

S 3 Testování statistických hypotéz – nezávislé výběry

testování hypotéz o průměru dvou nezávislých souborů

1. T – test pro nezávislé výběry, jestliže se rozptyly rovnají ($\sigma_1^2 = \sigma_2^2$)

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n-1} + \frac{s_2^2}{n-1}}}$$

2. T – test pro nezávislé výběry, jestliže se rozptyly nerovnají ($\sigma_1^2 \neq \sigma_2^2$) (Welshův test)

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{n_1 s_1^2 + n_2 s_2^2}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}$$

Obecná charakteristika jednotlivých etap:

- posouzení smysluplnosti aplikace statistických metod
- přesná formulace nulové hypotézy H_0
- zvolení hladiny významnosti (úzus 0,05)
- výpočet hodnoty statistického testu
- výpočet hodnoty p (*pravděpodobnost výsledků které ještě hovoří proti H_0*
pokud je hodnota větší nebo rovna 0,05 tak H_0 nemůžeme zamítnout (software, včetně Jamovi))
- posouzení statistické významnosti (je-li to naším cílem)
- posouzení věcné (praktické) významnosti
- interpretace výsledků

PŘÍKLAD 1

Ruční dynamometrií jsme měřili sílu stisku ruky u dvou **výběrových** souborů mužů: učitelské ($n_1 = 20$) a neučitelské ($n_2 = 30$) skupiny. Proved'te srovnání obou skupin. Vypočítali jsme tyto hodnoty (ze zdrojových dat z následné tabulky)

$$\begin{array}{lll} n_1 = 20 & \bar{x}_1 = 60 & s_1 = 3,2 \\ n_2 = 30 & \bar{x}_2 = 66 & s_2 = 3,6 \end{array}$$

Zdrojová data

Proband č.	Výkon	Stud.skupina	Proband č.	Výkon	Stud.skupina
1	60	učitel.	1	65	neučitel.
2	62	učitel.	2	66	neučitel.
3	58	učitel.	3	67	neučitel.
4	61	učitel.	4	65	neučitel.
5	60	učitel.	5	66	neučitel.
6	55	učitel.	6	70	neučitel.
7	65	učitel.	7	64	neučitel.
8	62	učitel.	8	59	neučitel.
9	61	učitel.	9	63	neučitel.
10	57	učitel.	10	72	neučitel.
11	64	učitel.	11	70	neučitel.
12	60	učitel.	12	66	neučitel.
13	61	učitel.	13	61	neučitel.
14	59	učitel.	14	72	neučitel.
15	58	učitel.	15	65	neučitel.
16	60	učitel.	16	67	neučitel.
17	52	učitel.	17	67	neučitel.
18	66	učitel.	18	63	neučitel.
19	59	učitel.	19	65	neučitel.
20	60	učitel.	20	66	neučitel.
			21	59	neučitel.
			22	70	neučitel.
			23	67	neučitel.
			24	68	neučitel.
			25	60	neučitel.
			26	61	neučitel.
			27	66	neučitel.
			28	69	neučitel.
			29	70	neučitel.
			30	70	neučitel.



Postup výpočtu v programu Jamovi.

Vložíme data z obou skupin do sloupce A, označíme je jako metrická (*Continuous*)

The screenshot shows the Jamovi software interface. At the top, there are two tabs: 'Data' and 'Analyses'. Below the tabs is a toolbar with various icons for data manipulation: Paste, Edit, Setup, Compute, Transform, Add, Delete, Filters, and Rows. The 'Data Variable' configuration panel is open, showing the variable name 'A', a description field, and the following settings:

- Measure type: Continuous
- Data type: Integer (auto)
- Missing values: (empty field)

Below the configuration panel is a data table with three columns labeled A, B, and C. The data is as follows:

	A	B	C
16	60	učitel	
17	52	učitel	
18	66	učitel	
19	59	učitel	
20	60	učitel	
21	65	neučitel	
22	66	neučitel	
23	67	neučitel	
24	65	neučitel	
25	66	neučitel	
26	70	neučitel	

At the bottom of the interface, there is a status bar showing: Ready, Filters 0, Row count 50, Filtered 0, Deleted 0, Added 50, Cells edited 150.

