Renáta Bednárová

STATISTIKA PRO EKONOMY

ZÁKLADNÍ STATISTICKÉ POJMY

- Statistika
- Statistický soubor
- Statistická jednotky
- Statistický znak

STATISTIKA

- Vědní obor, který se zabývá hromadnými jevy
- Hromadné jevy jsou přírodní, společenské či jiné jevy sledované ne jednotlivě, ale ve velkém počtu případů.
- Popisná statistika – ryzí základy

STATISTICKÝ SOUBOR

- Množina všech objektů statistického pozorování shromážděných na základě toho, zda mají jisté společné vlastnosti (též. datový soubor).

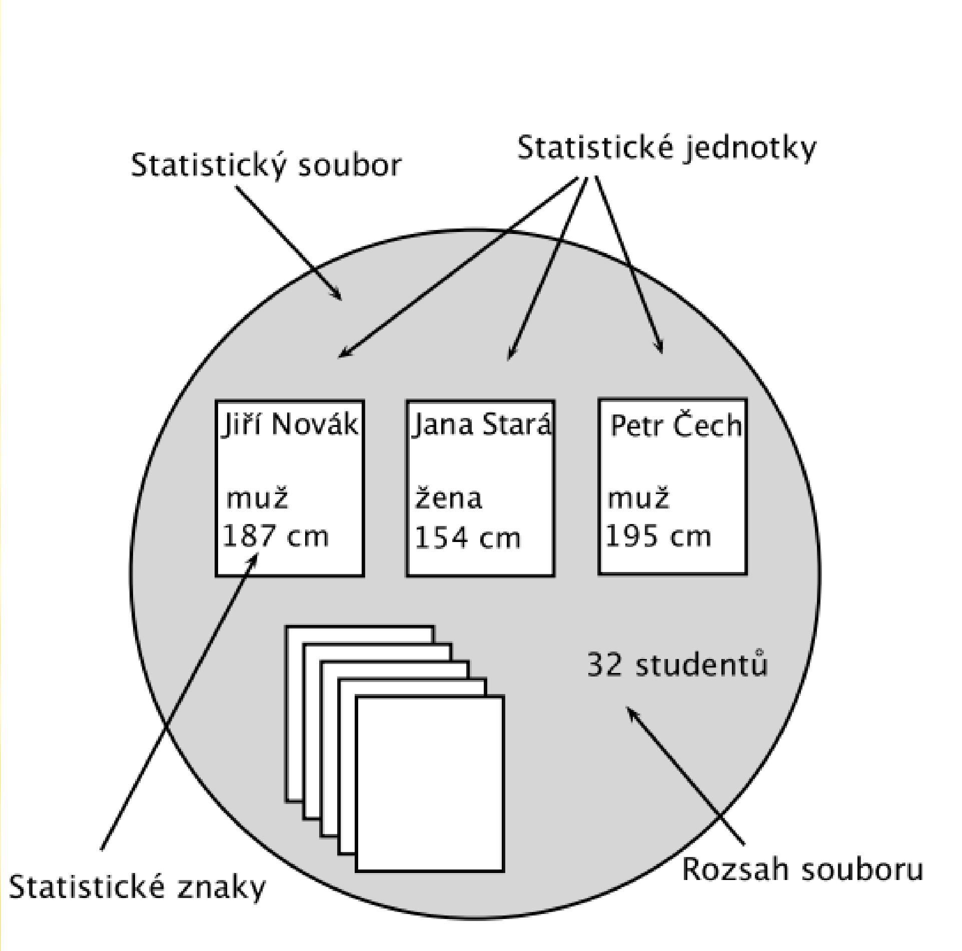
STATISTICKÁ JEDNOTKA

- Prvky množiny (statistického souboru).
- Značíme n

STATISTICKÝ ZNAK

- Společná vlastnost prvků statistického souboru, jejíž proměnlivost je předmětem statistického zkoumání.
- Značí se x
- Jednotlivé údaje statistického znaku se nazývají hodnoty znaku
- Dělíme na kvalitativní a kvantitativní

