

## LEKCE03

# NORMÁLNÍ ROZLOŽENÍ A ZÁKLADY TESTOVÁNÍ HYPOTÉZ. STATISTICKÁ INFERENCE ANEB ZOBECŇOVÁNÍ VÝSLEDKŮ Z VÝBĚROVÉHO NA ZÁKLADNÍ SOUBOR

### vzorový výsledek cvičení

**CVIČENÍ 3.1:**

V souboru EVS\_CR1999.sav zjistěte, zdali rozložení názoru na to, kdo by měl být odpovědný za zajištění bydlení (proměnná q54h), je normální. Řešte graficky i početně.

**PROCEDURA FREQUENCIES**

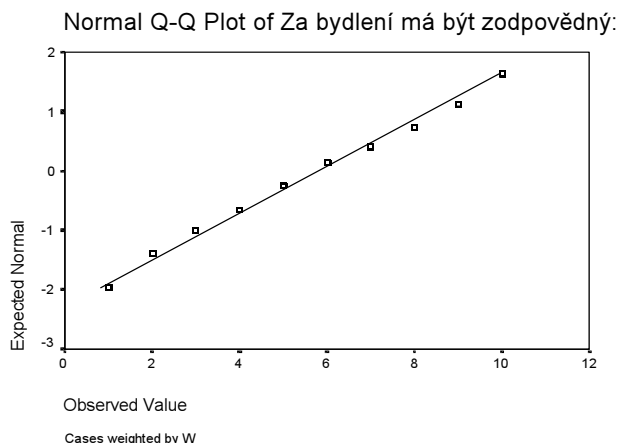
Q54H Za bydlení má být zodpovědný:

|         |                    | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|--------------------|-----------|---------|---------------|--------------------|
| Valid   | 1 každý člověk sám | 94        | 4.9     | 5.0           | 5.0                |
|         | 2                  | 119       | 6.2     | 6.3           | 11.3               |
|         | 3                  | 166       | 8.7     | 8.8           | 20.1               |
|         | 4                  | 199       | 10.4    | 10.6          | 30.7               |
|         | 5                  | 372       | 19.5    | 19.7          | 50.4               |
|         | 6                  | 209       | 11.0    | 11.1          | 61.5               |
|         | 7                  | 175       | 9.2     | 9.3           | 70.8               |
|         | 8                  | 240       | 12.6    | 12.7          | 83.5               |
|         | 9                  | 123       | 6.4     | 6.5           | 90.0               |
|         | 10 stát            | 189       | 9.9     | 10.0          | 100.0              |
|         | Total              | 1886      | 98.9    | 100.0         |                    |
| Missing | -2                 | 9         | .5      |               |                    |
|         | -1                 | 13        | .7      |               |                    |
|         | Total              | 22        | 1.1     |               |                    |
|         | Total              | 1908      | 100.0   |               |                    |

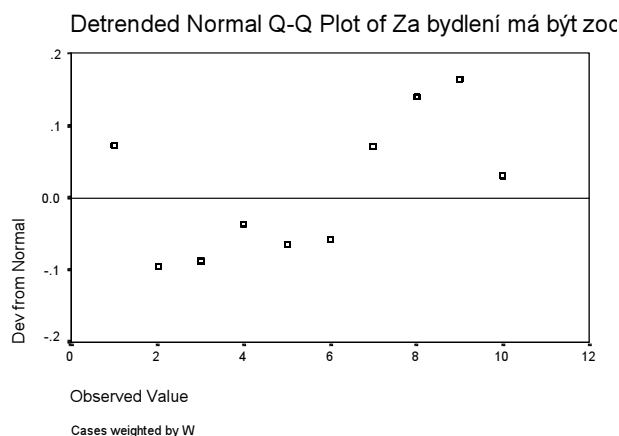


V proceduře FREQUENCIES jsme si vypočítali rozložení proměnné Q54h. I když se zdá být normálnímu velmi blízké, přece jen je nutné si to ověřit (zejména prvá strana rozložení modelu normálního rozložení příliš neodpovídá).

## Procedura EXPLORE:



Na grafu (normal Q-Q plot) lze vidět, že se pozorované hodnoty od očekávaných (pokud by bylo rozložení normální) odchyľují, o kolik se to však opticky jeví, ovlivňuje měřítko.



V tomto grafu odchylek je již zřejmější, že nepůjde o normální rozložení.

## Tests of Normality

|                                    | Kolmogorov-Smirnov <sup>a</sup> |      |      |
|------------------------------------|---------------------------------|------|------|
|                                    | Statistic                       | df   | Sig. |
| Q54H Za bydlení má být zodpovědný: | .123                            | 1886 | .000 |

a. Lilliefors Significance Correction

Test normality (Kolmogorov-Smirnov) nás pak definitivně přesvědčuje, že o normální rozložení nejde. Testujeme zde totiž nulovou hypotézu:

**naše empirické rozložení a (modelové) normální rozložení jsou shodná,**

Test nám říká, že při zamítnutí této hypotézy se dopouštíme jen nepatrného rizika chyby (menší jak 0,1%, zatímco by nám stačilo 5%, neboli hladina sig. 0,05).

Tento test je důležitý, vzpomeňte si totiž, že histogram nám sice naznačil odlišnost našeho normálního rozložení od normálního, ale neřekl nám nic o tom, zda je tato odlišnost dostatečně velká, aby byla statisticky významná.

## CVIČENÍ 3.2:

Stanovte interval spolehlivosti průměru odpovědí na otázku q54f z výzkumu EVS-ČR1999, zdali by u nás mělo *vzrůstat soukromé vlastnictví obchodu a průmyslu, nebo zda by se mělo v co nejširší míře zachovat vlastnictví státní*. Interval spolehlivosti stanovte se spolehlivostí 95 a 99 %. V odpovědi nad získaným výsledkem srovnajte českou situaci se situací v některých evropských zemích. K tomu použijte tabulku 1. Do ní dopočítejte intervaly spolehlivosti pro vámi zvolených pět zemí.