Do sloupce B zadáme příslušnost ke studijní skupině. Data označíme jako *Nominální*, typ dat *Text*.

The screenshot shows the 'DATA VARIABLE' dialog box for variable 'B'. The 'Measure type' is set to 'Nominal' and the 'Data type' is set to 'Text'. The 'Missing values' field is empty. On the right, the 'Levels' list contains 'učitel' and 'neučitel'.

Pomocí funkcí *Analyses* → *Exploration* → *Descriptives* přiřadíme data do položky *Variables* (proměnné, v našem příkladě jednotlivé výkony) a třídy (v našem příkladě učitel/neučitel) do položky (*Split by*)

The screenshot shows the 'Descriptives' dialog box. The 'Variables' list contains 'A' and the 'Split by' list contains 'B'. A search icon is visible in the top right corner of the dialog box.

V záložce *Statistics* zvolíme následné položky popisné statistiky včetně testu normality – *Shapiro- Wilkův test*.

The image shows the SPSS 'Descriptives' dialog box on the left and the resulting 'Descriptives' output window on the right. The dialog box has several sections with checkboxes:

- Sample Size:** N, Missing
- Percentile Values:** Quartiles, Cut points for: 4 equal groups
- Dispersion:** Std. deviation, Variance, Range, Minimum, Maximum, S. E. Mean
- Central Tendency:** Mean, Median, Mode, Sum
- Distribution:** Skewness, Kurtosis
- Normality:** Shapiro-Wilk

The output window shows the following data:

Descriptives	B		A	
	učitel	neučitel	učitel	neučitel
N	20	30		
Mean	60.0	66.0		
Median	60.0	66.0		
Standard deviation	3.21	3.61		
Shapiro-Wilk W	0.955	0.954		
Shapiro-Wilk p	0.454	0.214		

Ve výsledcích v pravé části zkontrolujeme výslednou hodnotu p Shapiro- Wilkova testu. Pokud je vyšší než 0,05 u obou skupin, konstatujeme normální rozdělení četností, což nás opravňuje k použití zamýšleného parametrického T testu pro nezávislé výběry. V opačném případě bychom použili neparametrickou obdobu T testu – tedy Mann- Whitney U test (viz seminář 9).

Pozn. Předpoklad, že oba výběry pocházejí z normálního rozdělení, nemusí být za každou cenu dodržen. Dle Sebery (2014), T-test totiž pracuje s průměry obou výběrů, a ty již při rozsahu výběru v řádu desítek mají přibližně normální rozdělení díky centrální limitní větě.

Výpočet hodnoty t testu pro nezávislé výběry: postupujeme přes nabídky *Analyses* → *T – tests* → *Independent Samples T - test*

The image shows the SPSS 'Analyses' menu with the 'T-Tests' option selected. A dropdown menu is open, showing three options: 'Independent Samples T-Test', 'Paired Samples T-Test', and 'One Sample T-Test'. In the background, the 'Variables' list shows variable 'A' and the 'Split by' list shows variable 'B'.

Zvolíme následné položky a v pravém okně budou zobrazeny výsledky:

The image shows the SPSS 'Independent Samples T-Test' dialog box on the left and the corresponding output window on the right. In the dialog box, 'A' is selected as the dependent variable and 'B' as the grouping variable. Under 'Tests', 'Student's t' is checked with an effect size of 0.707. Under 'Assumption Checks', 'Homogeneity test' is checked. The output window displays the following tables:

Independent Samples T-Test					
	Statistic	df	p	Cohen's d	
A	Student's t	-5.57	48.0	<.001	-1.72

Homogeneity of Variances Test (Levene's)				
	F	df	df2	p
A	0.765	1	48	0.386

Note: A low p-value suggests a violation of the assumption of equal variances.

