

Statistická analýza dat – vzorové zpracování

Statistics Calculators k volnému použití

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

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

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

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

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

Tabulka 1: Vzorová data (n = 15, je vyžadována **tečka** místo desetinné čárky)
Síla stisku dominantní (DH) a nedominantní ruky (NDH)

Č.	DH	NDH
1	26.8	21.1
2	43.3	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	26.23	22.35
SD	6.50	4.93

Postup (statistická analýza dat)

1. Ověření normality dat
2. Výpočet základních statistických charakteristik (M, SD, V)
3. Testování statistických hypotéz (t – test, U test)
4. Korelační počet (Pearson, Spearman)

Obdobné výpočty v Excel

Statistics Calculators např.

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

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

Výpočty vzorových dat (soubory HGS dom/nondom, Tabulka 1)

1. Ověření normality (např. K-S, Lilliefors, Shapiro-Wilk test)

Kolmogorovův-Smirnovův test (K-S) (pro malé soubory $n \leq 50$)

K-S test je metoda umožňující testovat, zda jedna jednorozměrná proměnná má předpokládané (nejčastěji normální) rozdělení četností.

Výpočet: <https://www.socscistatistics.com/tests/kolmogorov/default.aspx>

DH: The p-value is .52533. Your data does **not differ significantly from** that which is **normally distributed**.

*Hodnota statistiky testu K-S (D) je 0,19948. Hodnota p je 0,52533. Vaše data se **významně neliší od normální distribuce** (mají normálního rozložení četností).*

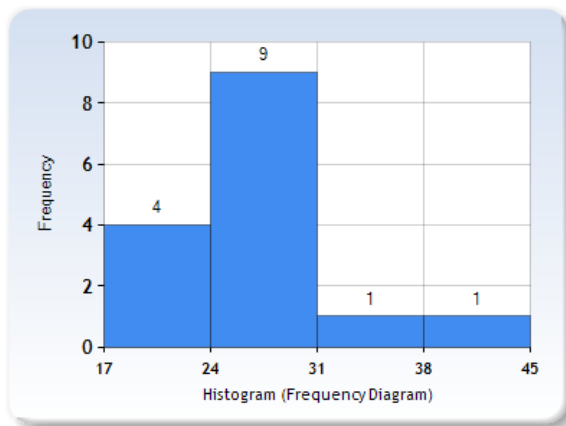
NDH: The p-value is .69091. Your data does **not differ significantly from** that which is normally distributed.

*Hodnota p je 0,69091. Vaše data se **významně neliší** od těch, která jsou běžně distribuována (mají normálního rozložení četností).*

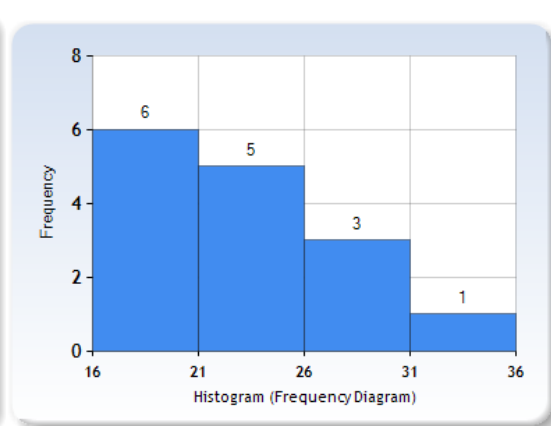
Histogram

<https://www.socscistatistics.com/descriptive/histograms/default.aspx>

DH



NDH



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

Mean, Median, Modus

<https://www.socscistatistics.com/descriptive/averages/default.aspx>

DH: Mean: 26.3, Median: 24.9, Mode: 26.8

NDH: Mean: 22.4, Median: 21.6, Mode: 17.6

Směrodatná (standardní) odchylka (SD)

<https://www.socscistatistics.com/descriptive/std/default.aspx>

DH: SD = 6.50; **NDH:** SD = 4.93

Variační koeficient

<https://www.socscistatistics.com/descriptive/coefficientvariation/default.aspx>

DH: Coefficient of Variation (CV) = $(s/M) * 100 = (6.5/26.27) * 100 = 24.74\%$

NDH: Coefficient of Variation = 22.06%

3 TESTOVÁNÍ STATISTICKÝCH HYPOTÉZ

3.1 Statistická významnost

T – test (parametrický test, prokázáno normální rozložení četností)

T-Test Calculator for 2 Independent Means (další volba for 2 Dependent Means)

<https://www.socscistatistics.com/tests/studentttest/default2.aspx>

One-tailed hypothesis: The t-value is 1.85803. The p-value is .036854. The result is significant at $p < .05$. *Nulová hypotéza se zamítá = měl by následovat výpočet Effect size.*

Two-tailed hypothesis: The t-value is 1.85803. The p-value is .073709. The result is not significant at $p < .05$. Výsledek není významný na $p < 0,05$. Nulová hypotéza se nezamítá.

