

Štatistické pojmy

Posledná aktualizácia: 28.05.2020

Pre jednoduchšie pochopenie štatistiky je potrebné vysvetliť niektoré základné štatistické pojmy.

Štatistický súbor

Štatistika sa zaoberá javmi, ktoré nazývame **hromadné javy**. Rozlišujeme dva druhy hromadných javov. Jeden druh je taký, čo sa vyskytuje vo veľkom počte objektov - výška, hmotnosť, farba očí, pohlavie, vek, profesia, počet detí a pod. Druhý druh je taký, ktorý je výsledkom veľkého počtu opakovaní, napr. opakované váženie nejakého predmetu, opakované hody kockou, mincou a pod. Objekty, ktoré skúma štatistika voláme **štatistický súbor** - napr. množina osôb, ktoré žijú na nejakom území.

Základný štatistický súbor rozsahu **N** predstavuje množinu všetkých štatistických jednotiek. V prípade, že nemožno skúmať základný súbor (z časových, finančných alebo iných dôvodov), vytvárame z neho **výberový súbor** podľa vopred stanovených pravidiel. V prieskume sa údaje zbierajú len za časť populácie, tzv. vzorku. Tieto údaje sa potom použijú na odhad charakteristík celej populácie. V tomto prípade sa musí zabezpečiť, aby vzorka reprezentovala príslušnú populáciu. Napríklad podiel osôb vo veku do 18 rokov alebo podiel žien a mužov vo vybranej vzorke domácností musí odrážať realitu v celkovom počte obyvateľov.

Výber musíme urobiť tak, aby vybrané objekty mali vlastnosť, ktorú má aj celý základný súbor. Túto vlastnosť nazývame **štatistický znak** (môže to byť pohlavie, profesia, počet detí, vek, vzdelanie a pod.). Rôzne hodnoty štatistického znaku voláme obmeny, varianty.

Štatistický znak

Štatistické znaky možno deliť podľa rôznych hľadísk. Základné delenie je na znaky kvalitatívne a kvantitatívne:

Kvalitatívne (kategoriálne) znaky slovne vyjadrujú vlastnosť štatistickej jednotky. Hodnoty kvalitatívnych znakov sa zvyknú označovať číslami, okódujú sa. Kvalitatívne znaky delíme na:

- **nominálne - názvové znaky**, ich hodnoty môžeme pomenovať, ale nevieme ich zoradiť do poradia. Môžeme o nich povedať, že sú rôzne alebo sa rovnajú (pohlavie, farba očí, štátna príslušnosť);
- **ordinálne - poradové znaky**, ich hodnoty môžeme prirodzene usporiadať do poradia, ale nevieme určiť, o koľko je jedna hodnota väčšia ako druhá (medaila - zlatá, strieborná, bronzová, hodnosť v armáde, kvalitatívne hodnotenie študenta - výborný, veľmi dobrý, dobrý, nevyhovet môžeme vyjadriť číselným hodnotením 1, 2, 3, 4).

Kvantitatívne (kardinálne, číselné) znaky sú reálne čísla, môžeme o ich hodnotách povedať, či sú rovnaké alebo rôzne, vieme ich usporiadať do poradia a vieme tiež určiť, o koľko je jedna hodnota väčšia ako druhá. Tieto znaky majú aj nulovú hodnotu a meraciu jednotku. Kvantitatívne znaky delíme na:

- **diskrétné znaky** – majú konečný, ale spočítateľný počet obmien a všetky obmeny sa dajú očíslovať prirodzenými číslami 1, 2, 3 napríklad počet detí v rodine, počet bodov v teste;

- **spojité znaky** – môžu mať rôznu číselnú hodnotu z istého intervalu (telesná výška, príjem).

Intervalové znaky (rozdielové) sú tie, ktorých hodnoty môžeme usporiadať do poradia, vieme určiť, o koľko je jedna hodnota väčšia ako druhá a aké sú rozdiely (intervaly) medzi hodnotami, ale nevieme určiť, koľkokrát je jedna hodnota väčšia ako druhá. Intervalový znak nemá prirodzený nulový bod, ale nulový bod možno určiť dohodou. Preto sa hodnoty intervalového znaku môžu sčítovať alebo odčítavať, ale nemôžu sa násobiť a deliť. Môžu nadobúdať kladné aj záporné hodnoty, napríklad: znak „Celziova teplota“ má dohodnutú nulu – teplotu topenia ľadu.

Pomerové znaky (podielové) sú špeciálne prípady intervalových znakov. Možno ich usporiadať do poradia a vieme určiť rozdiely medzi hodnotami. Majú tiež prirodzený nulový bod, absolútnu 0, ktorá znamená, že znak neexistuje. Ich hodnoty nemôžu byť záporné, možno ich sčítovať, odčítavať, násobiť aj deliť. Napríklad znak vek má absolútnu nulu 0 rokov. 60 ročný človek má o 40 rokov viac ako 20 ročný a pritom môžeme povedať, že je 3 krát starší ako dvadsiatnik.

Zisťovanie hodnôt štatistického znaku v štatistickom súbore sa volá štatistické zisťovanie. Pri štatistickom spracovávaní údajov často nahrádzame kvalitatívny znak kvantitatívnym.
(napríklad kvalitatívny znak pohlavie: muži = 1, ženy = 2).

Znak, ktorý má iba jednu nemennú hodnotu sa volá **konštanta**. V štatistike sa sledujú hlavne znaky, ktoré majú aspoň dve alebo viac obmien. Takýto znak má názov **premenná**.

Štatistické triedenie

Údaje získané štatistickým zisťovaním najlepšie zjednodušíme a sprehľadníme tým, že ich usporiadame do určitých skupín – tried.

Metóda usporiadania štatistických údajov do určitých skupín (tried) podľa určitého znaku sa nazýva triedenie.
triedenie = rozdelenie štatistických jednotiek do takých skupín (tried), aby čo najlepšie vynikli charakteristické vlastnosti skúmaných javov;
triediaci znak = štatistický znak, ktorý je kritériom triedenia štatistického súboru;
trieda = skupina štatistických jednotiek s rovnakou hodnotou (variantom) znaku.