Výsledná hodnota p pro T test je menší než 0,001. Srovnáním hodnoty p (s hladinou významnosti 0,05) tak zamítáme nulovou hypotézu H_0 a usuzujeme na statisticky významný rozdíl mezi oběma výběry.

Zatržením volby *Homogeneity test* jsme otestovali, zda se rozptyly obou souborů rovnají, či nikoliv. V tomto případě je hodnota p (0,386) pro Levenův test vyšší než 0,05 což indikuje rovnost rozptylů. Použití T testu pro nezávislé soubory s rovností rozptylů tak bylo korektní. V případě opačném, tedy pokud by hodnota p byla menší než 0,05 konstatovali bychom nerovnost rozptylů. V tomto případě se pro testování hypotézy o průměru používá T test pro nezávislé soubory s nerovností rozptylů, tzv. Welschův test (viz volba v předchozím obrázku). Tuto variantu nám přiblíží příklad 2 v tomto semináři.

Teorie

Věcná (praktická) významnost.

Doposud výzkumní pracovníci hodnotili věcnou významnost výhradně v naměřených jednotkách např. v cm, sekundách, bodech apod., což je i nadále nutné. Současně se však užívají statistické koeficienty „effect size“ které určují podíl „vysvětleného rozptylu“. Jsou to koeficienty, které budeme považovat za obsahově podstatné v relaci k ostatním nesledovaným vlivům a zpravidla jsou uvedeny v procentech.

Pro posouzení věcné významnosti máme k dispozici minimálně tři dostupné nástroje:

1. Statistickou významnost na určené hladině významnosti, zpravidla $p = 0,05$
2. Logický úsudek, kdy předem stanovíme minimální hodnotu velikosti v jednotkách měření
3. Stanovení procenta velikosti účinku „effect size“

Zpracováno volně dle Blahuše, (2000)

Postup výpočtu věcné (praktické) významnosti (effect size, ω^2)

vypočítá se podle vzorce: $\omega^2 = \frac{t^2 - 1}{t^2 + n_1 + n_2 - 1}$

$$\omega^2 = \frac{-5,97^2 - 1}{-5,97^2 + 20 + 30 - 1} = 0,409$$

Výsledek je větší než 0,1 a proto je sledovaný rozdíl věcně (prakticky) významný. Znamená to, že rozdíl ve výkonu mezi dvěma skupinami je ze 41 % ovlivněn příslušností ke studijní skupině. Jinými, zpravidla neznámými faktory je ovlivněno 59 % rozdílu.



Mezi další z použitelných koeficientů věcné významnosti patří Cohenovo d . Výpočet tohoto koeficientu je součástí programu Jamovi. V nabídce pro výpočet hodnoty T testu zaškrtneme volbu *Effect size*. Hodnota Cohenova d je poté uvedena ve výsledkové části. Viz předchozí obrázek z programu Jamovi. Výslednou uváděnou hodnotu interpretujeme v absolutní hodnotě, tedy v předchozím konkrétním příkladu se jedná o hodnotu 1,72. Rozdíl mezi skupinami tak dle následujícího schématu hodnotíme slovně jako velmi velký.

<i>Velikost efektu</i>	<i>d</i>	<i>Zdroj</i>
Velmi malý	0,01	Sawilowsky, 2009
Malý	0,20	Cohen, 1988
Střední	0,50	Cohen, 1988
Velký	0,80	Cohen, 1988
Velmi velký	1,20	Sawilowsky, 2009
Obrovský	2,00	Sawilowsky, 2009

PŘÍKLAD 2

Náhodné výběry žen studijních skupin Tv-Čj ($n_1 = 20$) a Tv-Z ($n_2 = 30$) dosáhly těchto průměrných výkonů ve vertikálním výskoku (Sargentův test).

$$n_1 = 20 \quad \bar{x}_1 = 62,2 \quad s_1 = 9,2$$

$$n_2 = 30 \quad \bar{x}_2 = 65,2 \quad s_2 = 13,6$$

Proveďte srovnání obou skupin. Zdrojová data jsou uvedena v následné tabulce.

