

# Základy reprezentativity a výběrové postupy v kvantitativním výzkumu

Mgr. Jana Obrovská, 20.9. 2016

# Kvantitativní a kvalitativní výzkum: podobnosti a rozdíly (opakování)

- deduktivní versus induktivní strategie výzkumu
- lineární versus cirkulární plán výzkumu
- náhodný versus záměrný výběr případů do zkoumaného vzorku
- zobecnitelnost zjištění versus platnost zjištění zejména pro daný případ/kontext (mnoho informací o málo jedincích versus málo informací o mnoha jedincích)
- snaha eliminovat efekt tazatele versus subjektivita a biografie výzkumníka jako součást výzkumného procesu
- Kdy se více hodí následovat kvantitativní výzkumnou strategii a kdy kvalitativní? U kvantitativních výzkumů zpravidla pokud se zabýváme relativně velkými populacemi a zjištění chceme generalizovat, zatímco kvalitativní strategii zpravidla u menších populací, kdy usilujeme o detailní zjištění, která nechceme zobecňovat na celou populaci.

# Plán kvantitativního výzkumu

- ▶ 1) Formulace teoretického nebo praktického sociálního problému
  - ▶ 2) Formulace teoretické hypotézy
  - ▶ 3) Formulace souboru pracovních hypotéz
  - ▶ 4) Rozhodnutí o populaci a vzorku
  - ▶ 5) **Pilotní studie**
  - ▶ 6) Rozhodnutí o technice sběru informací
  - ▶ 7) Konstrukce nástrojů pro tento sběr
  - ▶ 8) **Předvýzkum**
  - ▶ 9) Sběr dat
  - ▶ 10) Analýza dat
  - ▶ 11) Interpretace, závěry, teoretická zobecnění
- (podle Disman, 2011)*

# Kvanti, nebo kvali?

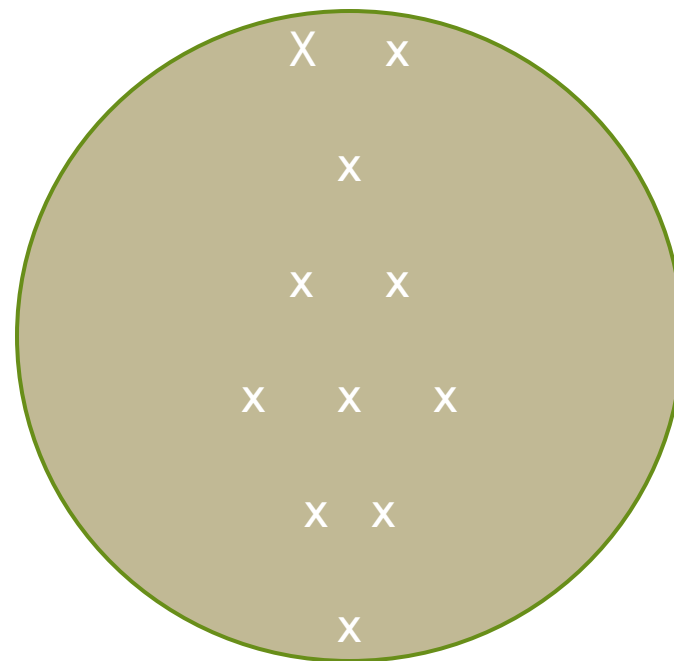
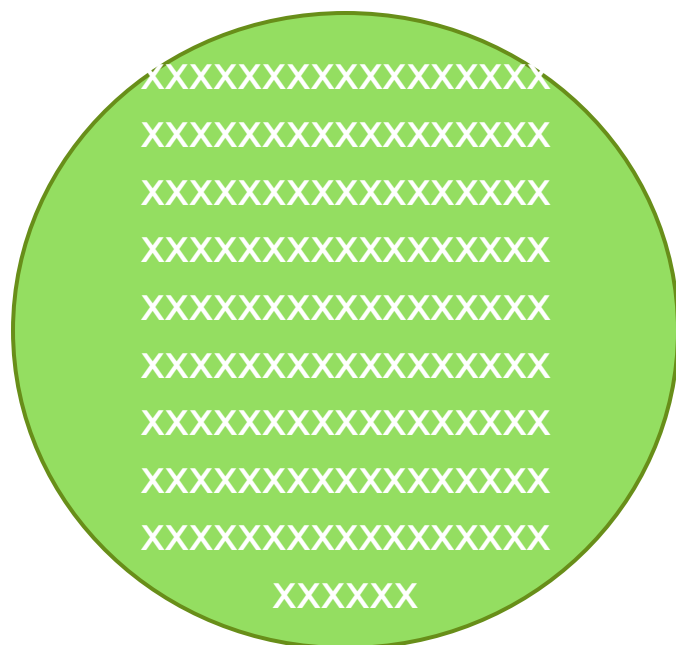
- ▶ Musíme dobře zvážit, jestli zvolit kvalitativní nebo kvantitativní design výzkumu. Volba se odvíjí od našeho tématu a zejména od výzkumných otázek, které si na začátku výzkumu položíme.
- ▶ Pokud si zvolíme kvantitativní paradigma, musíme dobře zvážit, jaká je velikost cílové skupiny, kterou chceme zkoumat, případně jestli jsme vůbec schopni její velikost zjistit
- ▶ Např. zajímá nás populace dospělých jedinců s Williamsovým syndromem. Má smysl zde usilovat o realizaci kvantitativního výzkumu, pokud jejich celkový počet bude 250?
- ▶ A co když nás bude zajímat populace jedinců s touto diagnózou, kteří jsou starší 50 let, kterých bude v celé populaci ČR jen 50?