Q54F Rozšiřovat/ zachovat by se mělo obchod. a prům.:

|         |                        | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|------------------------|-----------|---------|---------------|--------------------|
| Valid   | 1 soukromé vlastnictví | 199       | 10.4    | 10.7          | 10.7               |
|         | 2                      | 211       | 11.0    | 11.4          | 22.1               |
|         | 3                      | 262       | 13.7    | 14.2          | 36.3               |
|         | 4                      | 203       | 10.6    | 10.9          | 47.2               |
|         | 5                      | 363       | 19.0    | 19.6          | 66.8               |
|         | 6                      | 179       | 9.4     | 9.7           | 76.5               |
|         | 7                      | 128       | 6.7     | 6.9           | 83.4               |
|         | 8                      | 134       | 7.0     | 7.2           | 90.6               |
|         | 9                      | 80        | 4.2     | 4.3           | 94.9               |
|         | 10 státní vlastnictví  | 94        | 4.9     | 5.1           | 100.0              |
|         | Total                  | 1851      | 97.0    | 100.0         |                    |
| Missing | -2                     | 12        | .6      |               |                    |
|         | -1                     | 46        | 2.4     |               |                    |
|         | Total                  | 57        | 3.0     |               |                    |
|         | Total                  | 1908      | 100.0   |               |                    |

Považujeme škálu za kardinální. Pro výpočet průměru lze požit kteroukoliv z procedur DESCRIPTIVE STATISTICS (FREQUENCIES, DESCRIPTIVES či EXPLORE). Pro výpočet intervalu spolehlivosti jsou však první dvě nepříliš výhodné, počítají jen standardní chybu a interval bychom museli dopočítávat podle známých vlastností normálního rozložení (pro 95% je to  $\pm 2S.E.$  a pro 99% je to  $\pm 3S.E.$ ).

a) Příklad výstupu z procedury FREQUENCIES:

## Statistics

Q54F Rozšiřovat/ zachovat by se mělo obchod. a prům.:

|                    |         |      |
|--------------------|---------|------|
| N                  | Valid   | 1851 |
|                    | Missing | 57   |
| Mean               |         | 4.71 |
| Std. Error of Mean |         | .06  |
| Median             |         | 5.00 |
| Mode               |         | 5    |
| Std. Deviation     |         | 2.51 |

b) Příklad výstupu z procedury DESCRIPTIVES:

## Descriptive Statistics

|   | N         | Mean      |            | Std.      |
|---|-----------|-----------|------------|-----------|
|   | Statistic | Statistic | Std. Error | Statistic |
| Q54F Rozšiřovat/<br>zachovat by se mělo<br>obchod. a prům.: | 1851      | 4.71      | .06        | 2.51      |
| Valid N (listwise)  | 1851      |           |            |           |

Pozor: V obou případech jsou hodnoty S.E. zokrouhleny.

c) Příklad výpočtu z procedury EXPLORE:

- Pro 95%

|   |                                  | Statistic   | Std. Error |  |
|---|----------------------------------|-------------|------------|--|
| Q54F Rozšiřovat/<br>zachovat by se mělo<br>obchod. a prům.: | Mean                             | 4.71        | .06        |  |
|   | 95% Confidence Interval for Mean | Lower Bound | 4.60       |  |
|   |                                  | Upper Bound | 4.83       |  |
|   | 5% Trimmed Mean                  | 4.62        |            |  |
|   | Median                           | 5.00        |            |  |
|   | Variance                         | 6.317       |            |  |
|   | Std. Deviation                   | 2.51        |            |  |
|   | Minimum                          | 1           |            |  |
|   | Maximum                          | 10          |            |  |
|   | Range                            | 9           |            |  |
|   | Interquartile Range              | 3.00        |            |  |
|   | Skewness                         | .376        | .057       |  |
|   | Kurtosis                         | -.678       | .114       |  |

všimněte si, že  
4,83 je

$4,71 + 2 \cdot 0,06$  neboli  
 $2 \times S.E$  jak jsme ji  
vypočítali v  
DESCRIPTIVES  
a jak je uvedena i  
v této tabulce

a) Pro 99%

|   |                                  | Statistic   | Std. Error |  |
|---|----------------------------------|-------------|------------|--|
| Q54F Rozšiřovat/<br>zachovat by se mělo<br>obchod. a prům.: | Mean                             | 4.71        | .06        |  |
|   | 99% Confidence Interval for Mean | Lower Bound | 4.56       |  |
|   |                                  | Upper Bound | 4.86       |  |
|   | 5% Trimmed Mean                  | 4.62        |            |  |
|   | Median                           | 5.00        |            |  |
|   | Variance                         | 6.317       |            |  |
|   | Std. Deviation                   | 2.51        |            |  |
|   | Minimum                          | 1           |            |  |
|   | Maximum                          | 10          |            |  |
|   | Range                            | 9           |            |  |
|   | Interquartile Range              | 3.00        |            |  |
|   | Skewness                         | .376        | .057       |  |
|   | Kurtosis                         | -.678       | .114       |  |

| Země | Průměr | Std. Err. | N    | 95% Interval spolehlivosti | 99% Interval spolehlivosti |
|------|--------|-----------|------|----------------------------|----------------------------|
| ČR   | 4,71   | 2,51      | 1851 | 4,60 až 4,83               | 4,56 až 4,86               |

Pozor: Standardní odchylka nám k řešení našeho úkolu nepomůže, uvádíme ji jen proto, aby byl údaj o charakteristice rozložení postoje kompletní.

Všimněte si, že jsme za větší spolehlivost zaplatili menší přesnosti (na 99% víme, že průměr leží v intervalu 4,56 až 4,86, ale tento interval je rozsáhlejší než když nám stačí 95% jistota).

Tab. 1: Odpovědi na otázku q54f v různých evropských zemích

| Země    | Průměr | Std. Dev. | N    | 95% Interval spolehlivosti |
|---------|--------|-----------|------|----------------------------|
| Francie | 4,03   | 2,23      | 1498 |                            |
| SRN     | 4,05   | 2,14      | 1836 |                            |
| etc.    |        |           |      |                            |
| ČR      | 4,71   | 2,51      | 1851 | 4,60 až 4,83               |

Tento příklad nemá řešení, protože neuvádí S.E. ale St. Dev. Museli bychom si je v souborech jednotlivých zemí spočítat, jako jsme to učinili pro Českou republiku.

