

Uzorkovanje

Šesto predavanje

Populacija i uzorak

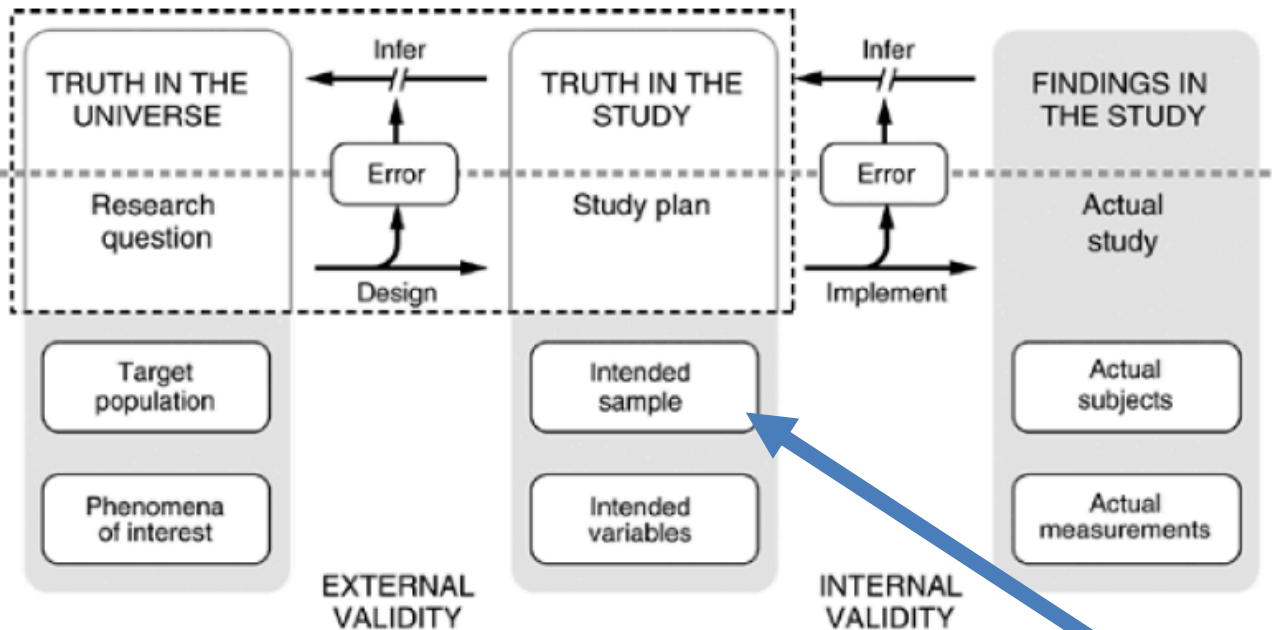


FIGURE 2.1. Choosing the research question and designing the study plan.

Šta je uzorkovanje?

- Ponekad jedinice namjerno i ciljano bираmo:
 - Primjer 1: Zašto su sudovi neefikasni?
 - Primjer 2: Zašto se mladi učlanjuju u političke partije?
 - Primjer 3: Koji su osnovni razlozi za mirno prihvatanje rezultata Referenduma 2006. godine u Crnoj Gori?
- **Uzorkovanje – metodološki specifičan način izbora finalnih jedinica koje reprezentuju čitavu populaciju.**
- Dobar uzorak mora da posjeduje sve ključne karakteristike koje posjeduje populacija.

Primjer

- U.S.– istraživanje o navikama građana po pitanju kulture – 50.000 ljudi popunilo anketu
 - Online anketa preko web sajta
- U.S. National Endowment for Arts – 12.000 ljudi popunilo anketu
 - Telefonska anketa
- Ista tema različiti rezultati

Uzorkovanje u kvantitativnim istraživanjima

- Prvi korak: definisati populaciju koja se uzorkuje
- Način uzorkovanja
 - Probability sampling
 - Svako ima jednaku šansu da se nađe u uzorku
 - Simple random sampling (lutrija sistem)
 - Stratifikovani random sampling (podjela na podgrupe, npr. teritorijalno i uzorkovanje na osnovu proporcija grupa)
 - Nonprobability sampling
 - Kvotni uzorak (odabir karakteristika)
- Response rate (50-80%)
 - Broj onih koji pristanu vrši distorziju

