

LEKCE 3 ŘEŠENÍ

Úkol 3.1: V souboru EVS99_cvicny.sav zjistěte, zdali rozložení názoru na to, kdo by měl být odpovědný za zajištění bydlení (proměnná q54h), je normální. Řešte graficky i početně.

Descriptives

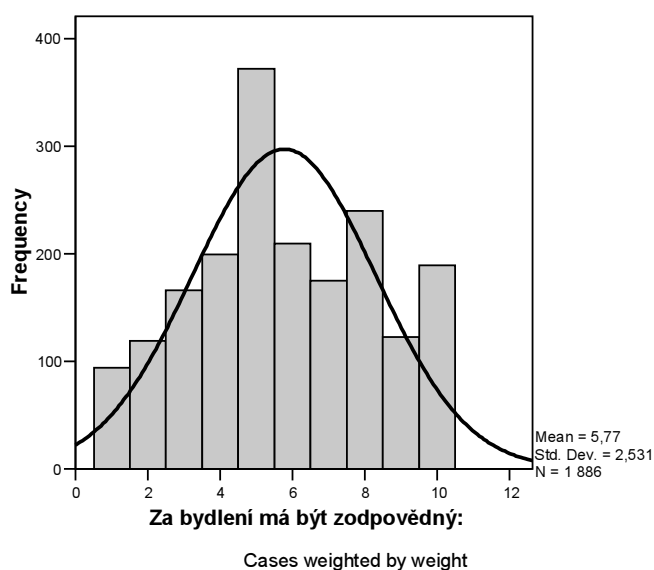
			Statistic	Std. Error
Za bydlení má být zodpovědný:	Mean		5,77	,058
	95% Confidence Interval for Mean	Lower Bound	5,65	
		Upper Bound	5,88	
	5% Trimmed Mean		5,80	
	Median		5,00	
	Variance		6,408	
	Std. Deviation		2,531	
	Minimum		1	
	Maximum		10	
	Range		9	
	Interquartile Range		4	
	Skewness		,003	,056
	Kurtosis		-,869	,113

Šikmost: blízká 0, tedy OK. Z-skór šikmosti: $0,003/0,056 = 0,05$, tedy OK.

Špičatost: Z-skór = 0,09, OK

Graf:

Histogram



Tests of Normality

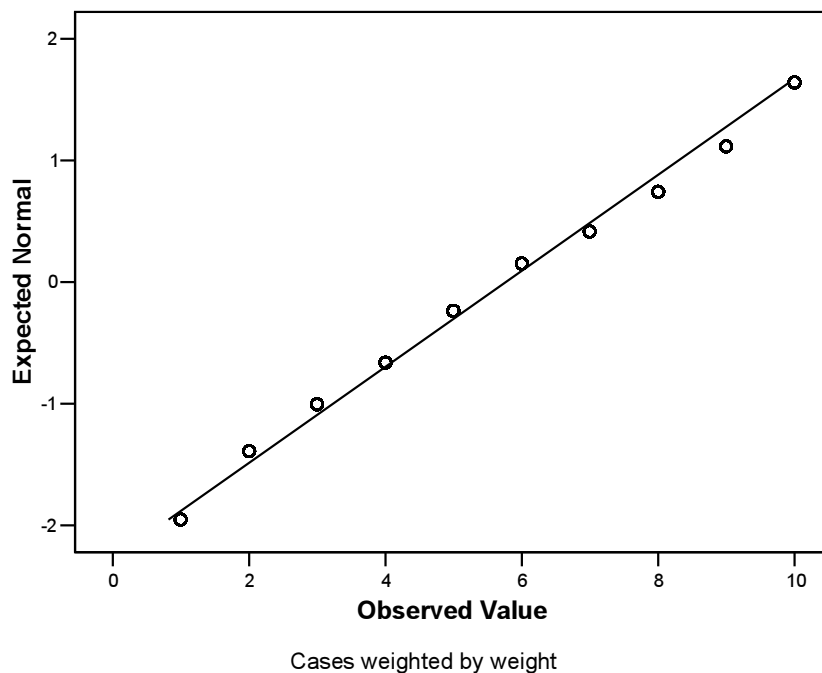
	Kolmogorov-Smirnov ^a		
	Statistic	df	Sig.
Za bydlení má být zodpovědný:	,123	1886	,000

a. Lilliefors Significance Correction

Signifikace vysoká, test říká, že rozložení není normální. Ale jelikož máme velké N (1896), není možné brát tuto signifikaci vážně. Neboť ve velkých souborech i malé odchylky od normálního rozložení vyjdou jako signifikantní.