## CVIČENÍ 3. 3:

Stanovte intervaly spolehlivosti pro průměrnou volbu mužů a pro průměrnou volbu žen na škále proměnné q54f. Překrývají se tyto intervaly nebo ne (co z výsledku plyne)?

a) Provedeme výběr (SELECT IF) pro muže a spočítáme požadované hodnoty (procedura EXPLORE).

| Descriptives  |                                     |             | Statistic | Std. Error |
|---|-------------------------------------|-------------|-----------|------------|
| Q54F Rozšiřovat/<br>zachovat by se mělo<br>obchod. a prům.: | Mean                                |             | 4.49      | 8.40E-02   |
|   | 95% Confidence<br>Interval for Mean | Lower Bound | 4.32      |            |
|   |                                     | Upper Bound | 4.65      |            |
|   |                                     |             |           |            |
|   | 5% Trimmed Mean                     |             | 4.38      |            |
|   | Median                              |             | 4.00      |            |
|   | Variance                            |             | 6.397     |            |
|   | Std. Deviation                      |             | 2.53      |            |
|   | Minimum                             |             | 1         |            |
|   | Maximum                             |             | 10        |            |
|   | Range                               |             | 9         |            |
|   | Interquartile Range                 |             | 4.00      |            |
|   | Skewness                            |             | .481      | .081       |
|   | Kurtosis                            |             | -.622     | .162       |

Lze to spočítat i pro 99% interval spolehlivosti:

| Descriptives  |                                     |             | Statistic | Std. Error |
|---|-------------------------------------|-------------|-----------|------------|
| Q54F Rozšiřovat/<br>zachovat by se mělo<br>obchod. a prům.: | Mean                                |             | 4.49      | 8.40E-02   |
|   | 99% Confidence<br>Interval for Mean | Lower Bound | 4.27      |            |
|   |                                     | Upper Bound | 4.70      |            |
|   |                                     |             |           |            |
|   | 5% Trimmed Mean                     |             | 4.38      |            |
|   | Median                              |             | 4.00      |            |
|   | Variance                            |             | 6.397     |            |
|   | Std. Deviation                      |             | 2.53      |            |
|   | Minimum                             |             | 1         |            |
|   | Maximum                             |             | 10        |            |
|   | Range                               |             | 9         |            |
|   | Interquartile Range                 |             | 4.00      |            |
|   | Skewness                            |             | .481      | .081       |
|   | Kurtosis                            |             | -.622     | .162       |

Totéž provedeme pro ženy:

| Descriptives  |                                     |             | Statistic | Std. Error |
|---|-------------------------------------|-------------|-----------|------------|
| Q54F Rozšiřovat/<br>zachovat by se mělo<br>obchod. a prům.: | Mean                                |             | 4.93      | 8.07E-02   |
|   | 95% Confidence<br>Interval for Mean | Lower Bound | 4.77      |            |
|   |                                     | Upper Bound | 5.09      |            |
|   |                                     |             |           |            |
|   | 5% Trimmed Mean                     |             | 4.87      |            |
|   | Median                              |             | 5.00      |            |
|   | Variance                            |             | 6.152     |            |
|   | Std. Deviation                      |             | 2.48      |            |
|   | Minimum                             |             | 1         |            |
|   | Maximum                             |             | 10        |            |
|   | Range                               |             | 9         |            |
|   | Interquartile Range                 |             | 4.00      |            |
|   | Skewness                            |             | .288      | .080       |
|   | Kurtosis                            |             | -.679     | .159       |

Lze to spočítat i pro 99% interval spolehlivosti:

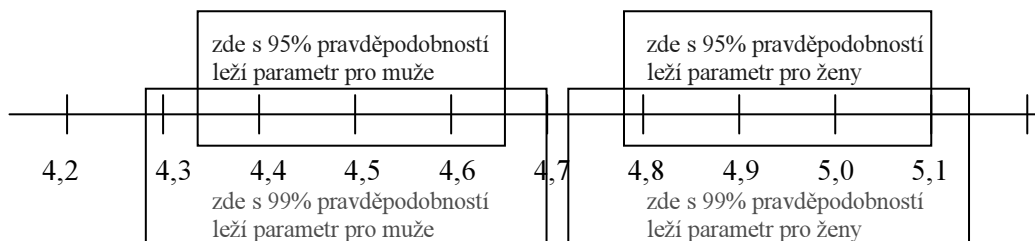
Descriptives

|   |                                  |             | Statistic | Std. Error |
|---|----------------------------------|-------------|-----------|------------|
| Q54F Rozšiřovat/<br>zachovat by se mělo<br>obchod. a prům.: | Mean                             |             | 4.93      | 8.07E-02   |
|   | 99% Confidence Interval for Mean | Lower Bound | 4.72      |            |
|   |                                  | Upper Bound | 5.14      |            |
|   | 5% Trimmed Mean                  |             | 4.87      |            |
|   | Median                           |             | 5.00      |            |
|   | Variance                         |             | 6.152     |            |
|   | Std. Deviation                   |             | 2.48      |            |
|   | Minimum                          |             | 1         |            |
|   | Maximum                          |             | 10        |            |
|   | Range                            |             | 9         |            |
|   | Interquartile Range              |             | 4.00      |            |
|   | Skewness                         |             | .288      | .080       |
|   | Kurtosis                         |             | -.679     | .159       |

Intervaly spolehlivosti průměru voleb na škále Q54f pro muže a ženy

|      | 95% interval spolehlivosti |                | 99% interval spolehlivosti |                |
|------|----------------------------|----------------|----------------------------|----------------|
|      | horní hranice              | spodní hranice | horní hranice              | spodní hranice |
| muži | 4,65                       | 4,32           | 4,70                       | 4,27           |
| ženy | 4,77                       | 5,09           | 5,14                       | 4,72           |

- b) Porovnáme intervaly. Všimněte si, že nepronikají ani 99% intervaly spolehlivosti (rozsáhlejší než 95%). Lze předpokládat na 99%, že se neliší jen statistiky (průměrné volby v našem souboru), ale i parametry (průměrné volby v základním souboru, z něhož byl náš výběr proveden).



### CVIČENÍ 3.4:

Určete věkové hranice oddělujících 2,28% nejstarších a 2,28% nejmladších osob v našem souboru. O co se jedná? Znamená to určit v rozložení věku hranice vymezuující 95% souboru (směrem nahoru i směrem dolů odřízneme dohromady 5%): průměrný věk  $\pm 2$  standardní odchylky. Jak i výpočet ukazuje, nebyl to šťastný nápad pokusit se simulovat výpočet na rozložení, které není normální (ve statistickém smyslu).

