

2 JEDNOROZMĚRNÁ POPISNÁ STATISTIKA – 2. ČÁST

Ke statistickému zkoumání je třeba mít k dispozici statistická data (údaje, pozorování), která získáváme prostřednictvím statistických šetření. *Šetření v sociálně-ekonomické oblasti* přitom vykazují určitá specifika, takže data z nich pocházející je třeba zpracovávat za pomoci vhodných statistických metod a postupů.

2.1 Zpracování dat

- *hodnoty proměnných = data = údaje = pozorování*
- většinou jde o větší množství údajů, které jsou více či méně nepřehledné
- prvním krokem při zpracování dat je proto jejich zpřehlednění (setřídění)
- zpřehlednění provádíme za pomoci *statistických tabulek a grafů*
- cílem tohoto procesu je, aby vynikly charakteristické rysy a zákonitosti souboru.

2.1.1 Statistické tabulky

Tabulka prostého rozdělení četností

- je vhodná při zpracování diskrétní proměnné s nepříliš velkým počtem obměn.

Tabulka 2.1

Obměna proměnné x_i	Četnost		Kumulativní četnost	
	absolutní n_i	relativní p_i	absolutní	relativní
x_1	n_1	p_1	n_1	p_1
x_2	n_2	p_2	$n_1 + n_2$	$p_1 + p_2$
.
.
.
x_k	n_k	p_k	n	1
Celkem	n	1	×	×

$$p_i = \frac{n_i}{\sum_{i=1}^k n_i} = \frac{n_i}{n} ; \quad \sum_{i=1}^k n_i = n ; \quad \sum_{i=1}^k p_i = 1$$

Tabulka intervalového rozdělení četností

- je vhodná v případě zpracování diskrétní proměnné s mnoha obměnami nebo spojitě proměnné
- klíčovým problémem je určení optimálního počet intervalů (k), na které rozdělíme variační rozpětí (R)
- k tomu slouží různá pravidla (např. Sturgesovo pravidlo: $k \approx 1 + 3,3 \log_{10} n$)
- při výpočtech lze pak každý interval zastoupit jeho středem
- výsledky takovýchto výpočtů jsou samozřejmě pouze přibližné.

2.2 Grafy

Další možností zřehlednění statistických údajů je jejich grafické znázornění. Existuje mnoho druhů grafů, vždy je třeba vybrat takový, který odpovídá charakteru dat.

Polygon četností

- spojnicový graf
- vhodný pro znázornění prostého rozdělení četností.

Histogram četností

- sloupkový graf
- vhodný pro znázornění intervalového rozdělení četností.

Výsečový graf (piechart)

- vhodný pro znázornění rozdělení četností nominální proměnné.

Sloupkový graf (barchart)

- vhodný pro znázornění rozdělení četností nominální proměnné.

Graf stem-and-leaf

- vhodný pro znázornění numerické proměnné s velkým množstvím obměn.

Krabicový graf

- vhodný pro znázornění extrémních hodnot a kvartilů
- užitečný pro srovnání určitého numerického znaku ve dvou či více souborech.