

UZORCI I NJIHOVA PRIMJENA

Mihaljević, Mario

Undergraduate thesis / Završni rad

2022

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Josip Juraj Strossmayer University of Osijek, Faculty of Economics in Osijek / Sveučilište Josipa Jurja Strossmayera u Osijeku, Ekonomski fakultet u Osijeku**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:145:096926>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-09-01**



Repository / Repozitorij:

[EFOS REPOSITORY - Repository of the Faculty of Economics in Osijek](#)



Sveučilište Josipa Jurja Strossmayera u Osijeku

Ekonomski fakultet u Osijeku

Preddiplomski studij, smjer: Marketing

Mario Mihaljević

UZORCI I NJIHOVA PRIMJENA

Završni rad

Osijek, 2022.

Sveučilište Josipa Jurja Strossmayera u Osijeku

Ekonomski fakultet u Osijeku

Preddiplomski studij, smjer: Marketing

Mario Mihaljević

UZORCI I NJIHOVA PRIMJENA

Završni rad

Kolegij: Statistika

JMBAG: 0010192499

e-mail: mmihalj3@efos.hr

Mentor: prof. dr. sc. Josipa Mijoč

Osijek, 2022.

Josip Juraj Strossmayer University of Osijek

Faculty of Economics in Osijek

Undergraduate Study of Marketing

Mario Mihaljević

STATISTICAL SAMPLES AND THEIR APPLICATION

Final paper

Osijek, 2022

IZJAVA

O AKADEMSKOJ ČESTITOSTI, PRAVU PRIJENOSA INTELEKTUALNOG VLASNIŠTVA, SUGLASNOSTI ZA OBJAVU U INSTITUCIJSKIM REPOZITORIJIMA I ISTOVJETNOSTI DIGITALNE I TISKANE VERZIJE RADA

1. Kojom izjavljujem i svojim potpisom potvrđujem da je ZAVRŠNI RAD isključivo rezultat osobnoga rada koji se temelji na mojim istraživanjima i oslanja se na objavljenu literaturu. Potvrđujem poštivanje nepovredivosti autorstva te točno citiranje radova drugih autora i referiranje na njih.
2. Kojom izjavljujem da je Ekonomski fakultet u Osijeku, bez naknade u vremenski i teritorijalno neograničenom opsegu, nositelj svih prava intelektualnoga vlasništva u odnosu na navedeni rad pod licencom *Creative Commons Imenovanje – Nekomercijalno – Dijeli pod istim uvjetima 3.0 Hrvatska*.
3. Kojom izjavljujem da sam suglasan/suglasna da se trajno pohrani i objavi moj rad u institucijskom digitalnom repozitoriju Ekonomskoga fakulteta u Osijeku, repozitoriju Sveučilišta Josipa Jurja Strossmayera u Osijeku te javno dostupnom repozitoriju Nacionalne i sveučilišne knjižnice u Zagrebu (u skladu s odredbama Zakona o znanstvenoj djelatnosti i visokom obrazovanju, NN br. 123/03, 198/03, 105/04, 174/04, 02/07, 46/07, 45/09, 63/11, 94/13, 139/13, 101/14, 60/15).
4. Izjavljujem da sam autor/autorica predanog rada i da je sadržaj predane elektroničke datoteke u potpunosti istovjetan sa dovršenom tiskanom verzijom rada predanom u svrhu obrane istog.

Ime i prezime studenta/studentice: Mario Mihaljević

JMBAG: 0010192499

OIB: 44916785824

e-mail za kontakt: mmihalj3@efos.hr

Naziv studija: Preddiplomski studij, smjer: Marketing

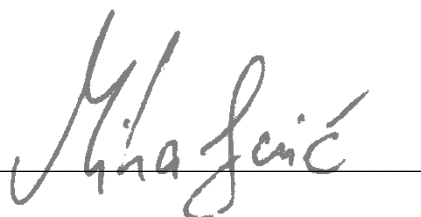
Naslov rada: Uzorci i njihova primjena

Mentor/mentorica rada: prof. dr. sc. Josipa Mijoč

Lektorica: Iva Barešić, mag. educ. philol. croat. et mag. educ. philo.