Tri važna elementa kod izbora finalnih jedinica

- Postojanje dobrog uzoračkog okvira (lista elemenata koji čine populaciju ili mehanizam da se do članova populacije dođe)
- Slučajni izbor elemenata sa liste
- Mehanizam da su bitne i relevantne grupe zastupljene u odabranom uzorku

Probability sampling

- Svaki element u populaciji ima poznatu vjerovatnoću da bude izabran
 - Ta vjerovatnoća ne mora nužno da bude ista

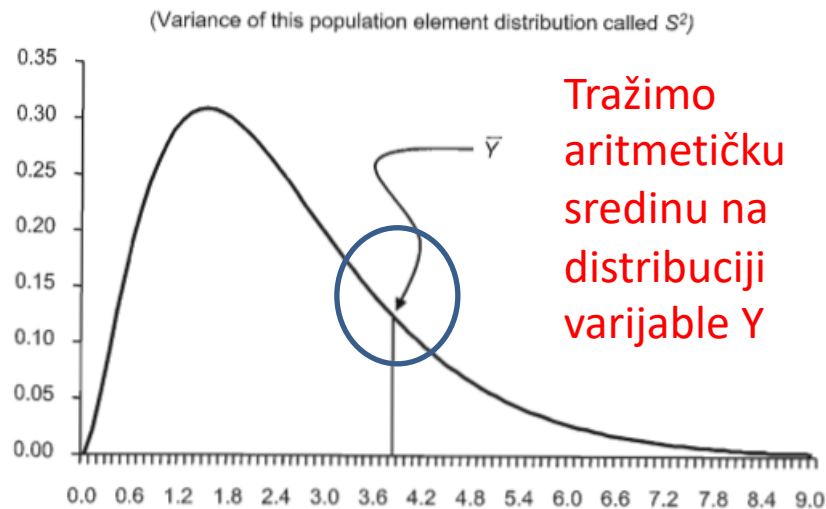
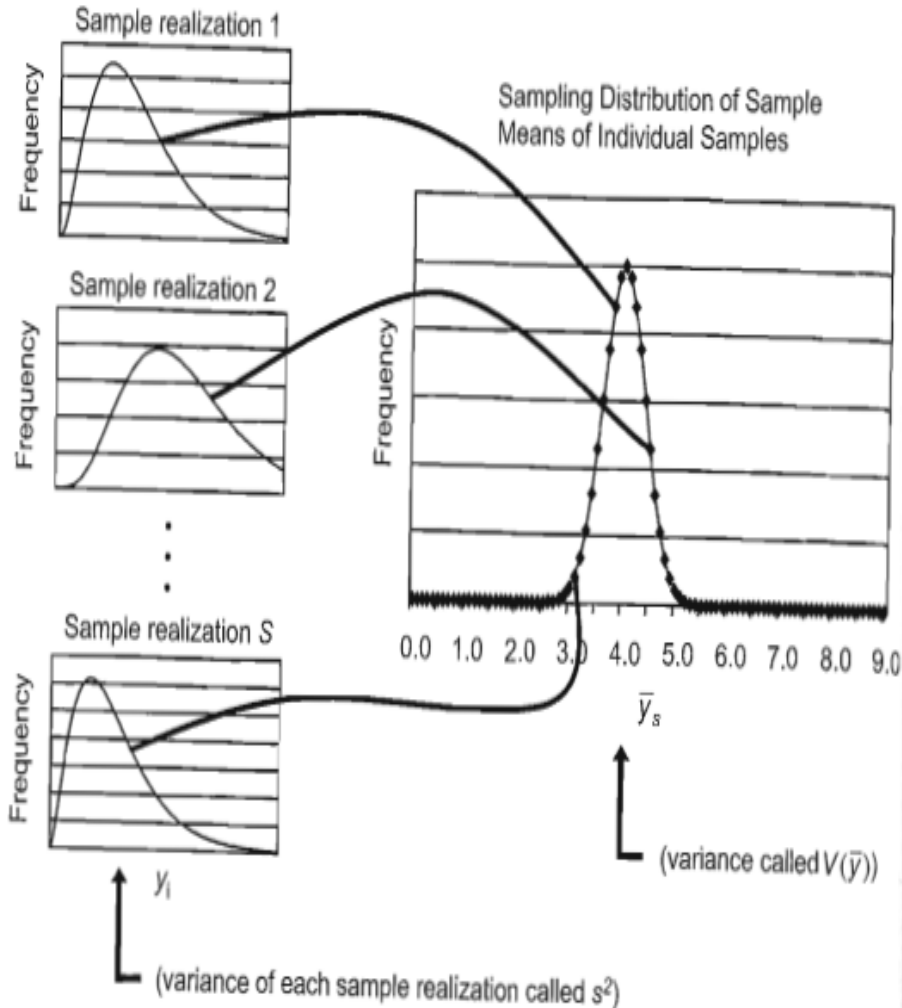


Figure 4.1 Unknown distribution for variable Y in frame population.