SCHÉMA ZÁKLADNÍCH STATISTICKÝCH POJMŮ



ROZDĚLENÍ ČETNOSTÍ

- Absolutní četnost
- Relativní četnost
- Skupinové rozdělení četnosti
- Formy grafického znázornění rozdělení četností

ABSOLUTNÍ ČETNOST

- Absolutní četností hodnoty x_j znaku x rozumíme počet n_j statistických jednotek jistého statistického souboru, které mají stejnou hodnotu x_j znaku x.
- Součet všech četností je rozsah statistického souboru.

RELATIVNÍ ČETNOST

- Relativní četnost hodnoty x_j je rovna podílu absolutní četnosti x_j a rozsahu n celého statistického souboru.
- Označujeme ji symbolem v_j
- Součet všech relativních četností všech různých hodnot statistického znaku je roven jedné
- RČ vyjadřujeme v praxi velice často v procentech.

TABULKA ROZDĚLENÍ ČETNOSTÍ A RELATIVNÍCH ČETNOSTÍ

- Pro přehledné znázornění absolutních a relativních četností hodnot znaků statistického souboru se využívají tabulky. Takovým tabulkám se říká tabulky rozdělení četností a relativních četností. Obecně tato tabulka může při zavedeném označení vypadat následovně:

TABULKA ROZDĚLENÍ ČETNOSTÍ A RELATIVNÍCH ČETNOSTÍ

x_j	n_j	v_j
x_1	n_1	$\frac{n_1}{n}$
x_2	n_2	$\frac{n_2}{n}$
$\dots$	$\dots$	$\dots$
x_k	n_k	$\frac{n_k}{n}$

PŘÍKLAD 1

U domácností byl zjišťován počet obytných místností s možnými výsledky: 1; 2; 3; 4; 5+ (5+ znamená 5 a více místností). Při tomto zjišťování bylo náhodné vybráno 25 respondentů s následujícími výsledky:

1,3,2,5+,2,2,3,5+,4,2,1,3,3,3,5+,2,3,3,4,4,2,3,4,2,3.

Vytvořte tabulku rozdělení četností a relativních četností.

SKUPINOVÉ ROZDĚLENÍ ČETNOSTÍ

- Statistický soubor o velkém rozsahu
- Zjednodušení souboru
- Hodnoty zkoumaného znaku, které jsou k sobě navzájem blízké, sdružovat do určitých skupin (tříd) tvořených zpravidla intervaly
- Zásady: konstantní šířka jednotlivých intervalů
- Sturgesův vzorec $k \doteq 1 + 3{,}3 \log n$, kde k je vhodný počet stejně velkých intervalů a n je rozsah statistického souboru

Počet třídicích intervalů v závislosti na rozsahu souboru určených Sturgesovým pravidlem

počet znaku	počet třídicích intervalů
1	1
2	2
3-5	3
6-11	4
12-23	5
24-46	6
47-93	7
94-187	8
188-376	9
377-756	10
...	...

PŘÍKLAD 2

U obcí Moravskoslezského kraje s počtem obyvatel větším než tisíc a menším než deset tisíc byl zjištěn počet narozených dětí za rok 2008. Výsledky jsou následující:

28; 28; 23; 51; 21; 25; 9; 6; 30; 18; 16; 15; 65; 14; 9; 40; 16; 23; 12; 21; 10; 10; 40; 38; 10;21; 31; 48; 19; 17; 16; 16; 11; 11; 27; 19; 20; 46:

Pomocí Sturgesova pravidla určete počet třídicích intervalů, vytvořte tabulku skupinového rozdělení četností a relativních četností. Závěrem se pokuste získané výsledky interpretovat.

