**STATISTIKA PRO EKONOMY KURZ**

Úroveň Bc. stupeň studia

**2. TÉMA**

**Základní statistické charakteristiky**

**Cíle kapitoly:**

1. Charakteristika polohy či úrovně

2. Charakteristika variability

3. Základní vlastnosti průměru a rozptylu

**1. Charakteristika polohy či úrovně**

Charakteristiky polohy popisují obecnou polohu či úroveň znaku. Udávají střed rozdělení četností.

*Aritmetický průměr populační*

$$mí=\frac{1}{N}\sum\_{i=1}^{N}x\_{i}$$

Jaká část z celkového úhrnu připadne na jednotku. **Výhody** aritmetického průměru jsou, že bereme v úvahu všechny hodnoty v souboru a získáme pouze jediný výsledek všech hodnot. **Nevýhodou** však je, že bere v úvahu všechny hodnoty v souboru, včetně těch extrémních (malých i velkých hodnot). Střední hodnota stanovená právě průměrem má poté velmi **zkreslující informace**. Totéž platí pro aritmetický průměr výběrový.

*Aritmetický průměr výběrový*

$$\overbar{x}=\frac{1}{n}\sum\_{i=1}^{n}x\_{i}$$

*Vážený průměr*

**Vážený průměr** se stanoví z dat agregovaných v tabulce rozdělení četností. Více zobecňuje aritmetický průměr o příslušné váhy. Váhy představují důležitost hodnot v souboru.

Platí následující vzorec (váhy – absolutní četnosti ni):

$$\overbar{x}=\frac{1}{n}\sum\_{i=1}^{k}x\_{i}n\_{i}$$

Platí následující vzorec (váhy – relativní četnosti pi):

$$\overbar{x}=\sum\_{i=1}^{k}x\_{i}p\_{i}$$

*Harmonický průměr*

**Harmonický průměr** obecnou charakteristikou značí převrácenou hodnotu průměru aritmetického. Nejčastěji se aplikuje při výpočtu průměrné rychlosti na úsecích stejné délky.

Stanoví se dle následujícího vztahu:

$$\overline{x\_{h}}= \frac{n}{\sum\_{i=1}^{n}\frac{1}{x\_{i}}}$$

*Modus*

**Modus** značí vždy hodnotu, která se v daném statistickém souboru vyskytuje nejčastěji. Pokud je výsledkem pouze jedna hodnota, značíme tzv. pojmem „unimodální“. Pokud je výsledkem více hodnot, jde o tzv. „bimodální“.

*Medián*

**Medián** $\tilde{x}$ je prostřední hodnota v seřazeném souboru hodnot. Dělí soubor (znaky) na dvě části. Medián se považuje vždy za druhý kvartil v souboru (znaků), tedy 50 % kvartil.

Aritmetický průměr, medián a modus se považují za tzv. střední hodnotu.

*Kvartily*

**Dolní a horní kvartil** $Q\_{1}, Q\_{3}$ jsou hodnoty v 1. a 3. čtvrtině seřazeného souboru hodnot. Rozděluje tedy soubor (znaky) celkově na 4 části. Druhý kvartil (Q2) odděluje obě poloviny souboru a vždy je značen jako medián.

*Decily*

**Decily** rozdělují soubor na deset částí. Umožňuje se tak srovnání dosažených výsledků testování.

*Percentily*

**Percentily** mají podobnou vypovídací schopnost jako decily a kvartily, pouze s tím rozdílem, že je soubor členěn na sto částí.

*Krabicový graf*



**Krabicový graf** slouží jako grafický nástroj pro zobrazení povahy (polohy) hodnot v příslušném statistickém souboru. Obsahem jsou hodnoty extrémně odlehlé, uvedení maxima, minima a příslušné kvantily.

**2. Charakteristika variability**

Charakteristiky variability popisují měnlivost neboli rozptýlenost (variabilitu) hodnot znaku. Tedy tzv. homogenitu hodnot znaku. Malá variabilita znamená malou vzájemnou různost hodnot znaku, v tomto případě je průměr dobrou mírou. Vysoká variabilita značí velkou vzájemnou odlišnost hodnot znaku, pak průměr není dobrá míra.

**Rozptyl** nám udává, jak moc jsou hodnoty ve statistickém souboru rozptýleny.

Určuje se zvlášť pro základní a výběrový soubor.