Probability sampling



- Aritmetička sredina **uzorka** sumira sve moguće vrijednosti uzorka koji se može izabrati (sample mean)
- Distribucija tih vrijednosti se naziva **sampling variance of the mean**
- Standard error of the mean (standardna greška mjerenja) je kvadratni korijen iz sampling variance of the mean
 - Označava granične vrijednosti u okviru datih intervala povjerenja
 - Primjer $\pm 3.1\%$ za interval povjerenja od 95%

Intervali povjerenja

- Statistička značajnost
 - Koja je vjerovatnoća da je razlika među grupama uzrokovana slučajem, a ne nekom konkretnom pojavom?
 - Konvencija 0.5%, 0.1%
 - 0.5% - interval povjerenja od 95% (čitaj: od 100 mogućih uzoraka 5 teorijski može biti van granice standardne greške mjerenja)
 - 0.1% interval povjerenja od 99% (čitaj: od 100 mogućih uzoraka 1 teorijski može biti van granice standardne greške mjerenja)
 - Prethodni primjer – standardna greška mjerenja $\pm 3.1\%$ za interval povjerenja od 95% i incidence od 50% znači – u 95 od 100 uzoraka koje izvučemo za datu populaciju srednja vrijednost koju izmjerimo na 50% tražimo će se naći između $+ 3.1\%$ i -3.1% , tj. između 46.9% i 53.1%

Utvrdjivanja greške mjerenja za određeni interval povjerenja

- Greška mjerenja zavisi od:
 - Veličine uzorka (što veći to bolji, ali rast kvaliteta „usporava“)
 - Procenta izmjerene vrijednosti
 - Najveća greška se može pojaviti kada je vrijednost izmjerena na 50%
 - Veličine populacije

Find Confidence Interval

Confidence Level: 95% 99%

Sample Size:

Population:

Percentage:

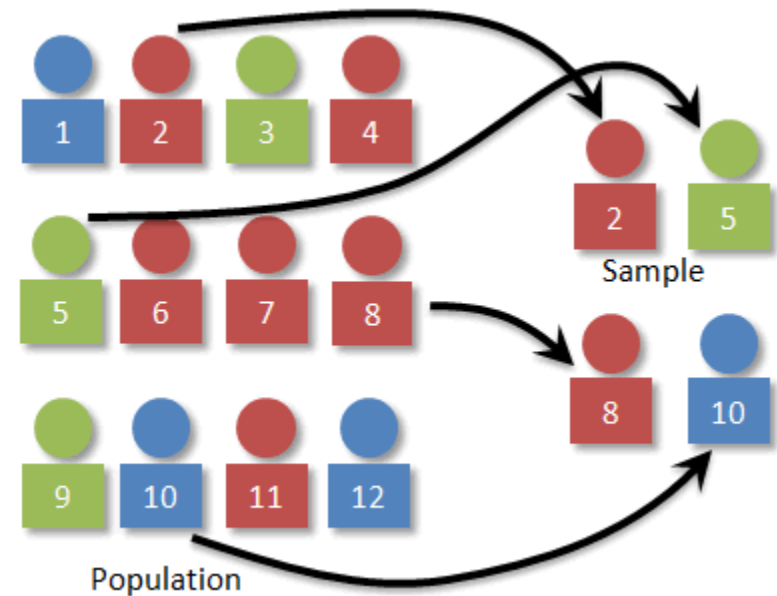
Confidence Interval:

Link za izračunavanje

- Napomena:
 - Confidence interval = greška mjerenja!
 - Confidence level = interval povjerenja!
- <https://www.surveysystem.com/sscalc.htm>

Probability sampling

- Simple random sampling
- Sistem lutrije
- Važno: svaki ispitanik mora da ima istu vjerovatnoću da bude izabran
- Kritika: Nepostojanje savršenih registara kao osnova



