

Statistická analýza dat – vzorové zpracování

Statistics Calculators k volnému použití

<https://www.socscistatistics.com/tests/>

(hrubá data s tečkou)

<https://www.statskingdom.com/>

(vkládá se M , SD , n)

<https://effect-size-calculator.herokuapp.com/>

https://www.psychometrica.de/effect_size.html

(hrubá data s desetinnou tečkou).

Tab. 1: Síla stisku dominantní (DH) a nedominantní ruky (NDH)

Č.	DH	NDH
1	26.8	21.1
2	33.3	32.4
3	17.6	17.6
4	26.8	26.5
5	24.9	21.6
6	28.9	25.1
7	21.2	16.7
8	36.0*	28.8
9	20.1	16.5
10	24.5	17.6
11	20.1	17.6
12	24.6	19.7
13	25.4	24.5
14	24.4	22.2
15	29.4	27.4
M	25.6	22.35
SD	4.95	4.93

* odlehlá hodnota

Postup (statistická analýza dat)

1. Ověření normality dat
2. Výpočet základních statistických charakteristik
3. Testování statistických hypotéz (t-test, U test)
4. Korelační počet (Pearson, Spearman)
5. Excel ... (viz přednášky Statistika 1 +2)

Statistics Calculators (preferovaných)

<https://www.statskingdom.com/>

Další možné a ověřené Calculators:

<https://www.socscistatistics.com/>

https://www.psychometrica.de/effect_size.html

<https://effect-size-calculator.herokuapp.com/>

<https://lbecker.uccs.edu/>

VÝPOČTY VZOROVÝCH DAT (SOUBORY HGS DOM/NONDOM, TABULKA 1)

Výpočty pomocí: <https://www.statskingdom.com/>

1. Ověření normality (např. Kolmogorov-Smirnov, Shapiro-Wilk test, Lilliefors,)

Kolmogorov-Smirnov test (K-S)

K-S test je metoda umožňující testovat, zda jedna jednorozměrná proměnná má předpokládané (nejčastěji normální) rozdělení četností (*pro malé soubory <50*). Limitací K-S testu je jeho *vysoká citlivost na extrémní hodnoty*. Test K-S má nízkou sílu a *nemělo by se o něm vážně uvažovat při testování normality* (Ghasemi, 2012).

Shapiro-Wilk test je založen na korelaci mezi daty a *poskytuje lepší výkon než K-S test*. S-W test je doporučován jako *nejlepší volba pro testování normality dat* (Ghasemi, 2012).
<https://www.statskingdom.com/shapiro-wilk-test-calculator.html>

Volba pro DH: S-W test; Significance level (α): 0,05; Outliers: include; Rounding: 2
Výpočet (Calculate):

Reporting Shapiro-Wilk test in APA Format

The S-W test did not show a significant departure from normality, $W(15) = .96$, $p = .699$ (*S-W test neprokázal významný odklon od normality, $W(15) = 0,96$, $p = 0,699$*). Outliers: 36

1. H0 hypothesis. Since $p\text{-value} > \alpha$, we accept the H0. It is assumed that the data is normally distributed. (*Protože $p\text{-hodnota} > \alpha$, akceptujeme H0. Předpokládá se, že data jsou normálně distribuována.*).

4. Effect size. The observed effect size **KS - D** is very small, 0.14. This indicates that the magnitude of the difference between the sample distribution and the normal distributions is **very small**.
(Pozorovaná **velikost účinku** KS – D je **velmi malá, 0,14**. To ukazuje, že velikost rozdílu mezi výběrovým rozdělením a normálním rozdělením je **velmi malá**).

The **Kolmogorov-Smirnov effect size** to measure the deviation from the normality: We know only the sample effect size!

We defined the following **effect levels** of the effect size:

Velikost efektu Kolmogorov-Smirnov pro měření odchylky od normálu: Známe pouze velikost efektu vzorku!

Definovali jsme následující úrovně efektu velikosti efektu:

Small - 0.0063, the effect size of Chi-squared(df=40)

Medium - 0.00224, the effect size of Chi-squared(df=20)

Large - 0.0427, the effect size of Chi-squared(df=5)

Obr. 1 a 2: DH – Histogram; Q-Q Plot ScatterChart (ukázka)

Histogram (při počtu $n = 15$ málo užitečný, lepší je **bodový graf** (Q-Q Plot ScatterChart)).

