

PROGRAM VÝUKY DSP 8.4.2021

- 1. Informace k vyvážšeným podkladů (docx)**
- 2. Informace k APA (odkazy, ukázka, docx)**
- 3. ResearchGate (ukázka)**
<https://www.researchgate.net/>
- 4. Výzkumný soubor (22, ppt)**
- 5. Omyly a myty korelačního počtu (38, ppt)**
- 6. Diagnostika: Test+TB+Normy (40, ppt)**

**6. Odborný seminář
METODOLOGIE A STATISTIKA
V KINANTROPOLOGICKÉM VÝZKUMU**



FSpS MU Brno 28. 6. 2016

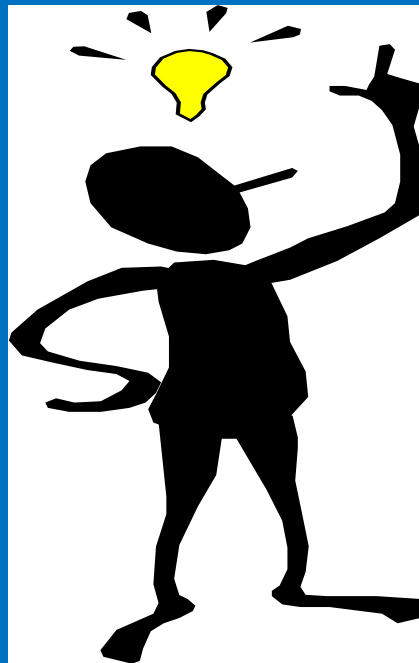
VÝZKUMNÝ SOUBOR: POPULACE & NÁHODNÝ VÝBĚR

Doc. RNDr. Jiří Zháněl, Dr.



Tento výstup vznikl na Masarykově univerzitě v rámci projektu „Posilování kompetencí akademických pracovníků v oblasti metodologie a statistické analýzy dat“ č. MUNI/51/02/2016 podpořeného z prostředků účelové podpory na Interní rozvojový projekt FSpS v roce 2016.

POPULACE nebo NÁHODNÝ VÝBĚR?



ZÁKLADNÍ resp. VÝBĚROVÝ SOUBOR

ZÁKLADNÍ SOUBOR (ZS)

(generální s., Grundgesamtheit, Population, N)

***ZS = množina všech jedinců (jednotek),
u/na kterých bychom teoreticky měli
šetření (výzkum) provádět.***

Jedinci s požadovanou vlastností/vlastnostmi.

Realizace obvykle nemožná, časově, logisticky a ekonomicky náročné.

ZÁKLADNÍ SOUBOR (ZS)

ZS má být vedle popisu jedinců a sledovaných znaků (vlastností) definován časovými a prostorovými souvislostmi.

Např.

- ✓ všichni studenti TV a sport v ČR 2016,
- ✓ všechny 5 leté děti s bydlištěm v CZ 1.1.2016,
- ✓ všichni členové fotbalové reprezentace 2016,

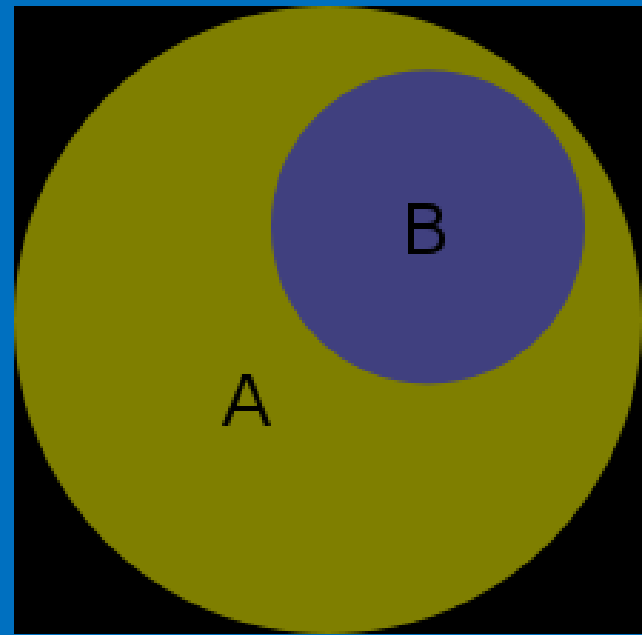
Realizace výzkumu ZS obvykle nemožná, časově a ekonomicky náročná.