FORMY GRAFICKÉHO ZNÁZORNĚNÍ ROZDĚLENÍ ČETNOSTÍ

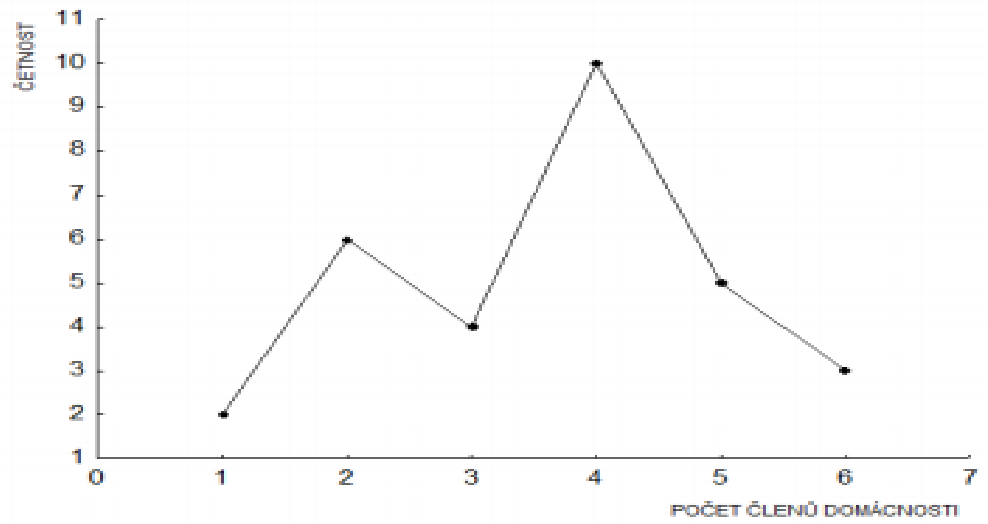
- V případě kvantitativního znaku se pro grafické znázornění četností používá především polygon četností nebo histogram
- Rozdělení četností hodnost kvalitativního znaku se graficky znázorňuje pomocí kruhového diagramu, kde různým hodnotám x_j statistického znaku odpovídají kruhové výseče, jejichž obsahy jsou přímo úměrné relativním četnostem v_j v procentech.

POLYGON ČETNOSTÍ

- Jinak zvaný spojnicový graf
- Propojení všech bodů v pravoúhlé soustavě, kde osa x vyjadřuje hodnoty znaků a osa y znázorňuje odpovídající četnost