Statistics

| VEK                |         |       |
|--------------------|---------|-------|
| N                  | Valid   | 1901  |
|                    | Missing | 7     |
| Mean               |         | 45.69 |
| Std. Error of Mean |         | .39   |
| Std. Deviation     |         | 16.83 |

Tyto hranice představují:

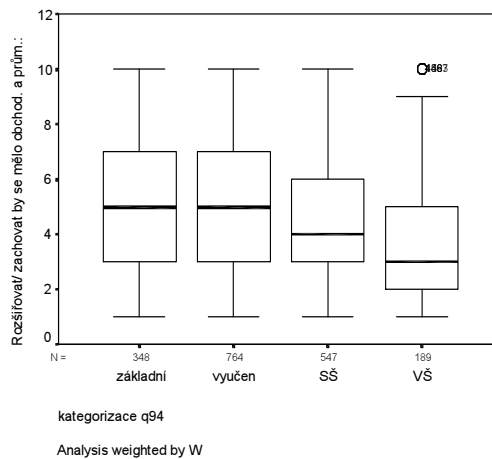
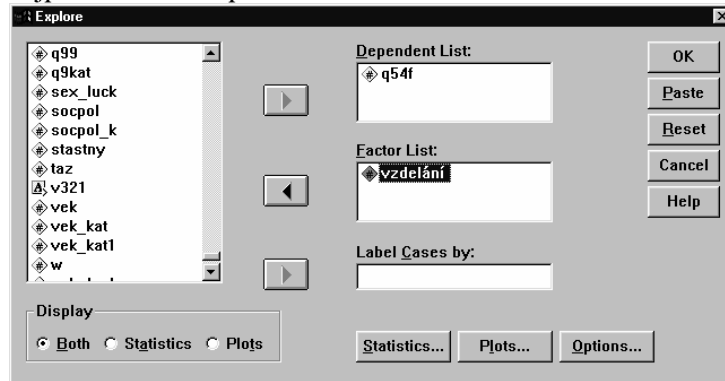
$$45,66 - 2 * 16,83 = 12,03$$

$$45,66 + 2 * 16,83 = 79,35$$

**CVIČENÍ 3.5:**

Co se týče umístění na škále proměnné q54f, vidíme rozdíl v průměrných volbách osob se základním vzděláním a osob s vysokoškolským vzděláním (proměnná vzdelani). Vyjádřete ovšem jejich pozici ve vztahu k průměru celého souboru pomocí standardní odchylky a obě hodnoty porovnejte (z-skóre a jejich průměrné hodnoty v obou skupinách).

Rozdíly si můžeme nejprve zobrazit v proceduře EXPLORE.

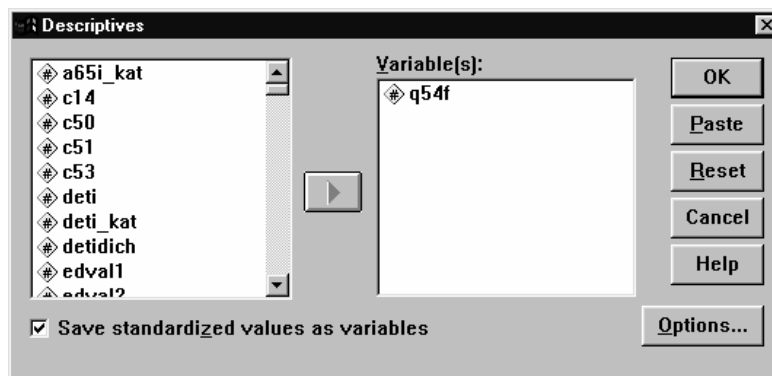


## Descriptives

| VZDELÁNÍ<br>kategorizac  |             |                                     |             | Statistic | Std.<br>Error |
|--|-------------|-------------------------------------|-------------|-----------|---------------|
| Q54F<br>Rozšiřovat/<br>zachovat by<br>se mělo<br>obchod. a<br>prům.: | 1 základní  | Mean                                |             | 5.34      | .13           |
|  |             | 95% Confidence<br>Interval for Mean | Lower Bound | 5.08      |               |
|  |             |                                     | Upper Bound | 5.60      |               |
|  |             | 5% Trimmed Mean                     |             | 5.32      |               |
|  |             | Median                              |             | 5.00      |               |
|  |             | Variance                            |             | 6.039     |               |
|  |             | Std. Deviation                      |             | 2.46      |               |
|  |             | Minimum                             |             | 1         |               |
|  |             | Maximum                             |             | 10        |               |
|  |             | Range                               |             | 9         |               |
|  |             | Interquartile Range                 |             | 4.00      |               |
|  |             | Skewness                            |             | .173      | .131          |
|  |             | Kurtosis                            |             | -.781     | .261          |
|  |             | 2 vyučen                            |             | Mean      |               |
| 95% Confidence<br>Interval for Mean                                  | Lower Bound |                                     |             | 4.69      |               |
|  | Upper Bound |                                     |             | 5.07      |               |
| 5% Trimmed Mean  |             |                                     |             | 4.81      |               |
| Median   |             |                                     |             | 5.00      |               |
| Variance   |             |                                     |             | 6.959     |               |
| Std. Deviation   |             |                                     |             | 2.64      |               |
| Minimum  |             |                                     |             | 1         |               |
| Maximum  |             |                                     |             | 10        |               |
| Range  |             |                                     |             | 9         |               |
| Interquartile Range  |             |                                     |             | 4.00      |               |
| Skewness   |             |                                     |             | .292      | .088          |
| Kurtosis   |             |                                     |             | -.812     | .177          |
| 3 SŠ   |             |                                     |             | Mean      |               |
|  |             | 95% Confidence<br>Interval for Mean | Lower Bound | 4.18      |               |
|  |             |                                     | Upper Bound | 4.58      |               |
|  |             | 5% Trimmed Mean                     |             | 4.29      |               |
|  |             | Median                              |             | 4.00      |               |
|  |             | Variance                            |             | 5.628     |               |
|  |             | Std. Deviation                      |             | 2.37      |               |
|  |             | Minimum                             |             | 1         |               |
|  |             | Maximum                             |             | 10        |               |
|  |             | Range                               |             | 9         |               |
|  |             | Interquartile Range                 |             | 3.00      |               |
|  |             | Skewness                            |             | .467      | .104          |
|  |             | Kurtosis                            |             | -.507     | .209          |
|  |             | 4 VŠ                                |             | Mean      |               |
| 95% Confidence<br>Interval for Mean                                  | Lower Bound |                                     |             | 3.53      |               |
|  | Upper Bound |                                     |             | 4.14      |               |
| 5% Trimmed Mean  |             |                                     |             | 3.71      |               |
| Median   |             |                                     |             | 3.00      |               |
| Variance   |             |                                     |             | 4.467     |               |
| Std. Deviation   |             |                                     |             | 2.11      |               |
| Minimum  |             |                                     |             | 1         |               |
| Maximum  |             |                                     |             | 10        |               |
| Range  |             |                                     |             | 9         |               |
| Interquartile Range  |             |                                     |             | 3.00      |               |
| Skewness   |             |                                     |             | .766      | .177          |
| Kurtosis   |             |                                     |             | .257      | .352          |



V proceduře DESCRIPTIVES zaškrtneme požadavek na výpočet Z-skóre.