VÝBĚROVÝ SOUBOR (VS)

(výběr, vzorek, Stichprobe, sample, n)

VS = náhodná podmnožina prvků základního souboru, která jej *REPREZENTUJE*.

VS získáváme metodami náhodného výběru, na základě výsledků VS činíme závěry platné pro základní soubor. Nejedná-li se o náhodný výběr, nelze závěry zobecňovat, platné jen pro VS.



VÝBĚROVÝ SOUBOR

- ✓ Může jít o **osoby**, ale také o **objekty** (např. školy, třídy, kluby, v meta-analýze jde o provedené studie).
- ✓ Zdůvodňujeme **způsob výběru** z populace a vhodný počet jedinců.
- ✓ Popisujeme **způsob**, jak získáme přístup k podmnožině populace.

Výběrový soubor získáváme některou z metod *náhodného výběru*, kdy **každý prvek základního souboru má stejnou možnost být vybrán.**

O vybrání/nevybrání rozhoduje náhoda.



Metody náhodného výběru

I. Losování

✓ losování statistických jednotek s jejich vrácením do osudí (u malých souborů),

✓ losování statistických jednotek bez vrácení do osudí (u velkých souborů),

II. **tabulka náhodných čísel**, nebo generátor náhodných čísel (počítačový program generující náhodná čísla).

Postup pomocí tabulky náhodných čísel

Např. ze základního souboru $N=540$ chceme vybrat $n=12$

1. V tabulce zvolíme libovolné číslo, od něj čteme uvedená čísla s potřebným počtem míst (např. $N=540 \Rightarrow$ trojmístná čísla)

2. Do výběru zahrnujeme ty jednotky základního souboru, jejichž přiřazená čísla jsou ta, která čteme v tabulce.

3. Čísla vyšší než rozsah základního souboru vynecháme.

4. Pokračujeme tak dlouho, než dosáhneme požadovaného rozsahu výběrového souboru.

N=540	85306	37114	22718	50584	92291	56575	24075	43889
	32066	43098	75738	94910	15403	89151	73322	18370
	63314	87302	49472	24885	79506	60638	07132	00908
	40287	52435	23926	92544	54099	31497	06853	22864
	30925	46148	20138	33874	56715	38424	38273	11361
n=12	27146	37012	43361	03173	97911	71313	44256	66609
	01674	47274	56350	37512	14883	99673	62298	33948
	76730	25043	16686	54737	57431	01786	20803	69465
	93941	84434	22384	13240	93617	51549	28532	57150
	<u>90475</u>	<u>10341</u>	<u>39703</u>	<u>83224</u>	<u>37858</u>	<u>61657</u>	<u>04184</u>	15597
86115	17196	24569	26820	66299	39960	02489	53079	
51156	74037	12501	94162	42006	16135	82797	31296	
59886	03051	78702	13402	74318	10870	72107	11550	
13960	95736	43637	60399	19080	60261	11207	73065	
39954	86726	91039	13884	25376	36880	02564	96978	
47906	99501	27753	69946	66875	25601	30038	78786	
66444	15979	83469	76952	50065	72802	70630	87336	
40177	01081	57788	08612	39886	42234	04905	83274	
46747	30655	41878	93610	51745	41771	61398	98154	
60888	18689	45966	25837	70906	60733	11765	09293	

III. Skupinový výběr

... užívá se, je-li základní soubor uspořádán do skupin (třídy, kluby), z nichž vybíráme – nutný je dostatečný počet skupin.

IV. Stratifikovaný výběr

... vychází z rozdělení základního souboru na skupiny („stráty“) z nichž se pak dělá náhodný výběr. Např. „vysokoškoláci“ (studující techniky, univerzity, uměleckých vysokých škol, atd.).