# Důležité pojmy - základní populace a výběrový soubor

- ▶ **Cílová populace (základní soubor)** je soubor jednotek, o kterém předpokládáme, že jsou pro něj naše závěry platné. Je to soubor osob, pro který chceme vyslovit závěry výzkumu.
- ▶ **Vzorek (výběrový soubor)** - skupina jednotek, které skutečně pozorujeme.
- ▶ Zkoumat celou populaci většinou nemůžeme - ta je studována pouze v případě cenzu jednou za deset let, kdy skutečně vyplňují dotazníky všechny osoby starší osmnácti let. Jde o tzv. **vyčerpávající výběr**.
- ▶ Redukovaná analýza reality (tj. z populace na vzorek) vede k tvrzením **pravděpodobnostního charakteru**

# Populace x vzorek

## Základní soubor x výběrový soubor



# Reprezentativita

- ▶ Možnost zobecnění výsledků - určuje, zda to, co bylo vyzkoumáno, je možné vztáhnout i na další objekty, které přímo nebyly předmětem zkoumání.
- ▶ **Generalizovat** informace lze pouze v případě, že výběrový soubor je zmenšeninou souboru základního, tzn. že oba soubory se neliší v rozložení žádné z myslitelných vlastností, jenom svojí velikostí. Proměnné, které nás ve výzkumu zajímají, musí být ve vzorku stejně distribuovány jako proměnné v naší populaci, abychom mohli zjištění na populaci aplikovat.
- ▶ Reprezentativnost znamená, že struktura vzorku musí imitovat složení populace tak dobře, jak je to jen možné.
- ▶ S rostoucí velikostí vzorku se rozdíl mezi strukturou populace a vzorku zmenšuje.
- ▶ **Reprezentativita** není určována jen počtem zkoumaných jednotek a návratností, ale i mechanismem jejich výběru ze základního souboru. **Zobecnění** jsou potom možná jen tehdy, pokud dodržíme principy **pravděpodobnostního** neboli **náhodného výběru**. Skrze postupy **náhodného výběru** se snažíme zajistit, aby vzorek byl co nejvíc podobný s výchozí populací.

