Výběr výzkumného vzorku

Základy pedagogické metodologie, Kateřina Vlčková, Centrum pedagogického výzkumu PdF MU

Subjekty výzkumu

- předmětem výzkumu jsou jevy, osoby, věci, procesy, archivní záznamy, knihy atd.
- lidé = subjekty výzkumu

Základní soubor (populace)

- základní soubor (populace) = všichni lidé, které výzkum zkoumá
- vymezení základního souboru musí být přesné, aby bylo jasné koho se zjištěné výsledky týkají

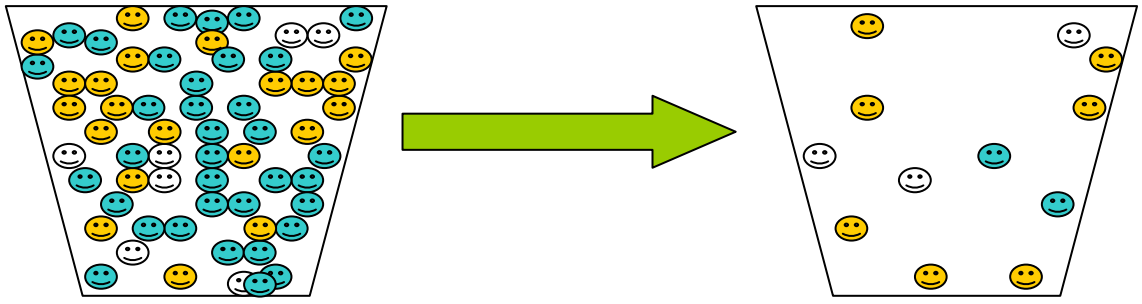
Př. základní soubor = všechny školy a žáci v ČR – to je velmi velký soubor, nelze z časových, finančních aj. důvodů všechny zkoumat

→ Vše, co zjistíme platí pouze pro tento základní soubor – pouze pro ČR, nikoli pro všechny žáky a školy na světě (a i pro ČR pouze tehdy, když byl zvolen adekvátní způsob výběru vzorku)

Př. základní soubor = všechny 9. třídy ZŠ v Jihomoravském kraji

→ Vše, co zjistíme platí pouze pro žáky 9. tříd ZŠ v JHM kraji, nelze závěry výzkumu formulovat obecně (žáci v ČR dělají to a to), je třeba je formulovat přesně (u zkoumaných žáků, žáků v JHM kraji se zjistilo...)

Základní soubor bývá obvykle velký (celá ČR, celý region atd.) a nelze jej z finančních časových aj. důvodů zkoumat celý. Je třeba vybrat výzkumný vzorek.

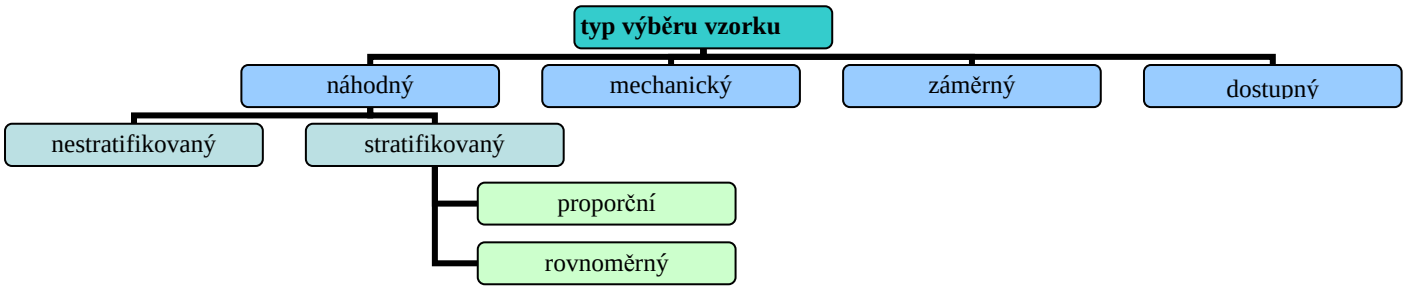


základní soubor — výzkumný vzorek

Výzkumný vzorek (výběrový soubor)

- má dobře reprezentovat základní soubor – má všechny náležitosti základního souboru, jen je menší, je zmenšeninou základního souboru
- Př. pokud jsou základním souborem všechny 3. ročníky čtyřletých gymnázií v kraji Vysočina, je třeba zjistit, kolik 3. ročníků v daném typu studijního programu v daném kraji je. Statistické údaje lze zjišťovat např. u českého statistického úřadu (*Statistická ročenka ČR* aj.) a ve *Statistické ročence školství* (www.uiv.cz)
- rozhodující pro možnost zobecňovaní výsledků získaných na výběrovém vzorku na základní soubor je způsob výběru vzorku