n_1	Výkon	Studijní aprobace	n_2	Výkon	Studijní aprobace
1	60	Tv-Čj	1	71	Tv-Geo
2	65	Tv-Čj	2	62	Tv-Geo
3	62	Tv-Čj	3	60	Tv-Geo
4	45	Tv-Čj	4	58	Tv-Geo
5	62	Tv-Čj	5	60	Tv-Geo
6	80	Tv-Čj	6	77	Tv-Geo
7	65	Tv-Čj	7	77	Tv-Geo
8	65	Tv-Čj	8	77	Tv-Geo
9	65	Tv-Čj	9	77	Tv-Geo
10	65	Tv-Čj	10	77	Tv-Geo
11	65	Tv-Čj	11	56	Tv-Geo
12	70	Tv-Čj	12	63	Tv-Geo
13	70	Tv-Čj	13	85	Tv-Geo
14	73	Tv-Čj	14	55	Tv-Geo
15	65	Tv-Čj	15	45	Tv-Geo
16	56	Tv-Čj	16	65	Tv-Geo
17	55	Tv-Čj	17	44	Tv-Geo
18	68	Tv-Čj	18	36	Tv-Geo
19	45	Tv-Čj	19	45	Tv-Geo
20	43	Tv-Čj	20	81	Tv-Geo
			21	55	Tv-Geo
			22	60	Tv-Geo
			23	77	Tv-Geo
			24	65	Tv-Geo
			25	75	Tv-Geo
			26	80	Tv-Geo
			27	45	Tv-Geo
			28	70	Tv-Geo
			29	93	Tv-Geo
			30	64	Tv-Geo



Při výpočtu v programu Jamovi postupujeme shodně jako v předchozím případě. Ve výsledkové části (viz obr.) můžeme vidět upozornění, že rozptyly obou souborů se nerovnájí. (*Levene's test is significant ($p < .05$), suggesting a violation of the assumption of equal variances*). Tuto skutečnost můžeme, popřípadě ověřit i zaškrtnutím volby *Homogeneity test* s výsledkem hodnoty $p = 0,041$.

V takovém případě Studentova t testu volíme Welchův test. p hodnota = 0,373 je větší než 0,05 nemůžeme tak zamítnout nulovou hypotézu, soubory se neliší. Z tohoto důvodu nestanovujeme ani významnost věcnou.

Independent Samples T-Test

Grouping Variable: B

Tests

- Student's
- Bayes factor
- Welch's
- Mann-Whitney U

Additional Statistics

- Mean difference
- Confidence interval: 95 %
- Effect size
- Confidence interval: 95 %
- Descriptives
- Descriptives plots

Hypothesis

- Group 1 \neq Group 2
- Group 1 $>$ Group 2
- Group 1 $<$ Group 2

Missing values

- Exclude cases analysis by analysis
- Exclude cases listwise

Assumption Checks

- Homogeneity test
- Normality test
- Q-Q plot

Independent Samples T-Test Results

	Statistic	df	p
Student's t	-0.834*	48.0	0.408
Welch's t	-0.899	48.0	0.373

* Levene's test is significant ($p < .05$), suggesting a violation of the assumption of equal variances.

Assumptions

Homogeneity of Variances Test (Levene's)

	F	df	df2	p
A	4.43	1	48	0.041

Note. A low p-value suggests a violation of the assumption of equal variances.

ÚKOL

Je statisticky významný rozdíl v hodnotách startovní reakce vrcholových sprinterů a sprinterek? (Je hodnota startovní reakce ovlivněna pohlavím?)

Jako vstupní data použijte startovní reakce závodníků v **rozběhích** na atletickém mistrovství světa v Dauhá 2019. Proved'te náhodný výběr 15 mužů a 15 žen. Data naleznete na:

<https://worldathletics.org/competitions/world-athletics-championships/world-athletics-championships-budapest-2023-7138987/timetable/bydiscipline>

I v případě detekce odchylky od normality rozdělení použijte z cvičných důvodů T test.