## Orientační návod pro vztah mezi velikostí základního a výběrového souboru (Gavora, 2010)

Základní soubor	Výběrový soubor
100	80
200	135
300	169
400	196
500	217
1000	278
1500	357
10000	370



- ▶ **Zobecnění jsou možná jen na naši cílovou populaci a na žádnou jinou, na což se často zapomíná.**
- ▶ Např. z reprezentativního vzorku pacientů a pacientek sexuologických poraden a ordinací se zobecňuje na sexuální chování celé české populace.
- ▶ Např. z reprezentativního vzorku účastníků demonstrace organizované Blokem proti islámu se usuzuje na xenofobní nálady celé české populace.

Pohádka pro odrostlejší děti 8.

## O hodně velkém vzorku, aneb Jak to nevyšlo

Byl jednou v Americe velice rozšířený týdeník, který se jmenoval Literary Digest. Byl u svých čtenářů hodně oblíben. Byl proslulý také tím, že spolehlivě předpovídal výsledky presidentských voleb. Jeho předpovědi byly založeny na obrovském vzorku dvou milionů voličů. (Dnes jsou podobné předpovědi založeny na vzorku tisíckrát menším.) Vzorek byl zkonstruován z mnoha zdrojů. Literary Digest si opatřil adresy voličů z celých USA. Používal pro to zdroje jako telefonní seznamy, městské adresáře, adresy držitelů řidičských průkazů, členské seznamy organizací, seznamy předplatitelů novin a časopisů atd.

Předpovědi byly přesné a úspěšné ve volbách 1920, 1924, 1928, 1932, a pak přišly volby v roce 1936. Literary Digest předpověděl, že presidentský kandidát Landon porazí Roosevelta rozdílem 14%. Přišel volební den a s ním i konec slávy Literary Digestu: Franklin Delano Roosevelt zvítězil drtivou většinou.

# Výběry zajišťující reprezentativitu

- ▶ **Náhodný výběr** neboli **pravděpodobnostní výběr** jako takový výběr, v němž má každý člen populace stejnou pravděpodobnost, že se dostane do vzorku.
- ▶ Náhodný vzorek reprezentuje všechny známé i neznámé vlastnosti populace.
- ▶ Pokud dodržujeme principy náhodného výběru, tak jsme schopni odhadnout, jak se vzorek liší od populace
- ▶ Závislost na tzv. opoře výběru; seznam jednotek základní populace, z něhož vzorek vybíráme
- ▶ V praxi tedy usilujeme o co možná největší vzorek, při zvladatelných finančních nákladech, aniž bychom ohrozili principy náhodného výběru

# Typy náhodného / pravděpodobnostního výběru

- ▶ **Prostý náhodný výběr** - „metoda klobouku“ - očíslováme všechny členy populace, z níž vytváříme vzorek, a použijeme generátor náhodných čísel (existují programy, kam můžeme přesně nastavit, v jakém rozmezí se náhodná čísla mají generovat - např. v rozsahu 1-300, pokud má naše populace 300 členů). Vybereme tolik čísel a podle nich i osob ze seznamu, kolik potřebujeme mít ve vzorku.
- ▶ **Systematický výběr** - ze seznamu je vybrán každý  $n$ -tý člen. Číslo  $n$  získáme tak, že vydělíme velikost populace velikostí vzorku. Aby byl postup skutečně náhodný, tak první člen, od něhož budeme vybírat každý  $n$ -tý člen, musí být vybrán náhodně. Seznamy nesmí být řazeny podle nějakého systematického kritéria.
- ▶ **Náhodný stratifikovaný výběr** - populace je rozdělena do skupin podle homogenního kritéria a v rámci těchto skupin je uplatněn prostý náhodný výběr nebo systematický výběr
- ▶ **Vícestupňový náhodný výběr** - jde o výběr prováděný ve dvou krocích. Nejprve jsou náhodně vybrána nějaká přirozeně existující uskupení a potom z těchto uskupení jsou náhodně vybírány jednotlivé členy. V tomto postupu je podstatné dodržet náhodnost i např. ve fázi výběru osoby v domácnosti (to je zpravidla úkol tazatele), ten by si mohl chtít usnadnit práci tím, že bude vybírat osobu, která je zrovna v dosahu.