Následně spočítáme průměrné hodnoty z-skóre pro jednotlivé vzdělanostní kategorie (proměnná VZDELANI). Následná tabulka ukazuje, že průměrná hodnota na škále Q54f se u osob se základním vzděláním pohybuje 0,25 standardní odchylky nad průměrem souboru a u osob s vysokoškolským vzděláním 0,35 standardní odchylky pod průměrem souboru.

## Descriptives

| VZDELÁNÍ  |            |                                  | Statistic                  | Std. Error                       |                            |                         |
|---|------------|----------------------------------|----------------------------|----------------------------------|----------------------------|-------------------------|
| ZQ54F<br>Zscore:<br>Rozšířovat/<br>zachovat by<br>se mělo | 1 základní | Mean                             | .2504568                   | 5.24E-02                         |                            |                         |
|   |            | 95% Confidence Interval for Mean | Lower Bound<br>Upper Bound | .1473556<br>.3535579             |                            |                         |
|   |            | 5% Trimmed Mean                  |                            | .2434655                         |                            |                         |
|   |            | Median                           |                            | .1144482                         |                            |                         |
|   |            | Variance                         |                            | .956                             |                            |                         |
|   |            | Std. Deviation                   |                            | .9777083                         |                            |                         |
|   |            | Minimum                          |                            | -1.47699                         |                            |                         |
|   |            | Maximum                          |                            | 2.10375                          |                            |                         |
|   |            | Range                            |                            | 3.58074                          |                            |                         |
|   |            | Interquartile Range              |                            | 1.5914421                        |                            |                         |
|   |            | Skewness                         |                            | .173                             | .131                       |                         |
|   |            | Kurtosis                         |                            | -.781                            | .261                       |                         |
|   |            | 2 vyučen                         | 2 vyučen                   | Mean                             | 6.702E-02                  | 3.80E-02                |
|   |            |                                  |                            | 95% Confidence Interval for Mean | Lower Bound<br>Upper Bound | -7.544E-03<br>.1415742  |
| 5% Trimmed Mean   |            |                                  |                            | 3.964E-02                        |                            |                         |
| Median  |            |                                  |                            | .1144482                         |                            |                         |
| Variance  |            |                                  |                            | 1.102                            |                            |                         |
| Std. Deviation  |            |                                  |                            | 1.0495396                        |                            |                         |
| Minimum   |            |                                  |                            | -1.47699                         |                            |                         |
| Maximum   |            |                                  |                            | 2.10375                          |                            |                         |
| Range   |            |                                  |                            | 3.58074                          |                            |                         |
| Interquartile Range                                       |            |                                  |                            | 1.5914421                        |                            |                         |
| Skewness  |            |                                  |                            | .292                             | .088                       |                         |
| Kurtosis  |            |                                  |                            | -.812                            | .177                       |                         |
| 3 SŠ  | 3 SŠ       |                                  |                            | Mean                             | -.1310039                  | 4.04E-02                |
|   |            |                                  |                            | 95% Confidence Interval for Mean | Lower Bound<br>Upper Bound | -.2103018<br>-5.171E-02 |
|   |            | 5% Trimmed Mean                  |                            | -.1699205                        |                            |                         |
|   |            | Median                           |                            | -.2834123                        |                            |                         |
|   |            | Variance                         |                            | .891                             |                            |                         |
|   |            | Std. Deviation                   |                            | .9438370                         |                            |                         |
|   |            | Minimum                          |                            | -1.47699                         |                            |                         |
|   |            | Maximum                          |                            | 2.10375                          |                            |                         |
|   |            | Range                            |                            | 3.58074                          |                            |                         |
|   |            | Interquartile Range              |                            | 1.1935815                        |                            |                         |
|   |            | Skewness                         |                            | .467                             | .104                       |                         |
|   |            | Kurtosis                         |                            | -.507                            | .209                       |                         |
|   |            | 4 VŠ                             | 4 VŠ                       | Mean                             | -.3478791                  | 6.11E-02                |
|   |            |                                  |                            | 95% Confidence Interval for Mean | Lower Bound<br>Upper Bound | -.4684707<br>-.2272874  |
| 5% Trimmed Mean   |            |                                  |                            | -.3989504                        |                            |                         |
| Median  |            |                                  |                            | -.6812729                        |                            |                         |
| Variance  |            |                                  |                            | .707                             |                            |                         |
| Std. Deviation  |            |                                  |                            | .8409185                         |                            |                         |
| Minimum   |            |                                  |                            | -1.47699                         |                            |                         |
| Maximum   |            |                                  |                            | 2.10375                          |                            |                         |
| Range   |            |                                  |                            | 3.58074                          |                            |                         |
| Interquartile Range                                       |            |                                  |                            | 1.1935815                        |                            |                         |
| Skewness  |            |                                  |                            | .766                             | .177                       |                         |
| Kurtosis  |            |                                  |                            | .257                             | .352                       |                         |