Grafické řešení:

Normal Q-Q Plot of Za bydlení má být zodpovědný:



Vychází velmi hezky, body jsou rozmístěny kolem přímky.

Závěr: Dané rozložení můžeme považovat za normální a já osobně bych se nerozpakoval při analýzách na tuto proměnnou pustit parametrické postupy.

Úkol 3.2: Popište všechny základní charakteristiky věkového rozložení (proměnná vek) v tomto souboru a uveďte, která hodnota věku odděluje 20% nejstarších respondentů.

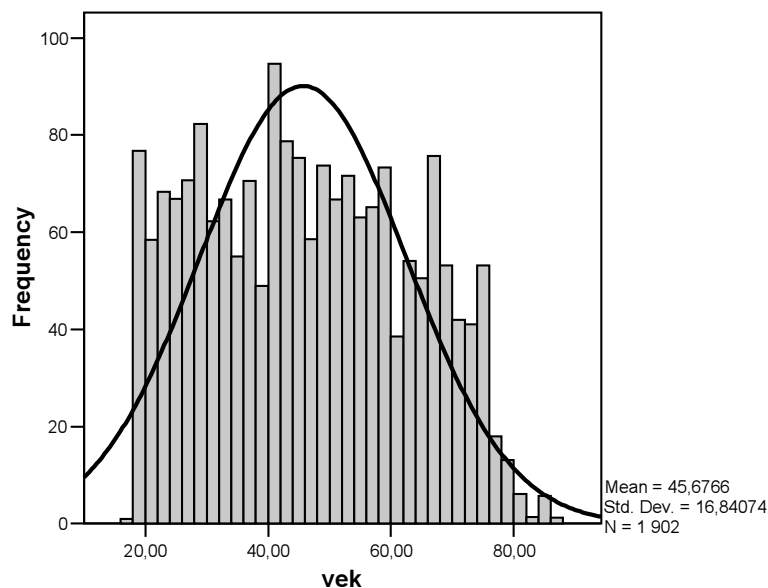
FREQUENCIES

```
VARIABLES=vek /FORMAT=NOTABLE
/PERCENTILES= 20
/STATISTICS=STDDEV VARIANCE MINIMUM MAXIMUM MEAN MEDIAN MODE SKEWNESS SESKEW K
URTOSIS SEKURT
/HISTOGRAM NORMAL
/ORDER= ANALYSIS .
```

Statistics

vek		
N	Valid	1902
	Missing	6
Mean		45,68
Median		45,00
Mode		18,00
Std. Deviation		16,84
Variance		283,6
Skewness		,142
Std. Error of Skewness		,056
Kurtosis		-1,030
Std. Error of Kurtosis		,112
Minimum		17,00
Maximum		88,00
Percentiles	20	28,00

Histogram



Cases weighted by weight

EXAMINE

```

VARIABLES=vek
/PLOT BOXPLOT NPLOT
/COMPARE GROUP
/PERCENTILES (5, 10, 25, 50, 75, 90, 95) HAVERAGE
/STATISTICS DESCRIPTIVES EXTREME
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.

```

Explore**Descriptives**

			Statistic	Std. Error
vek	Mean		45,6766	,38611
	95% Confidence Interval for Mean	Lower Bound	44,9194	
		Upper Bound	46,4339	
	5% Trimmed Mean		45,4578	
	Median		45,0000	
	Variance		283,610	
	Std. Deviation		16,84074	
	Minimum		17,00	
	Maximum		88,00	
	Range		71,00	
	Interquartile Range		28,00	
	Skewness		,142	,056
	Kurtosis		-1,030	,112

Extreme Values

			Case Number	Value
vek	Highest	1	1246	88,00
		2	1875	87,00
		3	1308	85,00
		4	1052	84,00
		5	1412	84,00 ^a
	Lowest	1	384	17,00
		2	387	18,00
		3	385	18,00
		4	377	18,00
		5	374	18,00 ^b