POLYGON ČETNOSTÍ

počet členů domácnosti	1	2	3	4	5	6
počet domácností	2	6	4	10	5	3

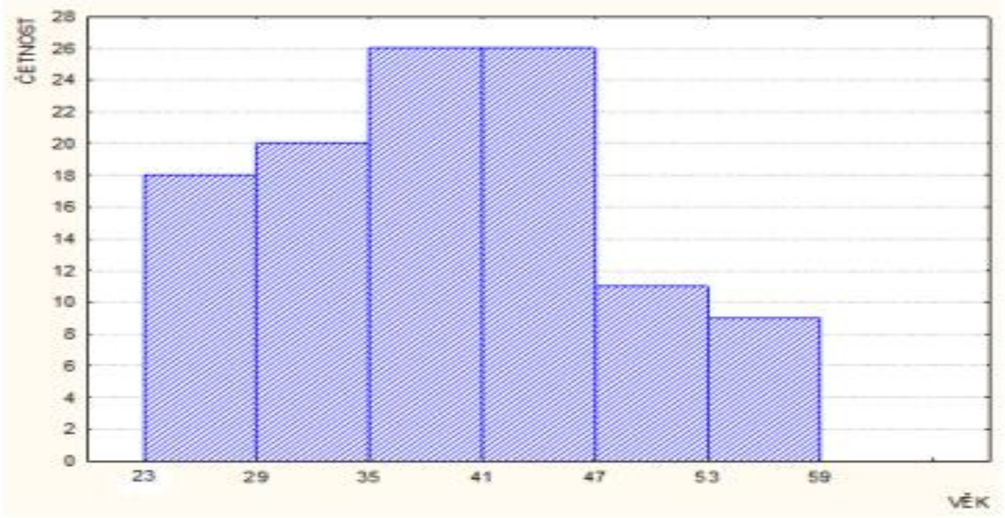


HISTOGRAM ČETNOSTÍ

- Jinak zvaný sloupkový diagram
- Osy stejně pojmenované jako u polygonu četností
- Používá se především pro skupinové rozdělení četností, je tvořen zpravidla rovnoběžníky, jejich základny mají délku zvolených intervalů a jejich výšky mají velikost příslušných třídních četností

HISTOGRAM ČETNOSTÍ

intervaly znaku x	četnosti n_j
$\langle 23, 29\rangle$	18
$(29, 35\rangle$	20
$(35, 41\rangle$	26
$(41, 47\rangle$	26
$(47, 53\rangle$	11
$(53, 59\rangle$	9
součet	110



KRUHOVÝ DIAGRAM

úroveň zájmu x	četnosti n_j	relativní četnosti v_j v %
jednoznačný nezájem	24	21,8
lehký nezájem	34	30,9
neutrální postoj	23	20,9
lehký zájem	21	19,1
jednoznačný zájem	8	7,3
součet	110	100 %

CHARAKTERISTIKY ZNAKU

- Typy statistických znaků podle stupně kvantifikace
- Aritmetický průměr
- Harmonický průměr
- Modus
- Medián
- Kvantily
- Rozptyl
- Směrodatná odchylka
- Variační koeficient
- Mezikvartilová odchylka

TYPY STATISTICKÝCH ZNAKŮ PODLE STUPNĚ KVANTIFIKACE

- Kvantitativní a kvalitativní
- Mezi kvalitativní znaky řadíme nominální a ordinální znaky
- Kvantitativní znaky naopak rozlišujeme intervalové a poměrové znaky

NOMINÁLNÍ ZNAKY

- Dovolují nám konstatovat pouze to, zda se nějaká hodnota statistického znaku rovná nějaké další hodnotě znaku jiné statistické jednotky, či nikoliv. Hodnotami mohou být slovní pojmenování nebo číselné kódy.

ORDINÁLNÍ ZNAKY

- jsou na škále znaků o stupínek výše než znaky nominální. Proto umožňují kromě posouzení rovnosti dvou hodnot znaku statistických jednotek také posouzení, zda je některá hodnota znaku větší nebo menší než hodnota znaku jiné statistické jednotky. To znamená, že hodnoty statistického znaku lze uspořádat do pořadí. U tohoto typu znaku nemá smysl se ptát, o kolik je daná hodnota znaku větší; smysl má pouze zeptat se, zda je větší nebo menší, případně zda se hodnoty rovnají. Ordinálními znaky mohou být čísla, text nebo i datum.

INTERVALOVÉ ZNAKY

- Jsou znovu o stupínek výše a lze s nimi provádět všechny operace, které jsme prováděli u znaků nominálních a ordinálních. Tyto znaky již patří do skupiny znaků kvantitativních. Intervalové znaky nám navíc umožňují interpretovat rozdíly mezi jednotlivými intervaly. To znamená, že rozdíl mezi hodnotami jedna a tři je stejný jako rozdíl mezi hodnotami pět a sedm. Intervalové znaky nám dovolují vypočítat, o kolik je jedna hodnota statistického znaku větší než druhá.

POMĚROVÉ ZNAKY

- jsou z těchto kategorií nejvýše, takže všechno, co platilo u znaků předchozích typů, platí i zde. Navíc u tohoto typu znaků platí, že stejný poměr mezi jednou dvojicí hodnot a druhou dvojicí hodnot znaku znamená i stejný podíl v rozpětí dané vlastnosti. Jinak řečeno má smysl uvažovat fakt, že hodnota je několikrát větší než jiná hodnota.