## CVIČENÍ 3.6:

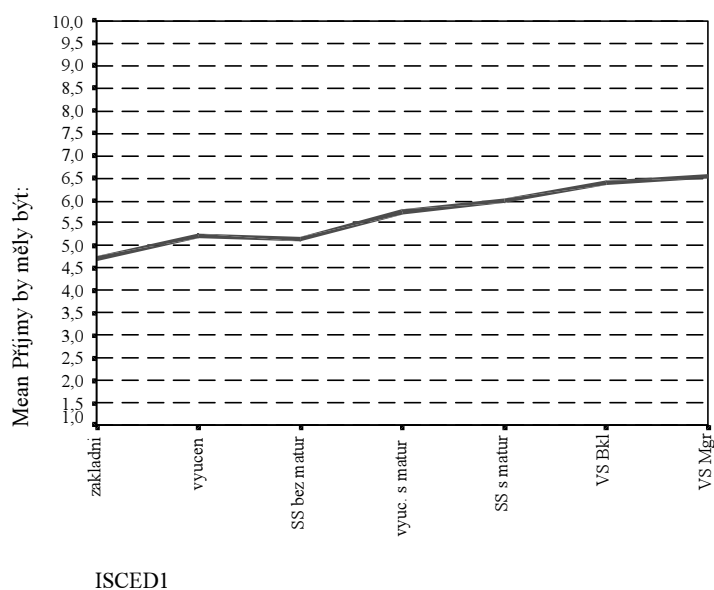
V datech zjistěte, zdali jsou rozdíly v postojích k příjmovým rozdílům (proměnná q54e) rozdílné u jednotlivých vzdělanostních skupin (proměnná isced1). Vypočítejte nejdříve průměry pro jednotlivé skupiny a zobrazte tyto průměry graficky. K odpovědi na otázku vypočítejte parametrický i neparametrický test a na základě testu normality se rozhodněte, který z vypočítaných testů je adekvátní.

Řešení:

## Report

Q54E Příjmy by mily být:

| ISCED1          | Mean | N    | Std. Deviation |
|-----------------|------|------|----------------|
| 2 základni      | 4,73 | 360  | 2,67           |
| 3 vyucen        | 5,20 | 602  | 2,85           |
| 4 SS bez matur  | 5,13 | 128  | 3,09           |
| 5 vyuc. s matur | 5,76 | 40   | 3,00           |
| 6 SS s matur    | 5,99 | 547  | 2,56           |
| 7 VS Bkl        | 6,38 | 23   | 2,26           |
| 8 VS Mgr        | 6,54 | 170  | 2,47           |
| Total           | 5,48 | 1870 | 2,77           |



## Parametrický test

## ANOVA

Q54E Příjmy by mily být:

|                | Sum of Squares | df   | Mean Square | F      | Sig. |
|----------------|----------------|------|-------------|--------|------|
| Between Groups | 625,004        | 6    | 104,167     | 14,134 | ,000 |
| Within Groups  | 13730,553      | 1863 | 7,370       |        |      |
| Total          | 14355,557      | 1869 |             |        |      |

## Multiple Comparisons

Dependent Variable: Q54E Přijmy by měly být:

Bonferroni

| (I) ISCED1      | (J) ISCED1      | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval |             |
|-----------------|-----------------|-----------------------|------------|------|-------------------------|-------------|
|                 |                 |                       |            |      | Lower Bound             | Upper Bound |
| 2 základni      | 3 vyucen        | -,47                  | ,18        | ,18  | -1,02                   | ,08         |
|                 | 4 SS bez matur  | -,40                  | ,28        | 1,00 | -1,25                   | ,45         |
|                 | 5 vyuc. s matur | -1,03                 | ,45        | ,49  | -2,42                   | ,35         |
|                 | 6 SS s matur    | -1,26*                | ,18        | ,00  | -1,82                   | -,70        |
|                 | 7 VS Bkl        | -1,65                 | ,58        | ,10  | -3,42                   | ,12         |
|                 | 8 VS Mgr        | -1,82*                | ,25        | ,00  | -2,58                   | -1,05       |
| 3 vyucen        | 2 základni      | ,47                   | ,18        | ,18  | -,08                    | 1,02        |
|                 | 4 SS bez matur  | ,08                   | ,26        | 1,00 | -,73                    | ,88         |
|                 | 5 vyuc. s matur | -,56                  | ,45        | 1,00 | -1,91                   | ,80         |
|                 | 6 SS s matur    | -,79*                 | ,16        | ,00  | -1,28                   | -,30        |
|                 | 7 VS Bkl        | -1,18                 | ,57        | ,85  | -2,93                   | ,57         |
|                 | 8 VS Mgr        | -1,34*                | ,24        | ,00  | -2,06                   | -,63        |
| 4 SS bez matur  | 2 základni      | ,40                   | ,28        | 1,00 | -,45                    | 1,25        |
|                 | 3 vyucen        | -,08                  | ,26        | 1,00 | -,88                    | ,73         |
|                 | 5 vyuc. s matur | -,63                  | ,49        | 1,00 | -2,14                   | ,87         |
|                 | 6 SS s matur    | -,87*                 | ,27        | ,02  | -1,68                   | -,06        |
|                 | 7 VS Bkl        | -1,25                 | ,61        | ,86  | -3,12                   | ,61         |
|                 | 8 VS Mgr        | -1,42*                | ,32        | ,00  | -2,38                   | -,45        |
| 5 vyuc. s matur | 2 základni      | 1,03                  | ,45        | ,49  | -,35                    | 2,42        |
|                 | 3 vyucen        | ,56                   | ,45        | 1,00 | -,80                    | 1,91        |
|                 | 4 SS bez matur  | ,63                   | ,49        | 1,00 | -,87                    | 2,14        |
|                 | 6 SS s matur    | -,23                  | ,45        | 1,00 | -1,59                   | 1,13        |
|                 | 7 VS Bkl        | -,62                  | ,71        | 1,00 | -2,78                   | 1,54        |
|                 | 8 VS Mgr        | -,78                  | ,48        | 1,00 | -2,24                   | ,67         |
| 6 SS s matur    | 2 základni      | 1,26*                 | ,18        | ,00  | ,70                     | 1,82        |
|                 | 3 vyucen        | ,79*                  | ,16        | ,00  | ,30                     | 1,28        |
|                 | 4 SS bez matur  | ,87*                  | ,27        | ,02  | ,06                     | 1,68        |
|                 | 5 vyuc. s matur | ,23                   | ,45        | 1,00 | -1,13                   | 1,59        |
|                 | 7 VS Bkl        | -,39                  | ,58        | 1,00 | -2,14                   | 1,36        |
|                 | 8 VS Mgr        | -,55                  | ,24        | ,43  | -1,28                   | ,17         |
| 7 VS Bkl        | 2 základni      | 1,65                  | ,58        | ,10  | -,12                    | 3,42        |
|                 | 3 vyucen        | 1,18                  | ,57        | ,85  | -,57                    | 2,93        |
|                 | 4 SS bez matur  | 1,25                  | ,61        | ,86  | -,61                    | 3,12        |
|                 | 5 vyuc. s matur | ,62                   | ,71        | 1,00 | -1,54                   | 2,78        |
|                 | 6 SS s matur    | ,39                   | ,58        | 1,00 | -1,36                   | 2,14        |
|                 | 8 VS Mgr        | -,16                  | ,60        | 1,00 | -1,99                   | 1,66        |
| 8 VS Mgr        | 2 základni      | 1,82*                 | ,25        | ,00  | 1,05                    | 2,58        |
|                 | 3 vyucen        | 1,34*                 | ,24        | ,00  | ,63                     | 2,06        |
|                 | 4 SS bez matur  | 1,42*                 | ,32        | ,00  | ,45                     | 2,38        |
|                 | 5 vyuc. s matur | ,78                   | ,48        | 1,00 | -,67                    | 2,24        |
|                 | 6 SS s matur    | ,55                   | ,24        | ,43  | -,17                    | 1,28        |
|                 | 7 VS Bkl        | ,16                   | ,60        | 1,00 | -1,66                   | 1,99        |

\*. The mean difference is significant at the .05 level.