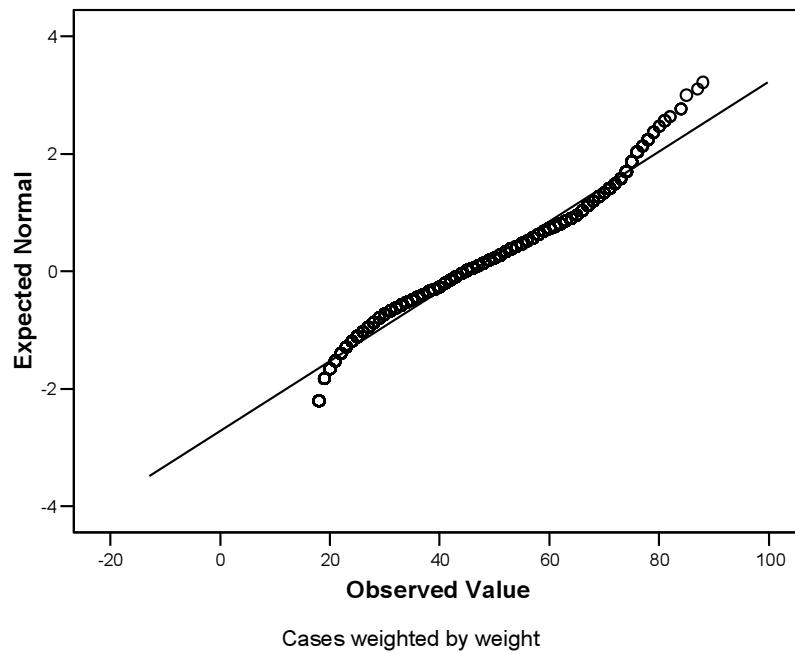
a. Only a partial list of cases with the value 84,00 are shown in the table of upper extremes.

b. Only a partial list of cases with the value 18,00 are shown in the table of lower extremes.