Mann-Whitney U Test Calculator (neparametrický test, neprokázáno normální rozložení četností)

<https://www.socscistatistics.com/tests/mannwhitney/default2.aspx>

DH: Mean = 26.27; SD = 6.50; NDH: Mean = 22.35; SD = 4.93

The U-value is 71. The critical value of U at $p < .05$ is 64. Therefore, the result is not significant at $p < .05$. Hodnota U je 71. Kritická hodnota U při $p < 0,05$ je 64. *Výsledek tedy není významný při $p < 0,05$.*

3.2 Věcná významnost (effect size, ES)

<https://www.socscistatistics.com/effectsize/default3.aspx>

Výpočet **Cohen's d** (0.2; 0.5; 0.8) = $(22.35 - 26.27) / 5.768661 = 0.68$ (medium)

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.socscistatistics.com/tests/pearson/default.aspx>

Výsledek: $r = 0.91$

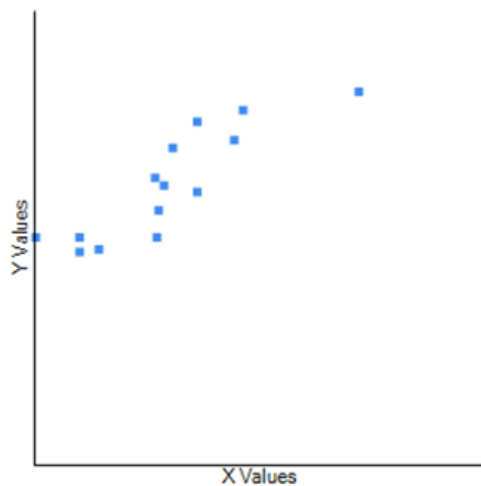
This is a strong positive correlation, which means that high X variable scores go with high Y variable scores (and vice versa).

Jedná se o silnou pozitivní korelaci, což znamená, že vysoké skóre proměnných X odpovídá vysokému skóre proměnných Y (a naopak).

The P-Value is $< .00001$. The result is significant at $p < .05$.

Hodnota P je $< 0,00001$. Výsledek je významný na $p < 0,05$.

Obrázek 1: Grafické znázornění (korelogram)



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

<https://www.socscistatistics.com/tests/spearman/default.aspx>

Výsledek: $r_s = 0.90$

By normal standards, the association between the two variables would be considered statistically significant. *Podle normálních standardů lze vztah mezi těmito dvěma proměnnými považovat za statisticky významný.*

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

Použití jiného statistics calculator

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

1. Ověření normality dat

Kolmogorov-Smirnov a Lilliefors test

(pro malé soubory $n \leq 50$)

Reporting results in APA Format

Hlášení výsledků ve formátu APA

Výsledky **Lillieforsova** testu ukázaly, že **existuje nevýznamný rozdíl od normálního rozdělení** ($D(15) = 0,2$, $p = 0,107$).

Výsledky **Kolmogorovova** testu ukázaly, že **existuje nevýznamný rozdíl od normálního rozdělení**, ($D(15) = 0,2$, $p = 0,518$).

1. H0 hypothesis

Since $p\text{-value} > \alpha$, we accept the H0.

It is assumed that the data distribution is normal.

Protože $p\text{-hodnota} > \alpha$, **akceptujeme H0**.

Předpokládá se, že **distribuce dat je normální**.

Shapiro-Wilkův test (pro velké soubory $n > 50$)

Shapiro-Wilkův test s použitím tabulek jako distribuce (pravostranný)

Protože $n \leq 50$ jsme k výpočtu $p\text{-hodnoty}$ použili Shapiro-Wilkovy tabulky.

Graf normálního rozdělení je pouze pro vizualizaci.

1. Hypotéza H0

Protože $p\text{-hodnota} < \alpha$, odmítáme H0.

Předpokládá se, že **data nejsou normálně** distribuována.

Jinými slovy, rozdíl mezi vzorkem dat a normálním rozdělením je dostatečně velký na to, aby byl statisticky významný.