

I.A.4 Několik rad, jak číst statistiky

- "přesná" čísla ve statistických zprávách**
- a) Dávejte si pozor na "přesná" čísla; typickým příkladem může být např. sčítání lidu, kdy čekáme - obvykle na půlnoc, abychom mohli prohlásit, že v naší republice žije 15832714 obyvatel. Bereme-li tato čísla za odhad počtu obyvatel, jsou naše zkoumání stejně legrační jako odhad příchodu Járy Cimrmana do Liptákova, "... pomocí analýzy prachu na jeho botách odhadnut na 2.12.1910 ve 14.00 plus minus 200 let" (hra Posel z Liptákova). Magie těchto "přesných" čísel ovládá skoro každého z nás, přitom si málokdo uvědomuje, jakých nepřesností jsme se dopustili. Zaokrouhlená hodnota, či vymezení intervalu možných hodnot (uvedení odhadu chyby) nepůsobí sice tolik přesvědčivě, ale je rozhodně blíže pravdě.
- procenta versus absolutní čísla**
- b) Samotná procenta říkají velmi málo nebo vůbec nic. Kdo zamlčuje absolutní čísla nebo neuvádí údaje pro srovnání, většinou chce něco skrýt. Většinou platí, že pan X uvádí jen takové údaje, které se hodí pro jeho tvrzení. Např. je-li jeho údaj (třeba produkce podniku) malý, je pro něj výhodnější uvádět údaj v procentech (např. zvýšení prodeje), který sám o sobě může být srovnatelný s ostatními, ale absolutní číslo by bylo velmi nepříznivé. Celkem snadno lze podvádět i obráceně. Všechna miliónová a miliardová čísla vypadají velmi působivě a lze je s úspěchem použít tam, kde poměrná čísla, či procenta nejsou přesvědčivá. Například deficit jednotlivých republik ve výši cca 17 miliard musí pro neinformovaného čtenáře vypadat hrozivě, naopak zjištění, že se jedná o 0,5 % národního produktu staví tento údaj do trochu jiné roviny.
- srovnatelné údaje**
- c) Porovnání údajů mezi sebou je snad nejdůležitější. V našem příkladu, kdy jsme zkoumali jaký je plat u žen, může závěr znít různě, ale z obecnějšího hlediska je skoro k ničemu; zajímavější údaj je třeba kupní síla (v porovnání s ostatními zeměmi), změna (vývoj) reálných mezd v čase, nebo alespoň porovnání s úrovní mezd u mužů atd.
- nezaměňovat ukazatele**
- d) Dávejte si pozor na nedbalý výklad uvedených hodnot. Zejména nesmíme zaměňovat minimální (maximální) hodnoty s průměrem, průměr s nejčastějšími hodnotami a naopak. Zjistíme-li, např., že průměrná výška (v našem výběru) je třeba 176 cm, může se stát, že nikdo v tomto výběru není vysoký právě 176 cm. Častou chybou je také přeceňování

modu rozdělení (zejména pro výběry malého rozsahu). V našem výběru může být nejčastější hodnota (modus) obsažena pouze dvakrát, např. 191 cm, což je spíše zavádějící údaj. Podobně se dá zacházet s čísly o produkci, dovozu, vývozu, atd., kde se většina lidí dá snadno poplést těmito údaji.

je porovnání
vhodné ?

e) Při zprávách pozor na formulace "... sice menší (větší) než, ale už (nyní následuje příznivý ukazatel). Téměř vždy se dá najít nějaký "příznivější" statistický ukazatel. Pokud je údaj s něčím porovnáván, nabízí se otázka, zda je to oprávněné, a zda má toto porovnání vůbec smysl (pozor při porovnávání průměrů - např. průměrné teploty mohou být stejné, ale tento údaj nás neopravňuje tvrdit, že dvě porovnávaná místa mají stejné podnebí). S tím souvisí otázka, zda se ve zprávě netvrdí víc, než se dá z předložených čísel poznat.

pozor na výběrová
zjišťování

f) Poslední poznámka se týká výběrových zjišťování. Pokud výběr nebyl proveden profesionály, všimněte si, zda výběr zachovává strukturu základního souboru (pokud se šetření týká např. národnostní (sociální) problematiky, potom musí respektovat složení obyvatelstva).

Důležité jsou i otázky - na některé dostanete vždy stejné (nebo zkrácené) odpovědi. Např. na otázku kolikrát denně se myjete, zda berete drogy atd. Pokud nás zajímá pravdivá odpověď, musíme ji zjišťovat nepřímo.

Dalším a častým druhem mystifikace je vlastní příprava experimentu - zjišťování. Např. se dá "prokázat", že v nemocnici umírají lidé častěji než doma. Zastíráme ovšem fakt, že v nemocnici leží pacienti vážněji nemocní, s různými komplikacemi, zatímco doma se léčí většinou jen lehce nemocní. Podobných příkladů je celá řada, pěkně se "konstruuje" zejména porovnání úspěšnosti léčení různých chorob doma a v nemocnici (ambulantně, či operativně). Příklady nalézáme i v tisku. J. Machek upozornil na krásný příklad možného statistického podvodu, kdy bylo na datech ukázáno, že lidé, kteří pijí denně alespoň sklenku alkoholu, se dožívají delšího věku (New York Times, září 1990). Proč se může jednat o skrytý podvod? Vtip je totiž v tom, že ti, kteří nepijí, k tomu mají většinou zdravotní důvody !!

Termíny

Nové termíny: výběrový (aritmický) průměr, medián, modus, kvantily, kvartily, decily, percentily, variabilita, rozpětí, mezikvartilové rozpětí, rozptyl, směrodatná odchylka, absolutní