

Statistika

Základní pojmy a cíle statistiky

Roman Biskup

(zapálený) statistik ve výslužbě, aktuálně analytik v praxi ;-)
roman.biskup(at)email.cz

12. února 2012



Obsah

Základní pojmy

Definice

Postup statistického zkoumání

Deskriptivní statistika

Induktivní statistika

Statistické šetření

Výběrová šetření



Definice Statistky jako předmětu

Tři významy pojmu statistika

- ▶ Vědní disciplína
- ▶ Praktická činnost
- ▶ Statistické údaje

#1

Statistiku lze definovat jako vědní obor, zabývající se metodami kvantitativního hodnocení vlastností **hromadných jevů** a **procesů**.

#2

Statistika zahrnuje jak získávání, tak i analýzu a interpretaci napozorovaných dat. Cílem statistického zpracování dat je podání informace o vlastnostech, povaze a zákonitostech projevujících se na pozorovaných datech.

#3

popisné statistiky, náhodné veličiny a jejich hodnoty



Definice pojmů I

Hromadný jev je jev, který se vyskytuje v masovém měřítku a může se neustále opakovat (alespoň 30 pozorování):

- ▶ opakované měření sledované vlastnosti jednoho předmětu, nebo
- ▶ sledování vlastností na předem definované množině objektů
- ▶ obyčejné pozorování, nebo
- ▶ experiment – „řízené pozorování“

Statistická jednotka je nositel sledovaných vlastností (**statistických znaků**) vymezených:

- ▶ prostorově,
- ▶ časově,
- ▶ věcně.

Statistický soubor je množina sledovaných statistických jednotek vyhovující prostorovému, časovému a věcnému hledisku;

- ▶ základní statistický soubor (populace) – konečný nebo nekonečný,
- ▶ výběrový statistický soubor (výběr) – vždy konečný.

Rozsah souboru je počet pozorování v statistickém souboru (n , N).



Definice pojmů II

Statistický znak je nositelem určité vlastnosti každé statistické jednotky souboru; hodnota statistického znaku se často nazývá pozorování.

- ▶ konstantní (identifikační) – vymezují statistický soubor,
- ▶ proměnlivý (variabilní) – předmětem zkoumání.
- ▶ měřený přímo,
- ▶ měřený nepřímo.
- ▶ nominální znak (alternativní / pomnožný),
- ▶ ordinální znak (kvalitativní znak, zavedené uspořádání),
- ▶ metrický znak (stupnice bez referenční hodnoty),
- ▶ kardinální znak (stupnice s referenční hodnotou) – matematické operace jsou relevantní:
 - ▶ diskrétní (konečný nebo spočetný),
 - ▶ spojitý (nespočetný).



Etapy statistického zkoumání

Statistické šetření

Vytvoření datového souboru

POPULACE

VÝBĚR



Statistické zpracování

Vyhodnocení datového souboru

Deskriptivní statistika

Induktivní statistika



Statistický rozbor

Interpretace výsledků s důsledným zamyšlením nad kvantitativním hodnocením
objektu statistického šetření

výsledky deskriptivní statistiky

výsledky induktivní statistiky



Deskriptivní statistika

- ▶ Popis statistických souborů;
- ▶ možnost vzájemného srovnání:
 - ▶ míry polohy, variability, šikmosti, špičatosti
 - ▶ rozdělení,
 - ▶ indexy;
- ▶ vlastnosti skupiny → jedinec, skupina jedinců (obecné → partikulární)



Induktivní statistika

- ▶ Vyslovování závěrů o statistickém souboru na základě výběru:
 - ▶ odhady parametrů,
 - ▶ ověřování (testování) hypotéz,
 - ▶ modelování závislostí,
 - ▶ vysvětlování souvislostí;
- ▶ výběr → vlastnosti skupiny (partikulární → obecné)



Statistické šetření jako podklad statistické indukce

► Populační šetření (proč není často realizovatelné):

 N

1. základní soubor je jen hypotetický,
2. destrukce statistické jednotky při měření statistického znaku,
3. časová a finanční náročnost měření a zpracování.

► Výběrové šetření

 n

- Cíl výběrového šetření: Prostřednictvím výběrového souboru odhadnout vlastnosti základního souboru.
- Prostředky výběrového šetření:
 - Odhady (teorie odhadu),
 - testování hypotéz,
 - regresní a korelační analýza, ...
- Související otázky:
 - správné vytvoření výběrové souboru:
 - rozsah výběru,
 - spolehlivost odhadu,
 - reprezentativnost výběrového souboru.
 - nalezení teoretických modelů rozdělení odhadovaných charakteristik.

 $n \ll N$


Výběrová šetření

- ▶ Výběrové šetření (dle rozsahu):
 - ▶ $n < 30(50)$ – malý výběrový soubor
 - ▶ $n \geq 30(50)$ – velký výběrový soubor
- ▶ Výběrové šetření (dle kritérií):
 - ▶ náhodný (zařazení jednotky je náhodné),
 - ▶ záměrný (zařazení jednotky na základě zvoleného hlediska),
 - ▶ na základě dostupnosti (nutné zpětné popsání základního souboru).
- ▶ Výběrové šetření (dle opakovatelnosti výběru statistické jednotky):
 - ▶ výběr s vracením,
 - ▶ výběr bez vracení (prostý náhodný výběr).
- ▶ Výběrové šetření (dle uspořádání základního souboru):
 - ▶ jednoduchý výběr,
 - ▶ složený (sofistikovaný) výběr (stratifikovaný, záměrný, ...)
 - možnost subjektivního zatížení,
 - nehomogenita výběru – výběr z různých populací,
 - nebezpečí nereprezentativnosti.



Vybrané metody výběru

Náhodný výběr Všechny statistické jednotky mají stejnou šanci být ve výběru, každá „výběrová“ n -tice je stejně pravděpodobná – **základní podmínka statistické indukce**.

Vícestupňový výběr Výběr z výběru ;-).

Metoda základního masivu Výběr se zaměřuje na „hlavní“ statistické jednotky.

Záměrný (úsudkový) výběr Výběr je vytvořen záměrně většinou odborníky.

Stratifikovaný (oblastní) výběr Základní soubor je rozdělen na disjunktní podsoubory a z nich je proveden výběr.

Systematický (mechanický) výběr Základní soubor je seřazen a je vybírána každá k -tá statistická jednotka ($k = \frac{N}{n}$). První statistická jednotka je vybrána náhodně z prvních k jednotek.

Anketa Výběr je osloven dotazníkem – návratnost dotazníků.

