

POUŽITÍ Z-SKÓRŮ

(příklad inspirován Andy Fieldem, 2004)

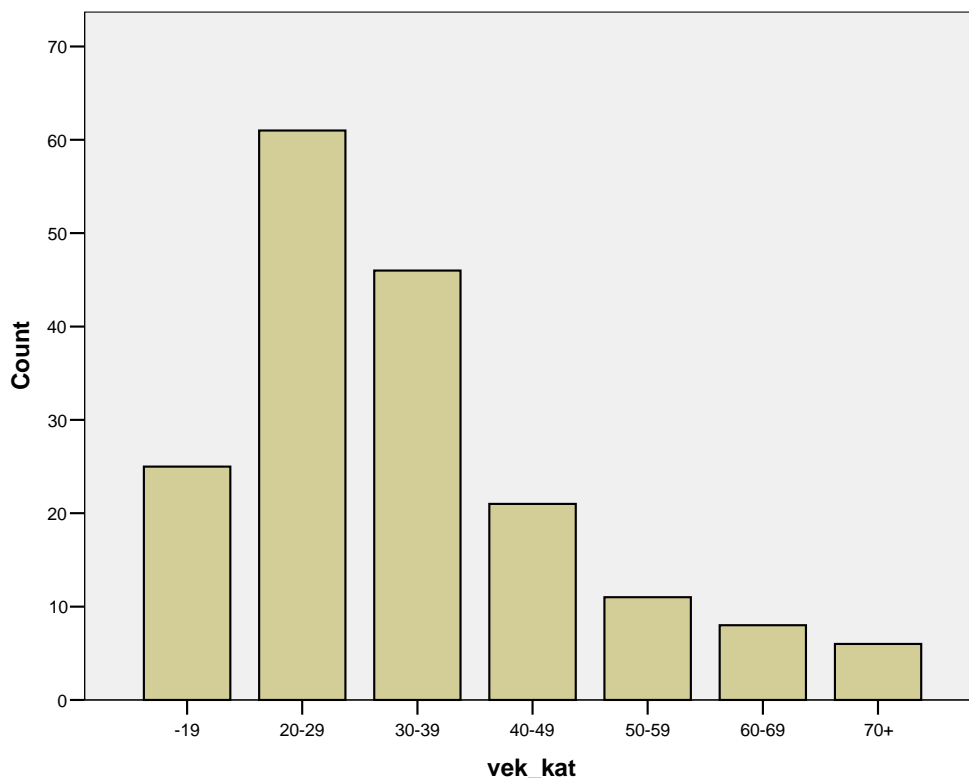
Ladislav Rabušic

Mějme data o sebevraždě – skoky z Nuselského mostu roce 2005 (fiktivní data, SPSS soubor je k dispozici v příloze).

Z individuálních dat jsme udělali novou, kategorizovanou proměnnou. Rozložení ukazuje tabulka: Celkem se v roce 2005 odehrálo 178 sebevražd v příslušném věkovém rozložení.

		vek_kat			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 -19	25	14,0	14,0	14,0
	2 20-29	61	34,3	34,3	48,3
	3 30-39	46	25,8	25,8	74,2
	4 40-49	21	11,8	11,8	86,0
	5 50-59	11	6,2	6,2	92,1
	6 60-69	8	4,5	4,5	96,6
	7 70+	6	3,4	3,4	100,0
Total		178	100,0	100,0	

Graf: Vidíme, že nejčastější věkovou skupinou, která páchala sebevraždy v daném roce byla skupina 20-29 let a že po šedesátce je počet sebevražd skokem z mostu už jen malý.



Na rozložení variant proměnné se můžeme dívat také z hlediska pravděpodobnosti a můžeme se ptát: Jaká je pravděpodobnost, že lidé ve věku 70 a více let spáchají sebevraždu skokem z mostu? Pohledem na graf nebo do tabulky řekneme: „No, velmi malá“. A jaká je pravděpodobnost, že to bude člověk kolem pětadvacítky? „Velmi vysoká.“ Rozložení tedy můžeme skutečně vnímat i z pohledu pravděpodobnosti, že nějaká situace nastane. My dokonce můžeme tuto pravděpodobnost přesně stanovit.

Využijeme k tomu vlastností standardizovaného normálního rozložení, tedy takového, které má průměr 0 a směrodatnou odchylku 1. Pro jeho jednotlivé hodnoty statistické spočítali pravděpodobnost výskytu.

Malým praktickým problémem je, že v empirických datech nikdy nenalezneme rozložení s průměrem 0 a směrodatnou odchylkou 1. Např. naše rozložení sebevražd má průměr, jak uvidíme za chvíli, 33,74 roků a směrodatnou odchylku 14,78. Co s tím? Pomůžeme si z-skóry.

Vzorec pro výpočet z-skórů:

$$Z = \frac{x_i - \bar{x}}{s_x}$$

x_i hodnota konkrétního případu

\bar{x} průměr

s_x směrodatná odchylka

Pro výpočet z-skórů tak musíme znát průměr a směrodatnou odchylku. Necháme je SPSS vypočítat.

Statistics

vek		
N	Valid	178
	Missing	0
Mean		33,74
Median		30,00
Std. Deviation		14,779

Průměr je 33,74, směr. odch. 14.78.

Nyní se můžeme ptát: Jaká je pravděpodobnost, že člověk, jemuž je 60 roků, spáchá sebevraždu skokem z nuselského mostu?

Řešení:

Vypočítáme z-skór pro 60-letého muže: $60 - 33,74 = 26,26$. Podělíme směrodatnou odchylkou: $26,26 / 14,78 = 1,78$.

Pro tuto hodnotu vyhledáme v tabulkách standardizovaného normálního rozložení příslušnou pravděpodobnost: je to 0,0375 tedy 0,038 neboli 3,8% šance, že 60letý muž spáchá sebevraždu skokem z nuselského mostu. Nebo řečeno jinými slovy, 96,2% šance, že člověk tohoto věku sebevraždu nespáchá.

Namísto obstarožního vyhledávání v tabulkách můžeme použít i z-skórovou kalkulačku, která je na této internetové adrese:

<http://www.fourmilab.ch/rpkp/experiments/analysis/zCalc.html>

Ukázka z výstupu:

The screenshot shows a web browser window titled "z score Calculator - Mozilla Firefox". The address bar shows the URL "http://www.fourmilab.ch/rpkp/experiments/analysis/zCalc.html". The page content is as follows:

Calculate probability Q from z

One of the most common probability calculations is determining, given the measured z value from an experiment or set of experiments, the probability of the result being due to chance. Enter the z value in the box below, press the **Return** key or the **Calculate** button, and the probability will appear in the Q box.

Given $z =$

The chance probability, Q , is: or about one in

Calculate z from probability Q

To determine the z score indicating a probability Q of non-chance occurrence for an experiment, enter Q in the box below and press the **Return** key or the **Calculate** button.

Given probability $Q =$

The browser's status bar at the bottom shows "Hotovo". The Windows taskbar at the very bottom includes the Start button and several open applications: Total Com..., z-skory - ..., Dokument..., and 2 Firefox.