**SOC108/708**  
**LEKCE 3: NORMÁLNÍ ROZLOŽENÍ**

**Neparametrický test**

Ranks

|                          | ISCED1          | N    | Mean Rank |
|--------------------------|-----------------|------|-----------|
| Q54E Příjmy by měly být: | 2 základni      | 359  | 796,47    |
|                          | 3 vyucen        | 604  | 873,60    |
|                          | 4 SS bez matur  | 130  | 826,03    |
|                          | 5 vyuc. s matur | 37   | 1058,27   |
|                          | 6 SS s matur    | 543  | 1030,47   |
|                          | 7 VS Bkl        | 26   | 1103,65   |
|                          | 8 VS Mgr        | 169  | 1167,33   |
|                          | Total           | 1868 |           |

Test Statistics<sup>a,b</sup>

|             | Q54E Příjmy by měly být: |
|-------------|--------------------------|
| Chi-Square  | 90,826                   |
| df          | 6                        |
| Asymp. Sig. | ,000                     |

a. Kruskal Wallis Test

b. Grouping Variable: ISCED1

**Dodatek 1:**

**Stanovte interval spolehlivosti pro odpovědi na otázku q7\_10 z výzkumu EVS-ČR1999, zdali by lidé nechtěli mít za sousedy lidi s nemocí AIDS. Kromě toho ve své odpovědi srovnajte český postoj k lidem s AIDS s ostatními evropskými zeměmi (viz tab. 2).**

Spočítáme si FREQUENCIES:

Q7\_10 Nemocní AIDS

|       |              | Frequency   | Percent      | Valid Percent | Cumulative Percent |
|-------|--------------|-------------|--------------|---------------|--------------------|
| Valid | 0 nezmínil/a | 1515        | 79,4         | 79,4          | 79,4               |
|       | 1 nechce     | 393         | 20,6         | 20,6          | 100,0              |
|       | <b>Total</b> | <b>1908</b> | <b>100,0</b> | <b>100,0</b>  |                    |

Spočítáme si v EXPLORE:

Descriptives

|                    |                                  | Statistic                                    | Std. Error |
|--------------------|----------------------------------|--|------------|
| Q7_10 Nemocní AIDS | Mean                             | 20,58  | ,93        |
|                    | 95% Confidence Interval for Mean | Lower Bound<br>18,77<br>Upper Bound<br>22,40 |            |
|                    | 5% Trimmed Mean                  | 17,32  |            |
|                    | Median                           | ,00  |            |
|                    | Variance                         | 1635,537                                     |            |
|                    | Std. Deviation                   | 40,44  |            |
|                    | Minimum                          | 0  |            |
|                    | Maximum                          | 100  |            |
|                    | Range                            | 100  |            |
|                    | Interquartile Range              | ,00  |            |
|                    | Skewness                         | 1,456  | ,056       |
|                    | Kurtosis                         | ,121   | ,112       |

95% interval spolehlivosti zahrnuje hodnoty 18,77 až 22,4. Pozor, jde o speciální případ výpočtu, jež je možný u dichotomické proměnné (považujeme ji za kardinální).

Tab. 2: Postoje k lidem s AIDS v různých evropských zemích

|                            |                       |       | V61 people with AIDS |             | Total  |
|----------------------------|-----------------------|-------|----------------------|-------------|--------|
|                            |                       |       | 0 not mentioned      | 1 mentioned |        |
| COUNTRY<br>country<br>code | 9901 France           | Count | 1476                 | 139         | 1615   |
|                            |                       | Row % | 91,4%                | 8,6%        | 100,0% |
|                            | 9902 UK               | Count | 759                  | 241         | 1000   |
|                            |                       | Row % | 75,9%                | 24,1%       | 100,0% |
|                            | 9903 Germany          | Count | 1602                 | 213         | 1815   |
|                            |                       | Row % | 88,3%                | 11,7%       | 100,0% |
|                            | 9905 Austria          | Count | 1241                 | 281         | 1522   |
|                            |                       | Row % | 81,5%                | 18,5%       | 100,0% |
|                            | 9906 Italy            | Count | 1379                 | 621         | 2000   |
|                            |                       | Row % | 69,0%                | 31,1%       | 100,0% |
|                            | 9907 Spain            | Count | 948                  | 252         | 1200   |
|                            |                       | Row % | 79,0%                | 21,0%       | 100,0% |
|                            | 9909 Netherl          | Count | 922                  | 81          | 1003   |
|                            |                       | Row % | 91,9%                | 8,1%        | 100,0% |
|                            | 9910 Belgium          | Count | 1688                 | 224         | 1912   |
|                            |                       | Row % | 88,3%                | 11,7%       | 100,0% |
|                            | 9911 Denmark          | Count | 960                  | 59          | 1019   |
|                            |                       | Row % | 94,2%                | 5,8%        | 100,0% |
|                            | 9913 Sweden           | Count | 948                  | 67          | 1015   |
|                            |                       | Row % | 93,4%                | 6,6%        | 100,0% |
|                            | 9916 Northern Ireland | Count | 681                  | 319         | 1000   |
|                            |                       | Row % | 68,1%                | 31,9%       | 100,0% |
|                            | 9917 Ireland          | Count | 763                  | 242         | 1005   |
|                            |                       | Row % | 75,9%                | 24,1%       | 100,0% |
|                            | 9918 Estonia          | Count | 589                  | 416         | 1005   |
|                            |                       | Row % | 58,6%                | 41,4%       | 100,0% |
|                            | 9919 Latvia           | Count | 719                  | 294         | 1013   |
|                            |                       | Row % | 71,0%                | 29,0%       | 100,0% |
|                            | 9920 Lithuani         | Count | 459                  | 559         | 1018   |
|                            |                       | Row % | 45,1%                | 54,9%       | 100,0% |
|                            | 9921 Poland           | Count | 592                  | 503         | 1095   |
|                            |                       | Row % | 54,1%                | 45,9%       | 100,0% |
|                            | 9922 Czechia          | Count | 1511                 | 397         | 1908   |
|                            |                       | Row % | 79,2%                | 20,8%       | 100,0% |
|                            | 9923 Slovakia         | Count | 738                  | 593         | 1331   |
|                            |                       | Row % | 55,4%                | 44,6%       | 100,0% |
|                            | 9924 Hungary          | Count | 233                  | 767         | 1000   |
|                            |                       | Row % | 23,3%                | 76,7%       | 100,0% |
|                            | 9925 Romania          | Count | 606                  | 540         | 1146   |
|                            |                       | Row % | 52,9%                | 47,1%       | 100,0% |
|                            | 9926 Bulgaria         | Count | 494                  | 506         | 1000   |
|                            |                       | Row % | 49,4%                | 50,6%       | 100,0% |
|                            | 9928 Greece           | Count | 835                  | 307         | 1142   |
|                            |                       | Row % | 73,1%                | 26,9%       | 100,0% |
|                            | 9929 Russia           | Count | 1168                 | 1332        | 2500   |
|                            |                       | Row % | 46,7%                | 53,3%       | 100,0% |
|                            | 9934 Slovenia         | Count | 672                  | 334         | 1006   |
|                            |                       | Row % | 66,8%                | 33,2%       | 100,0% |
|                            | 99031 SRN             | Count | 810                  | 112         | 922    |
|                            |                       | Row % | 87,9%                | 12,1%       | 100,0% |
|                            | 99032 DDR             | Count | 792                  | 101         | 893    |
|                            |                       | Row % | 88,7%                | 11,3%       | 100,0% |