# Problémy s pravděpodobnostními výběry

► Velmi často ale nejsme schopni postupy náhodného výběru použít kvůli:

- Nákladnosti
- Protože nemáme seznamy všech členů populace (tzv. opory výběru)
- Protože máme nízkou návratnost dotazníků (tzv. redukce negativním samovýběrem, s níž se velmi obtížně bojuje).

..... Proto volíme nepravděpodobnostní typy výběrů.

# Výběry nezajišťující reprezentativitu

- ▶ **Kvótní výběr** - imituje ve struktuře vzorku známé vlastnosti populace. Kvótní výběr se většinou omezuje jen na několik málo proměnných, zpravidla pohlaví, věk, povolání, dosažené vzdělání apod.
  - ▶ **Účelový výběr/Záměrný** - je založen na úsudku výzkumníka o tom, co by mělo být pozorováno a co je možné pozorovat
  - ▶ **Samovýběr/Anketa** - odpovídá ten, kdo má zájem. Je problematická, protože u ní nejsme schopni určit, pro jakou populaci jsou naše závěry platné (problém zkreslení samovýběrem)
  - ▶ **Sněhová koule** - je založena na postupu, při kterém nás původní informant vede k dalším informátorům. Tato technika je explicitně účelová a používá se zejména v kvalitativním výzkumu, případně v případě, že zkoumáme populace, které existovaly přechodně (např. účastníci nějaké události apod.).
- U nepravděpodobnostních výběrů je velké omezení pro generalizaci našich závěrů. Nesmíme předstírat, že jsou naše závěry platné pro „každého ve vesmíru.“

# O jaké typy výběru se jedná?

- ▶ Postavím se na Náměstí Svobody a oslovím každého třetího náhodného kolemjdoucího do svého vzorku.
- ▶ Nejprve náhodně vyberu okresy, v jejich rámci obce, v jejich rámci např. volební obvody a až potom z nich náhodně vybírám jedince.
- ▶ Zajímá mě populace bezdomovců, nechci se ale omezit jen na ty, kteří obývají azylové domy a jiné typy zařízení. Proto svůj vzorek konstruuji tak, že oslovím skupinu bezdomovců, kteří postávají na hlavním nádraží a ti mi doporučí oslovit další bezdomovecké skupiny podle míst jejich nejčastějšího výskytu.

# Pilotní studie

- ▶ cílem je zjistit, zda je náš výzkum v dané populaci vůbec realizovatelný
- ▶ ověřuje, zda informace, kterou požadujeme, v naší populaci vůbec existuje a zda je dosažitelná
- ▶ je prováděna např. na malé skupině vybrané z populace, kterou hodláme studovat
- ▶ často kvalitativní techniky
- ▶ důležitá hlavně tehdy, pokud nemáme opravdu hlubokou znalost o zkoumané populaci
- ▶ Opomenutí pilotní studie hrozí výrazně ohroženou validitou našich zjištění (v duchu rčení, když dva říkají to samé, nemusí to pro ně to stejné znamenat).



# Předvýzkum

- ▶ účelem předvýzkumu je odzkoušení nástrojů (např. dotazníku), které jsme pro náš výzkum zkonstruovali
- ▶ počet zapojených osob je zde zpravidla vyšší než u pilotní studie
- ▶ nejčastěji testujeme srozumitelnost otázek, popř. zkoušíme i analytický postup
- ▶ předvýzkum bychom neměli v žádném případě v našem výzkumném postupu opominout