V. Záměrný výběr

O zařazení do výběrového souboru nerozhoduje náhoda, nýbrž výzkumník vybírá jedince, které považuje za typické (subjektivní výběr) podle určitého záměru (kritéria).

Např. IQ přes 130b, tenisté ATP 100, hráči fotbalového oddílu U14, atd.

Další podrobnosti v odborné literatuře.

KONCEPT REPRESENTATIVNOSTI

Výběrový soubor (VS)

- ✓ reprezentuje **populaci** (základní soubor) tak dobře, jak dobře je vybrán,
- ✓ je reprezentativní, když výsledky mají prokazatelně **platnost i pro populaci**,
- ✓ **struktura (složení) VS** musí být silně podobná základnímu souboru (viz př.),
- ✓ umožňuje rychlou realizaci a vyhodnocení výzkumu.

KONCEPT REPRESENTATIVNOSTI

Odhad volebních výsledků USA (1936)

= legendární příklad nepodařeného NV.

Dotazníky na adresy 10 mil. Američanů z telefonních seznamů, členských karet klubů a spolků, vráceno 2,4 mil. dotazníků (24%):

*Odhad 43 % Roosevelt, skutečnost 62%. **Proč ???***

Zdroje chyb:

1) vyšší vrstva = telefon a karty

2) nižší vrstva = souhlas s politikou Roosvelta

Jak dosáhnout reprezentativity?

Provést náhodný ze všech jednotek základního souboru za situace, kdy všechny jednotky mají stejnou pravděpodobnost vybrání [Knoke, Bohrnstedt, Mee 2002]

Rozsah výběrového souboru?

Blaikie (2003): 300 je adekvátní, 500 je lepší a 1000 by bylo ještě lepší.

Bortz & Schuster (2010): Tolik, kolik je možno!!!

Soukup & Rabušic (2007): **základní soubor alespoň 100krát větší než výběrový soubor.**

Rozsah výběrového souboru?

1. Teorie **statistické indukce** (zobecňování z VS na ZS) byla vyvinuta pro případy **velkých NV z velkých ZS**, proto není možné její postupy v jiných situacích aplikovat.
2. **Velikost VS** závisí na požadované přesnosti odhadu, ale i na finančních a časových možnostech.
3. Se vzrůstajícím **rozsahem VS** stoupá **přesnost odhadu** parametrů ZS, stoupají ale také náklady.

Rozsah výběrového souboru?

Co znamená požadavek velkého náhodného výběru?

- (1) aby měl NV **dostatečný počet** jednotek,
- (2) aby bylo provedeno **vybírání ze ZS náhodně**,
- (3) aby šlo o náhodná výběr.

Pouze náhodné výběry, kde je vybíráno alespoň v řádu desítek jednotek cca **od 30–50**, lze označit za „velké“ (Soukup & Rabušic, 2007).

A co je to velký základní soubor? ?

Je nutné zajistit, aby **základní soubor** byl alespoň **100krát větší** než zamýšlený **výběrový soubor**.

Není-li toto splněno, lze samozřejmě statisticky testovat, ale **s jinými** než běžně dostupnými **vzorci**. V takovýchto situacích již zpravidla nelze užívat tolik oblíbené statistické balíky.

(Soukup & Rabušic, 2007).

Jiné postupy: Power analysis. Effect size, atd.

Rozsah výběrového souboru?

Př. Soubor M a Ž, jsou rozdíly ve výsledcích testu síly významné? Jaký rozsah VS?

Určení s ohledem na **sílu testu (ST)** a **velikost efektu (VE)** (Cohen, 1988).

Volba výzkumníka $ST=0,8$ (V) $VE=0,8$ (V)

Tabulka 1. Doporučený rozsah pro nezávislé výběrové soubory ($p=0,05$, oboustranný t-test).

ST/ VE	0,2	0,5	0,8
0,7	620	102	42
0,8	788	128	52
0,9	1054	172	68

Rozsah výběrového souboru?

Další způsoby určení rozsahu VS:

přípustná chyba odhadu, konfidenční interval, ... K nastudování!



Děkuji za pozornost