Volba pro NDH: S-W test; Significance level (α): 0,05; Outliers: include; Rounding: 2
Výpočet (Calculate):

Reporting Shapiro-Wilk test in APA Format

The S-W test did not show a significant departure from normality, $W(15) = .93$, $p = .268$
S-W test *neprokázal významný odklon od normality, $W(15) = 0,93$, $p = 0,268$*

1. H0 hypothesis. Since $p\text{-value} > \alpha$, we accept the H0. It is assumed that the data is normally distributed. (*Protože $p\text{-hodnota} > \alpha$, **akceptujeme H0**. Předpokládá se, že data jsou normálně distribuována*).

4. Effect size. The observed effect size KS - D is medium, **0.17**. This indicates that the magnitude of the difference between the sample distribution and the normal distributions is medium. (*Pozorovaná velikost účinku KS - D je střední, **0,17**. To ukazuje, že velikost rozdílu mezi distribucí vzorku a normální distribucí je střední.*)

Since the null assumption cannot be rejected, you may ignore the effect size.

(Protože nulový předpoklad nelze zamítnout, můžete velikost efektu ignorovat).

Obr. 3 a 4: DH – Histogram; Q-Q Plot ScatterChart (ukázka)

Histogram (při počtu $n = 15$ málo užitečný, lepší je **bodový graf** (Q-Q Plot ScatterChart)).

2 Výpočet základních statistických charakteristik (M, Me, Mo, SD, V)

Mean, Median, Modus, směrodatná (standardní) odchylka (SD)

<https://www.statskingdom.com/mean-median-mode-calculator.html>

Reporting in APA style

DH: M = 25.6, SD = 4.95, Mdn = 24.9, IQR = 22.8–27.85 (IQR = 5.05)

NDH: M = 22.35, SD = 4.93, Mdn = 21.6, IQR = 17.6–25.8 (IQR = 8.2)

Variční koeficient

<https://www.statskingdom.com/coefficient-of-variation-calculator.html>

Sample CV (Coefficient of Variation) = $(s/M) \cdot 100$

DH: CV = 0.1935 (19.35%)

NDH: CV = 0.2206 (22.06%)

3 TESTOVÁNÍ STATISTICKÝCH HYPOTÉZ

3.1 Statistická významnost

3.1.1 t – test (parametrický test, *bylo-li prokázáno normální rozložení četnosti*)

1 Two Sample T-Test Calculator (Pooled-Variance)

What is the two-sample equal variances (pooled-variance) t-test? Co je dvouvýběrový t-test se stejnými rozptyly? Dvouvýběrový t-test *se stejnými rozptyly* nazývaný také nezávislý t-test porovnává populační průměry (průměry) dvou nezávislých skupin. Test předpokládá, že *směrodatné odchylky jsou pro obě skupiny totožné*.

2 Two Sample T-Test Calculator (Welch's T-test)

What is Welch's t-test? Co je Welchův t-test?

Dvouvýběrový t-test *s nestejnými rozptyly* nazývaný také Welchův t-test porovnává populační průměry (průměry) dvou nezávislých skupin. Test předpokládá, že *směrodatné odchylky nejsou pro obě skupiny totožné*.

Volba: <https://www.statskingdom.com/150MeanT2uneq.html>

DH: M = 25.6; SD = 4.95; NDH: M = 22.35; SD = 4.93

1. H0 hypothesis.

Since $p\text{-value} > \alpha$, H0 cannot be rejected.

The average of Group-1's population is assumed to be equal to the average of Group-2's population.

(Protože p -hodnota $> \alpha$, **nelze H_0 zamítnout**. Předpokládá se, že průměr populace skupiny-1 se rovná průměru populace skupiny-2.).

2. P-value.

The p-value equals **0.08237**. The larger the p-value the more it supports H_0 .
(P -hodnota. P -hodnota se rovná 0,08237. Čím větší je p -hodnota, tím více podporuje H_0).

4. Effect size

The observed effect size d is large, **0.66**. This indicates that the magnitude of the difference between the average and average is large.
(Pozorovaná velikost účinku d je velká, **0,66**. To naznačuje, že velikost rozdílu mezi průměrem a průměrem je velká.).

Since the null assumption cannot be rejected, you may ignore the effect size.
(Protože nulový předpoklad nelze zamítnout, můžete velikost efektu ignorovat).

