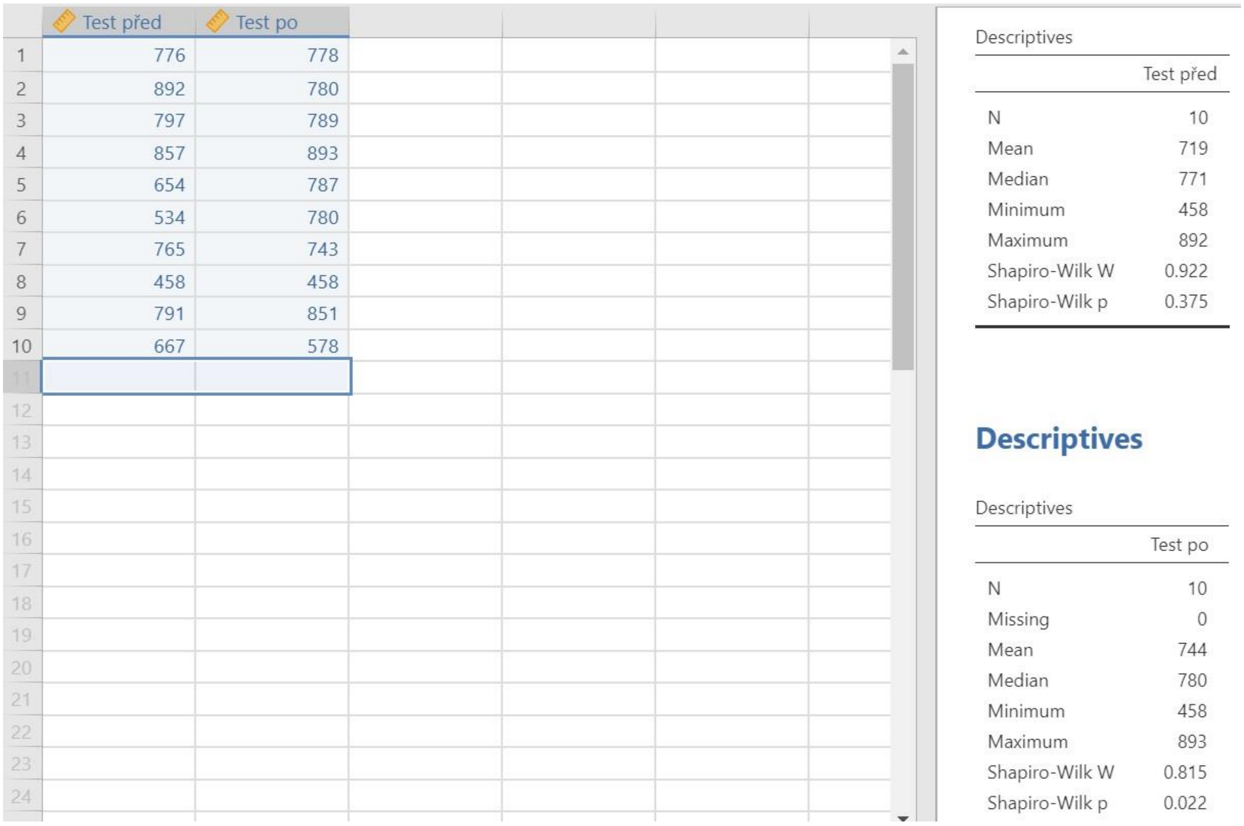
V párovém T testu se nulová a alternativní hypotéza vztahují k průměru. U Wilcoxonova párového testu se hypotézy vztahují k mediánu.