Probability sampling

- Stratified random sampling
 - Ako je „podgrupa“ relevantna za populaciju
 - Želimo da svaka bude jednako ili proporcionalno zastupljena u uzorku

Stratum	A	B	C
Veličina populacije	100	200	300
Odabrana „proporcija“	1/2	1/2	1/2
Veličina uzorka	50	100	150



Primjer kreiranja uzorka za CG:

- **Korak 1: Definisanje populacije**
 - Punoljetna populacija sa biračkim pravom
- **Korak 2: Definisanje relevantnih stratumata**
- **Korak 3: Definisanje relevantnog okvira za uzorkovanje**
 - Biračka mjesta

Region	Proporcija	Obrazloženje
Sjever	32.3%	Geografija, ekonomija, stepen razvijenosti, etnički sastav...
Centar	43.9%	
Jug	23.8%	

Grad	Biračko mjesto	Broj anketa
Podgorica	105-B DOM OMLADINE G.GORICA(R-I)	20
Podgorica	CENTAR ZA OBRAZ. I OSPOS. "1 JUN"(M-I)	20
Podgorica	Pošta br. 4 (A-K)	13
Podgorica	84-B O["BO@IDAR VUKOVI] PODGOR" (R-I)	20
Podgorica	106-A DOM OMLAD DONJA.GORICA (LJ-I)	23
Podgorica	O["Vladimir Nazor"	19
Podgorica	79-A JU O["PAVLE ROVINSKI" (M-I)	20
Podgorica	O["OKTOIH" (JU@NI ULAZ (M-I)	23
Podgorica	UO "POD ORAH" VL. ^ELJ] SAFET (M-I)	17
Podgorica	24-B UO "SMEDEREVO" (O-I)	17
Podgorica	FUDBALSKI KLUB "KOM" (A-Z)	17
Podgorica	116-B O["MAHMUT LEKI]" TUZI (N-I)	17
Rožaje	DESNA OBALA IBRA I/C	13
Rožaje	BUKOVICA	13
Rožaje	DESNA OBALA IBRA II/C	7
Rožaje	KALA^E	20
Rožaje	DESNA OBALA IBRA - II-	23
Tivat	Lepetane	7
Tivat	GRADIO[NICA II	13
Tivat	TIVAT -LAMELE,ST.BLOK	17
Tivat	SELJANOVO	20
Ulcinj	OBJEKAT-SMAILOVI]-	8
Herceg Novi	IGALO 3	13
Herceg Novi	PODI	20
Cetinje	ŠAH KLUB UL.NJEGOJEVA - CETINJE	13
Cetinje	KUD "NJEGOŠ" - CETINJE	13

Primjer kreiranja uzorka za CG:

- Korak 4: Izbor domaćinstva:
 - „random walk“ ili kreiranje registra
- Korak 5:
 - Sistem posljednjeg rođendana
 - Kišove tablice

Household	Eligible People							
	1	2	3	4	5	6	7	8+
1st	1	1	1	1	1	1	1	1
2nd	1	2	2	2	2	2	2	2
3rd	1	1	3	3	3	3	3	3
4th	1	2	1	4	4	4	4	4
5th	1	1	2	1	5	5	5	5
6th	1	2	3	2	1	6	6	6
7th	1	1	1	3	2	1	7	7
8th	1	2	2	4	3	2	1	8
9th	1	1	3	1	4	3	2	1
10th	1	2	1	2	5	4	3	2



"Do you really care about my opinion or am I just a random sample?"

Kako „čitamo“ podatke o uzorku u kvantitativnim istraživanjima

- Veličina uzorka
 - Nije intuitivno!
 - Logika: Veći uzorak manja standardna greška mjerenja
 - Primjer:
 - Šta znači to da je greška mjerenja $\pm 3.1\%$ za interval povjerenja od 95% i incidence od 50%?

Populacija Za nivo povjerenja od 95%	Veličina uzorka	Greška mjerenja (Standardna greška mjerenja)
498.305	50	± 13.8
	100	± 9.8
	500	± 4.3
	1000	± 3.1
	1500	± 2.5
	2000	± 2.1
	3000	± 1.7

Od čega zavisi dobar uzorak?

- Od broja elemenata koji se biraju
- Od vjerovatnoće da svi elementi mogu biti izabrani
- Od toga da li se pojedinci biraju nezavisno ili po grupama
- Od toga je li uzorak dizajniran na način da predstavi relevantne podgrupe u populaciji