Česká republika  
nezmíněno: .....1515  
zmíněno: .....393  
nezmíněno: ...79,4%  
zmíněno: .....20,6%

Poznámka: Všimněte si, že ve společném souboru jsme se dopočetli 20,8% a v národním souboru 20,6%, což jde asi na vrub jiného zacházení s missing values apod. Nehraje to velkou roli.

Česká populace nepatří ani k těm nejtolerantnějším (Francie, SRN, Nizozemí etc.), ale ani k těm s velmi nízkou tolerancí (Maďarsko, Rusko, Bulharsko, Slovensko, etc.) k postiženým osobám, spíše někam do středu (spolu s Velkou Británií, Rakouskem, Španělskem etc.).

**Dodatek 2:**

Stanovte intervaly spolehlivosti pro jednotlivé odpovědi na otázku q17\_4, zdali pracovat je povinností vůči společnosti.

Pozor! Tato úloha je jiného typu než předchozí, neboť jde o kategorizovanou proměnnou.

Q17\_4 Pracovat je povinnost

|         |                                | Frequenc<br>y | Percent | Valid Percent | Cumulative<br>Percent |
|---------|--------------------------------|---------------|---------|---------------|-----------------------|
| Valid   | 1 rozhodně souhlasí            | 358           | 18,8    | 19,0          | 19,0                  |
|         | 2 souhlasí                     | 831           | 43,6    | 44,0          | 62,9                  |
|         | 3 ani souhlas ani<br>nesouhlas | 368           | 19,3    | 19,5          | 82,4                  |
|         | 4 nesouhlasí                   | 278           | 14,6    | 14,7          | 97,1                  |
|         | 5 rozhodně souhlasí            | 55            | 2,9     | 2,9           | 100,0                 |
|         | Total                          | 1889          | 99,0    | 100,0         |                       |
| Missing | -2 neodpověděl/a               | 5             | ,2      |               |                       |
|         | -1 neví                        | 14            | ,7      |               |                       |
|         | Total                          | 19            | 1,0     |               |                       |
| Total   |                                | 1908          | 100,0   |               |                       |

Výpočet si zjednodušíme tím, že ho, podle předepsaného vzorce, provedeme v Excelu:

The screenshot shows an Excel spreadsheet with two tables. The first table is a pivot table showing the frequency and percentage distribution of responses for 'Q17\_4 Pracovat je povinnost'. The second table is a calculated table showing the 95% confidence interval for each response category.

|       |                        | Frequency | Percent | Valid Percent | Cumulative Percent |
|-------|------------------------|-----------|---------|---------------|--------------------|
| Valid | 1 rozhodně souhlasí    | 358       | 18,8    | 19,0          | 19,0               |
|       | 2 souhlasí             | 831       | 43,6    | 44,0          | 62,9               |
|       | 3 ani souhlas ani nesc | 368       | 19,3    | 19,5          | 82,4               |
|       | 4 nesouhlasí           | 278       | 14,6    | 14,7          | 97,1               |
|       | 5 rozhodně souhlasí    | 55        | 2,9     | 2,9           | 100,0              |
|       | Total                  | 1889      | 99,0    | 100,0         |                    |

| Q17_4 Pracovat je povinnost |       |            |      |       |       |
|-----------------------------|-------|------------|------|-------|-------|
|                             |       | std. Error | CI   | CI    |       |
|                             | Freq. | Valid %    | 95%  | dolní | dolní |
| 1 rozhodně souhlasí         | 358   | 19,0       | 1,81 | 17,2  | 20,8  |
| 2 souhlasí                  | 831   | 44,0       | 2,28 | 41,7  | 46,3  |
| 3. Ani, ani                 | 368   | 19,5       | 1,82 | 17,7  | 21,3  |
| 4 nesouhlasí                | 278   | 14,7       | 1,63 | 13,1  | 16,3  |
| 5 rozhodně nesouhlasí       | 55    | 2,9        | 0,77 | 2,1   | 3,7   |
| Total                       | 1889  | 100,0      |      |       |       |

Ve výše uvedené tabulce vidíme, v jakém intervalu se skutečně pohybují procenta osob s různou mírou souhlasu v základním souboru, odhadujeme-li je z výsledků našeho výběrového souboru, jenž byl z tohoto základního souboru náhodně vybrán (například je-li v našem výběrovém souboru 19% osob, které s předloženým výrokiem „rozhodně souhlasí“, víme že s 95% pravděpodobností je jich ve skutečnosti v našem základním souboru ne méně než 17,2% a ne více než 20,8%).