ALTERNATIVNÍ ZNAKY

- Alternativními znaky rozumíme znaky, které mohou nabývat pouze dvou hodnot a žádná další hodnota statistického znaku není možná. Do této kategorie bychom zařadili znaky, které vyjadřují přítomnost či absenci určitého jevu (přítomnost = 1, absence = 0) nebo například znak rozlišující pohlaví zkoumaných jedinců.

CHARAKTERISTIKA POLOHY

- Charakteristikami polohy (zvané také úrovně či střední hodnoty) hodnot znaku rozumíme čísla, která nám podávají určitou informaci o "střední hodnotě" sledovaného znaku. Patří mezi ně zejména aritmetický, geometrický a harmonický průměr, modus, medián a kvantily. Charakteristiky polohy se musíme naučit také vhodně volit. Kdybychom měli k dispozici statistický soubor a měli bychom ho jedním číslem charakterizovat, musíme obezřetně zvolit správnou charakteristiku, aby nedošlo k nesprávné představě o daném statistickém souboru

ARITMETICKÝ PRŮMĚR

Aritmetický průměr $\bar{x}$ hodnot $x_1, x_2, \ldots, x_n$ znaku x je definován jako podíl součtu hodnot znaku a jejich počtu (rozsahu souboru) n, tj. je určen vzorcem:

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

PŘÍKLAD 3

Na základě statistických hlášení o narození a úmrtí a souborů stěhování zpracovává Český statistický úřad v návaznosti na výsledky posledního sčítání lidu každoroční bilanci počtu obyvatel České republiky za všechny obce. My máme nyní k dispozici počet obyvatel v jednotlivých krajích (pro zlepšení práce s daty byly počty obyvatel zaokrouhleny a jsou uváděny v tisících):

Hlavní město Praha 1 233, Středočeský 1 230, Jihočeský 636, Plzeňský 569, Karlovarský 308, Ústecký 835, Liberecký 437, Královéhradecký 554, Pardubický 515, Vysočina 515, Jihomoravský 1147, Olomoucký 642, Zlínský 591, Moravskoslezský 1 250.

Vypočtěte aritmetický průměr počtu obyvatel žijících v jednom kraji České republiky.

HARMONICKÝ PRŮMĚR

Harmonickým průměrem x_H hodnot znaku $x_1, x_2, \ldots, x_n$ rozumíme podíl rozsahu souboru a součtu převrácených hodnot znaku, tj. platí:

$$\bar{x_H} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}} = n : \sum_{i=1}^{n} \frac{1}{x_i}.$$

PŘÍKLAD 4

- Auto jede rychlostí 100 km/h z bodu A do bodu B a nazpátek z bodu B do A rychlostí 160 km/h. Jaká je "korektní" průměrná rychlost, za jakou řidič auta urazil celou vzdálenost?

GEOMETRICKÝ PRŮMĚR

Geometrickým průměrem x_G hodnot zkoumaného znaku $x_1, x_2, \ldots, x_n$ rozumíme n – *tou* odmocninu ze součinu hodnot $x_1, x_2, \ldots, x_n$, proto:

$$\bar{x}_G = \sqrt[n]{x_1 x_2 \ldots x_n}.$$

PŘÍKLAD 5

Farmář má pole pro pěstování svých plodin obdélníkového tvaru z rozměry 9 m a 4 m. Nyní však farmář z nějakého důvodu potřebuje, aby mělo pole čtvercový tvar. Jaké jsou rozměry tohoto čtverce, jestliže má plocha zůstat stejná?

POUŽITÁ LITERATURA

- FIEDOR, David. Bakalářská práce - Statistika pro střední školy, Brno. Masarykova univerzita 2010.
- BUDÍKOVÁ, Marie; MIKOLÁŠ, Štěpán; OSECKÝ, Pavel. Popisná statistika. Brno: Masarykova univerzita v Brně, 2001.