Základní typy výběru výzkumného vzorku



Náhodný výběr

- každý prvek základního souboru má stejnou šanci dostat se do výběrového souboru (**reprezentativní výběr**)
- jediný faktor který hraje roli v tomto výběru je náhodnost
- náhodný výběr ve smyslu statistické pravděpodobnosti
- z hlediska teorie pravděpodobnosti se jedná o nejlepší výběr – reprezentativní výběr
- provádí se **losováním** (žáky očíslujeme a losujeme) nebo pomocí **tabulky náhodných čísel** sestavené počítačem
- ukázka k čemu vede chybný výběr viz Gavora Úvod do ped. výzkumu, 2000, s. 62.

Náhodný stratifikovaný výběr

- znaky pro stratifikaci volíme dle potřeb výzkumu (př. pohlaví, věk, vzdělání, socioekonomický statut, prospěch,..)

proporční výběr

- rozložíme základní vzorek na podskupiny a poté dle proporcionálního zastoupení vybereme vzorek náhodně
- př. rozlišíme malé a velké školy, školy městské a venkovské, školy dle vzdělávacího programu
- př. využívá se při výzkumu volebních preferencí (znakem pro stratifikaci je např. pohlaví, věk, vzdělaností úroveň, sídlo, forma výdělečné činnosti atd.)

rovnoměrný výběr

- rozložíme základní vzorek na podskupiny a poté vybereme ve stejném počtu z každé skupiny subjekty do vzorku

Mechanický výběr

- vybíráme každou n-tou osobu
- př. ze seznamu žáků, z telefonního seznamu
- nebezpečí: seřazení prvků nemusí být náhodné, může se projevit nějaká pravidelnost ve výběru (periodicita), a tím dochází ke zkreslení výběru a pak i výsledků výzkumu

Záměrný výběr

- určíme relevantní znaky základního souboru, které jsou podstatné pro náš výzkum
- podstatné znaky stanovíme na základě teorie, výzkumu, zkušeností apod. – jedná se o kvalifikovaný postup (skupina expertů výběr navrhuje)
- vyhledávají se subjekty, které daný znak mají
- nedostatky: zvětšováním souboru se nezlepšují jeho vlastnosti (ověření vlastností výběru lze zkontrolovat na dalším výběrovém souboru vybraném ale náhodným výběrem)

Dostupný výběr

- výzkumník bere „co je po ruce“ (školy, na které je „puštěn“; školu v sousedství; své bývalé učitele; žáky, kteří jsou ochotni odpovědět apod.)
- nedostatky: silné zkreslení výběru od základního souboru, závěry platí pouze pro ty subjekty, které se výzkumu skutečně účastnili, nelze nic zevšeobecňovat

Velikost výzkumného vzorku

- neboli rozsah výzkumného vzorku – **n** (např. n = 220)
- jak velký potřebujeme vzorek závisí na požadované spolehlivosti výsledků, požadované přesnosti a variabilitě základního souboru
- můžeme rozhodnout, jakou minimální kvalitu výsledků v určitém statistickém testu požadujeme (např. pro t-test: p = 0,05 .... 5% chyby akceptujeme; síla vztahu mezi zkoumanými proměnnými alespoň 0,8 ....silnější vztah) a spočítat si minimální velikost vzorku (více k tomu v literatuře k tématu Approach of Statistical Power), síla vztahů klesá mnohem rychleji, když k malému vzorku (např. n < 200) přidáváme další lidi než když přidáváme další lidi k velkému vzorku
- pokud je základní soubor nehomogenní (skládá se z více podskupin), je třeba vzorek větší (př. dotazník – na každou skupinu asi 50-60 lidí)
- čím více proměnných chceme zkoumat, tím větší by měl být rozsah vzorku (př. chceme zkoumat pohlaví, ročníky školy atd.)
- pokud ve výzkumu působí neznámé nebo nekontrolované (nepozorované) proměnné (např. v experimentu), je třeba mít větší vzorek, aby se výskyt nekontrolované proměnné ve vzorku znáhodnil
- čím závažnější jsou důsledky výzkumu (např. závažné politické rozhodnutí např. o zrušení nekvalitních škol), tím důležitější je správná a dostatečná velikost vzorku
- velikost vzorku závisí také na časových a finančních možnostech výzkumníka – dotazníků může být více než časově náročných videonahrávek vyučovacích hodin

Literatura

Sestaveno dle: Gavora, P. *Úvod do pedagogického výzkumu*, Brno : Paido, 2000.