 *Rozptyl pro výběrový soubor*

$$s^{2}=\frac{1}{n-1}\sum\_{i=1}^{n}\left(\overbar{x}-x\_{i}\right)^{2}$$

Základní soubor bere v úvahu veškeré měřené hodnoty. Z toho vyplývá, že ve jmenovateli nebude obsažen faktor, kdy od „n“ je odečítána 1. Tedy:

$$s^{2}=\frac{1}{n}\sum\_{i=1}^{n}\left(\overbar{x}-x\_{i}\right)^{2}$$

Směrodatná odchylka pro výběrový soubor se stanoví jako odmocnina z rozptylu. Matematické vyjádření je následující:

$$s=\sqrt{\frac{1}{n-1}\sum\_{i=1}^{n}\left(\overbar{x}-x\_{i}\right)^{2}}$$

Pro základní soubor platí následující vzorec:

$$s=\sqrt{\frac{1}{n}\sum\_{i=1}^{n}\left(\overbar{x}-x\_{i}\right)^{2}}$$

Důležité je si uvědomit, že rozptyl a směrodatná odchylka se zjišťuje zvlášť pro populační a výběrový soubor. V rámci populačního souboru není ve jmenovateli odečítána od „n“ mínus 1. Tento rozdíl slouží především k další statistické analýze, a to v rámci zjištění vybraných statistických charakteristik u intervalů spolehlivosti či jiných statistických metod.

Variační koeficient je charakterizován z různého úhlu pohledu. Z praktického hlediska slouží jako statistický nástroj pro určování výhodnosti investic. To znamená, že bere v úvahu jak faktor střední hodnoty, tak i riziko investice. Na druhé straně slouží jako nástroj pro identifikaci, zda analyzovaný statistický soubor v sobě neobsahuje příliš extrémní hodnoty. Pokud je variační koeficient ve výsledné hodnotě vyšší jak 50 %, potom aritmetický průměr není vhodné aplikovat pro zjištění střední hodnoty. Data jsou v tomto případě nehomogenní. To znamená, že nevýhody (vypovídací schopnost) aritmetického průměru jako střední hodnota má zkreslující informativní účinky. V tomto případě by bylo vhodnější aplikovat medián či modus.

Variační koeficient se stanoví dle vztahu:

$$v=\frac{s}{\overbar{x}}$$

**3. Základní vlastnosti průměru a rozptylu**

*Vlastnosti aritmetického průměru:*

* přičteme-li k jednotlivým hodnotám znaku konstantu, zvýší se o tuto konstantu i aritmetický průměr,
* aritmetický průměr konstanty je opět roven konstantě,
* násobíme-li jednotlivé hodnoty znaku konstantou, je touto konstantou násoben i průměr,
* součet jednotlivých odchylek od průměru je nulový,
* součet čtverců odchylek hodnot znaku od jeho aritmetického průměru je minimální.

*Vlastnosti rozptylu:*

* rozptyl je vždy v kladné hodnotě,
* populační vážený rozptyl lze rozložit na meziskupinový a vnitroskupinový rozptyl,
* rozptyl konstanty je roven nule,
* přičteme-li ke všem hodnotám znaku konstantu, rozptyl se nezmění,
* násobíme-li všechny hodnoty znaku konstantou, rozptyl je násoben čtvercem této konstanty.

**Shrnutí:**

1. Charakteristika polohy či úrovně 2. Charakteristika variability 3. Základní vlastnosti průměru a rozptylu

Průměry, medián, modus, kvantily.

Populační, výběrový rozptyl, směrodatná odchylka, variační koeficient.





****

****

**Zdroje:**

ADAMEC, V., L. STŘELEC, a D., HAMPEL, 2017. Ekonometrie I: učební text. Druhé nezměněné vydání. Brno: Mendelova univerzita v Brně. ISBN 978-80-7509-480-3. (s. 21-22)

HINDLS, R., 2018. Statistika v ekonomii. [Průhonice]: Professional Publishing. ISBN 978-80-88260-09-7. (s. 21-37)

MOŠNA, F., 2017. Základní statistické metody. Praha: Univerzita Karlova v Praze. ISBN 978-80-7290-972-8. (s. 4-9)

STUCHLÝ, J., 2015. Statistické analýzy dat: vysokoškolská učebnice. České Budějovice: Vysoká škola technická a ekonomická v Českých Budějovicích. ISBN 978-80-7468-087-8. (s. 16-32)