PŘÍKLAD

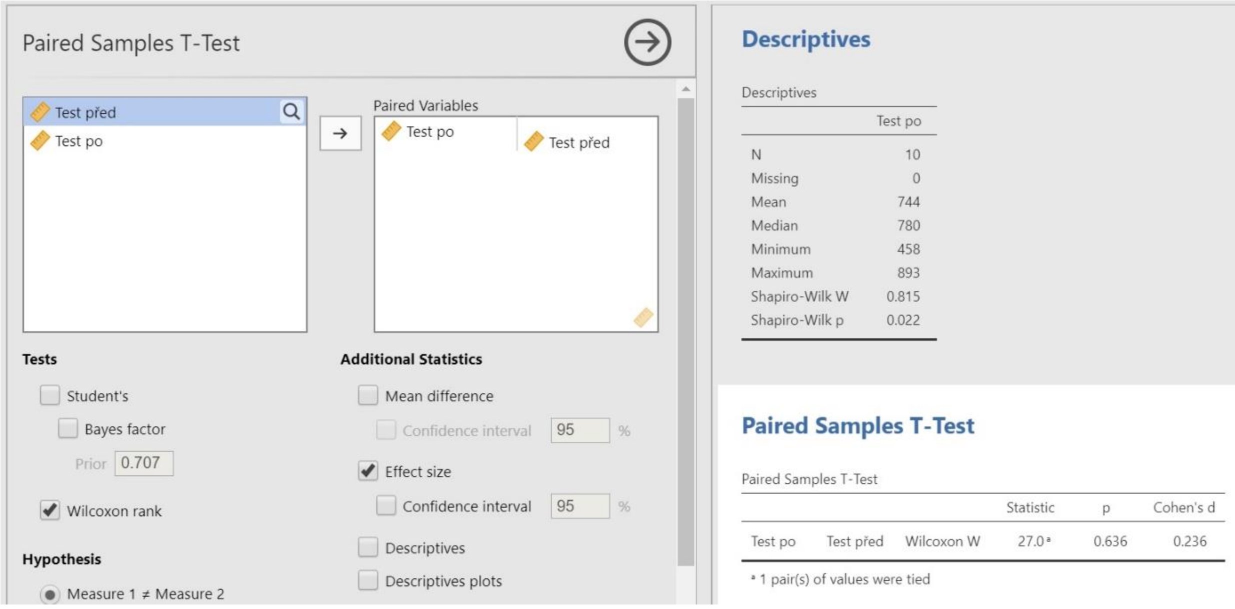
Data z výše uvedeného příkladu jsou uvedena v následující tabulce. Ověřte, zda existuje statisticky významný rozdíl ve výkonech před a po aplikaci tréninkového plánu na rozvoj explozivních silových schopností u osmi náhodně vybraných probandů.

Měření před (cm)	Měření po (cm)
776	772
892	947
797	687
857	893
654	769
534	549
765	743
458	359
791	851
667	578

Po analýze dat (*Analyses → Exploration, → Descriptives)* detekujeme porušení normality dat u souboru dat z druhého měření ($p < 0,05$).



Pro výpočet tedy použijeme Wilcoxonův test. *Analyses → T- Tests → Paired Samples T-test.* Zaškrtneme volbu *Wilcoxon rank.*



Výsledná hodnota p (0,636) >0,05. Rozdíl mezi výkony před a po aplikaci tréninkového plánu je tedy statisticky nevýznamný. Stimulace silových schopností se ukázala jako neúčinná.

V případě že by naším úkolem bylo stanovit věcnou významnost, využili bychom k interpretaci Cohenovo d.

ÚKOLY

1) V předmětu „Rozvoj pohybových schopností" absolvovali studenti v rámci kontroly studia závěrečný test. Chceme posoudit, zda se výsledky testu liší podle oboru studia. Náhodně bylo vybráno 10 studentů z každého studijního oboru. V tabulce je u každého z nich uvedena bodová hodnota, kterou v testu dosáhl. Rozhodněte, zda je mezi studijními obory statisticky významný rozdíl v úrovni vědomostí učiva daného předmětu.

TVS prezenční	TVS kombi	Učitelství ZŠ	Učitelství SŠ
19	27	13	30
25	30	23	23
18	22	24	31
18	29	30	28
15	22	11	22
24	21	21	20
29	24	20	13
16	13	21	24
23	22	15	25
13	15	28	15