Tests of Normality

	Kolmogorov-Smirnov ^a		
	Statistic	df	Sig.
vek	,067	1902	,000

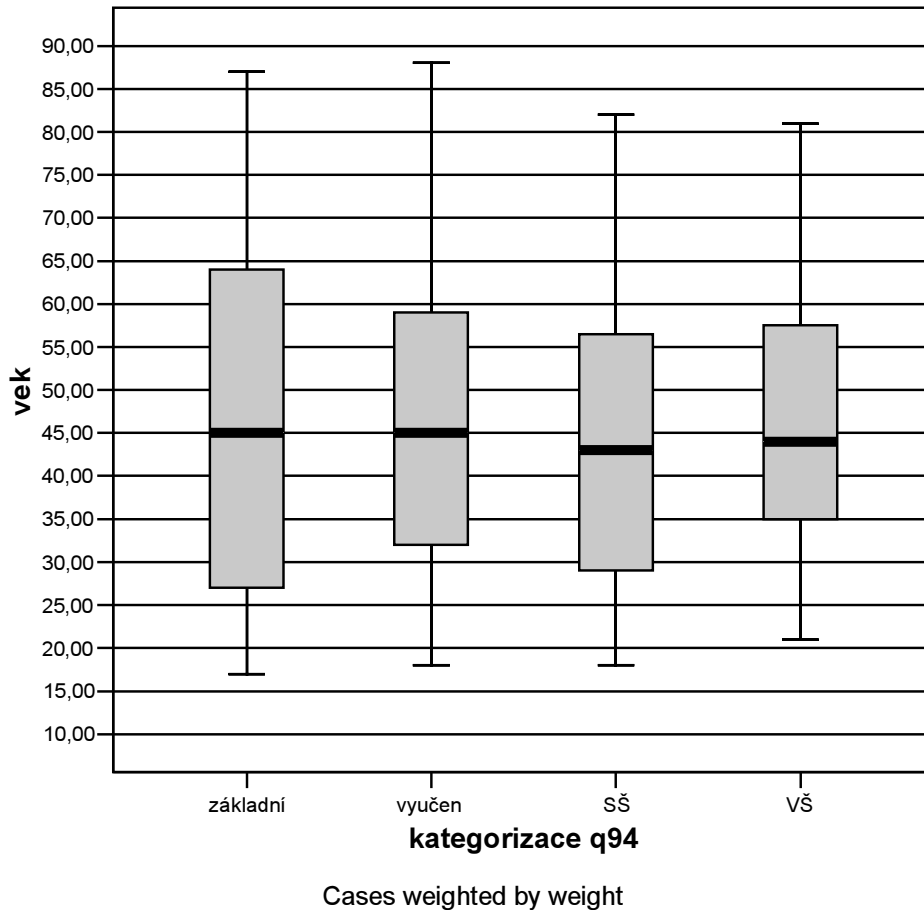
a. Lilliefors Significance Correction

Normal Q-Q Plot of vek

Úkol 3.3: Znázorněte věk graficky pro jednotlivé vzdělanostní kategorie (proměnná *vzdelani*) tak, abyste mohli porovnat jejich věkové mediány a interkvartilové rozpětí. Která ze vzdělanostních kategorií má nejvyšší medián a která největší interkvartilové rozpětí?

EXAMINE

```
VARIABLES=vek BY vzdelání /PLOT=BOXPLOT/STATISTICS=NONE/NOTOTAL.
```



Nejvyšší medián má základní vzdělání spolu s vyučenými (z grafu to moc nejde poznat, ale z tabulky z procedury *Tables – Basic Tables*, kde je výpočet mediánu, je jasné, že nejvyšší medián je u základního vzdělání. Ukažte tento numerický postup ve cvičení. Nejvyšší interkvartilové rozpětí má kategorie základní vzdělání, však má také největší směrodatnou odchylku.

		Mean	Median	Percentile 25	Percentile 75	Count	Col %	Std Deviation
kategorizace q94	základní	45,48	45,85	27,00	64,00	368	19,3%	19,76
	vyučen	46,58	45,00	32,00	59,00	782	41,1%	16,46
	SŠ	44,06	43,00	29,00	57,00	556	29,2%	15,90
	VŠ	46,90	44,00	35,00	58,00	197	10,3%	14,55

The screenshot shows the SPSS Data Editor window with a data table. The 'Analyze' menu is open, and the 'Basic Tables...' option is selected. The data table contains the following columns: vek, q109, q110, q1, v321, c51, c53, w, vzdělání, and vek. The rows represent data points for various years and locations.

Year	q109	q110	q1	v321	c51	c53	w	vzdělání	vek
1878	-3	1							
1879	4121	3	1						
1880	5220	2	1						
1881	2451	3	1						
1882	-3	1							
1883	3433	7	1						
1884	-3	1							
1885	-3	8							
1886	-3	1							
1887	-3	1							
1888	-3	7							
1889	8332	1							
1890	-3	2	1						
1891	-3	2	10000	2	2	2	2	2	2
1892	-3	2	10000	2	1	3	1	6	Bernartice
1893	-3	1	5050	2	3	4	6	3	Karlovy Vary
1894	-3	1	5400	4	3	2	6	3	Karlovy Vary
1895	-3	-2	-2	4	3	2	3	3	Nejdek
1896	-3	9	-2	2	2	3	6	3	Karlovy Vary
1897	-3	2	-2	5	4	1	7	5	Ústí nad Labem
1898	-3	3	-2	2	2	2	4	2	Nymburk
1899	7222	3	11000	2	1	2	2	2	Všetaty
1900	3439	7	-2	4	3	2	7	8	Ostrava-Poruba
1901	-3	2	10000	4	3	2	4	4	Český Krumlov
1902	-3	5	9000	2	2	3	1	8	Líšnice
1903	-3	3	10000	2	1	3	1	8	Líšnice
1904	-3	1	5250	2	1	3	6	5	Teplice
1905	-3	-2	-2	5	4	1	7	3	Plzeň
1906	-3	1	4000	2	1	3	7	7	Brno
1907	-3	-1	-2	2	2	1	7	8	Olomouc
1908	-3	7	-2	2	2	2	6	4	České Budějovice

The 'Basic Tables' dialog box is shown with the following settings:

- Summaries:** vek
- Subgroups:** Down: vzdělání
- Across:** (empty)
- Separate Tables:** (empty)
- Options:**
 - All combinations (nested)
 - Each separately (stacked)

Buttons on the right include OK, Paste, Reset, Cancel, Help, Statistics..., Layout..., Totals..., Format..., and Titles...