3.1.2 Mann-Whitney U Test Calculator

(neparametrický test, neprokázáno normální rozložení četností)

https://www.statskingdom.com/170median_mann_whitney.html

DH: $M = 25.6$; $SD = 4.95$; NDH: $M = 22.35$; $SD = 4.93$

1. H_0 hypothesis

Since p -value $> \alpha$, H_0 cannot be rejected. The randomly selected value of Group1's population is assumed to be equal to the randomly selected value of Group2's population.
(Protože p -hodnota $> \alpha$, **nelze H_0 zamítnout**. Předpokládá se, že náhodně vybraná hodnota populace skupiny 1 se rovná náhodně vybrané hodnotě populace skupiny 2).

2. P-value

The p-value equals **0.08856**. The larger the p-value the more it supports H_0 .
(P -hodnota se rovná **0,08856**. Čím větší je p -hodnota, tím více podporuje H_0).

4. Effect size

The observed common language effect size, $U1/(n1n2)$, is **0.68**, this is the probability that a random value from Group1 is greater than a random value from Group2.
(Pozorovaná velikost efektu, $U1/(n1n2)$, je **0,68**, což je pravděpodobnost, že náhodná hodnota ze skupiny 1 je větší než náhodná hodnota ze skupiny 2.)

3.2 Věcná významnost (effect size, ES)

<https://www.statskingdom.com/effect-size-calculator.html>

DH: $M = 25.6$; $SD = 4.95$; NDH: $M = 22.35$; $SD = 4.93$

Reporting results in APA Format

The effect size for this analysis ($d = 0.66$) was found to exceed Cohen's (1988) convention for a medium effect ($d = 0.5$).

(Bylo zjištěno, že velikost účinku pro tuto analýzu ($d = 0,66$) překračuje Cohenovu (1988) konvenci pro střední účinek ($d = 0,5$).

Test	small	medium	large
d	0,2-0,49	0,5-0,79	0,80 a více
r	0,1-0,29	0,3-0,49	0,50 a více

4 Korelační počet

4.1 Pearsonův korelační koeficient (pro metrická data s normálním rozložením četností)

<https://www.statskingdom.com/correlation-calculator.html>

Vložíme data DH a NDH

Results (výsledek): Pearson correlation coefficient (r) = **0.8931** (0.893)

Reporting correlation in APA Format

Results of the pearson correlation indicated that there is a significant large positive relationship between X and Y, ($r(13) = .893$, $p < .001$).

(Výsledky Pearsonovy korelace ukázaly, že mezi X a Y existuje významně velký pozitivní vztah ($r(13) = 0,893$, $p < 0,001$).

1. H0 hypothesis

Since the p -value $< \alpha$, **H0 is rejected**.

The population's correlation is considered to be not equal to the expected correlation (0).

In other words, the difference between the sample correlation and the expected correlation is big enough to be statistically significant.

*(Protože p -hodnota $< \alpha$, **H0 je zamítnuta**. Má se za to, že korelace populace není rovna očekávané korelaci (0). Jinými slovy, rozdíl mezi výběrovou korelací a očekávanou korelací je dostatečně velký na to, aby byl statisticky významný).*

Obr. 5: Grafické znázornění (korelogram)

2. Correlation effect size

The correlation value is also the correlation effect size. Define the level of the effect size is only a rule of thumb. Following Cohen's guidelines ([Cohen 1988 - pg 413](#))

(Hodnota korelace je zároveň velikostí korelačního účinku. Definovat úroveň velikosti efektu je pouze orientační pravidlo (Cohen 1988 - str. 413).

Correlation value(r)	Level
$ r < 0.1$	Very small
$0.1 \leq r < 0.3$	Small
$0.3 \leq r < 0.5$	Medium
$0.5 \leq r $	Large

4.2 Spearmanův korelační koeficient (pro ordinální data, resp. při neprokázaném normálním rozložením četností)

<https://www.statskingdom.com/correlation-calculator.html>

Results (výsledek): Spearman's rank correlation coefficient (r_s) = **0.8977** (0.897)

Reporting correlation in APA Format

Results of the spearman correlation indicated that there is a significant large positive relationship between X and Y, ($r(13) = .898$, $p < .001$).

(Výsledky Spearmanovy pořadové korelace ukázaly, že existuje významný velký pozitivní vztah mezi X a Y, ($r(13) = 0,898$, $p < 0,001$).

1. H0 hypothesis

Since the p -value $< \alpha$, H0 is rejected.

The population's correlation is considered to be not equal to the expected correlation (0).

2. Correlation effect size: $r = 0,898$ (ES large)

5 Výpočty lze provádět také v software Excel (viz příklady v přednáškách, resp. Statistica, SPSS atd.).