U Osijeku, 23. rujna 2022. godine

Potpis



Uzorci i njihova primjena

SAŽETAK

Uzorci su dijelovi populacije, a izabiru se u skladu s postavljenim ciljem istraživanja, istraživačkim pitanjima te su usko povezani s teorijskim i empirijskim odlikama istraživačkog problema. Vrste uzoraka impliciraju primjenu statističkih analiza, a vrste distribucija koje se očekuju na temelju provedenog izbora jedinica u uzorak utječu na primjenu statističkih procedura. Odabir vrste uzoraka u ekonomskim istraživanjima ovisi o odabiru populacije, broju varijabli, kao i o metodama koje se u istraživanju namjeravaju primijeniti. Opisivanje uzorka podrazumijeva uvid u demografsku strukturu uzorka te osnovne karakteristike uzorka povezane s problemom istraživanja. Mjerenje normalnosti distribucije jedan je od preduvjeta primjene procedura parametrijske statistike. Cilj je završnog rada opisati način na koji se biraju uzorci te kriterije kojima se istraživači koriste kako bi odabrali vrste uzoraka ovisno o tipu istraživanja koji provode.

Ključne riječi: vrste uzoraka, okvir uzorka, opis uzorka, statističke metode

Statistical samples and their application

ABSTRACT

Samples are parts of the population, and they are chosen in accordance with the set research goal, research questions, and are closely related to the theoretical and empirical characteristics of the research problem. The types of samples imply the application of statistical analyses, and the types of distributions expected based on the selection of units in the sample affect the application of statistical procedures. The choice of the type of samples in economic research depends on the selected population, the number of variables, and the methods that are intended to be applied in the research. Describing the sample implies an insight into the demographic structure of the sample and the basic characteristics of the sample related to the research problem. Measuring the normality of the distribution is one of the prerequisites for the application of parametric statistics procedures. The aim of the final paper is to describe the way in which samples are selected and the criteria that researchers use to select the types of samples, depending on the type of research they are conducting.

Key words: types of samples, framework of sample, sample description, statistical methods

SADRŽAJ

1. Uvod.....	1
2. Metodologija rada	3
3. Metoda uzoraka	4
3.1. Populacija	4
3.2. Podjele statističkih metoda	5
3.2.1. Procjena aritmetičke sredine.....	6
3.2.2. Procjena totala osnovnog skupa	10
3.2.3. Procjena proporcije osnovnog skupa	10
3.3. Proces izabiranja uzoraka	11
4. Karakteristike uzoraka	12
4.1. Veličina uzoraka	12
4.2. Reprezentativnost uzoraka.....	13
4.3. Vrste uzoraka.....	15
4.3.1. Metode uzorkovanja uz primjenu vjerojatnosti	15
4.3.2. Metode neprobabilističkog uzorkovanja	17
4.3.3. Primjeri uzorkovanja na studentima EFOS-a	20
5. Rasprava.....	22
6. Zaključak.....	24
Literatura	25
Popis priloga.....	27

1. Uvod

Završni rad donosi pregled teorijskih spoznaja kojima se opisuju uzorci i potreba za primjenom metode uzoraka u ekonomskim i poslovnim istraživanjima. Omair (2014:145) opisuje statistiku kao „znanost o prikupljanju, analizi, prezentiranju i tumačenju podataka. Razvoj statistike potaknut je s namjerom analize stanja i globalnog uspoređivanja nacionalnih podataka". Suvremena statistika omogućava pretvaranje velikih količina podataka dostupnih u mnogim primijenjenim područjima u korisne informacije, a ova potreba potaknula je teorijski i praktični razvoj statistike (Andrade, 2020).

Metode uzorkovanja dio su područja poznatog kao opisna statistika koja se ponajprije bavi metodama predstavljanja i tumačenja podataka pomoću grafikona, tablica i numeričkih sažetaka. Podaci dobiveni statističkom analizom zapisuju se brojevima ili riječima. Opisivanje podataka usmjereno je samo na opis jedne varijable, a primjena inferencijalne statistike, odnosno metoda procjena i testiranje hipoteza, postupci su koji se koriste za donošenje statističkih zaključaka, najčešće na temelju dviju ili više varijabli. Područja poput zdravstvene zaštite, biologije, kemije, fizike, obrazovanja, inženjeringa, poslovanja i ekonomije uvelike se koriste statističkim zaključcima (Andrade, 2013).

Uzorkovanje je dio kvantitativnog istraživačkog procesa (slika 1). „Istraživački proces temelji se na – prethodnim istraživanjima i izučavanju literature, odabiru metodologije istraživačkog rada, prikupljanju podataka i analizi podataka i tumačenju rezultata“ (Horvat i Mijoč, 2019:37). Uzorci se obično primjenjuju s namjerom kako bi se dobila slika o populaciji u cjelini, bez potrebe za promatranjem svakog člana te populacije.

Slika 1: Istraživački proces



Izvor: Horvat i Mijoč (2019: 36)

Kada se provodi istraživanje koje je usmjereno analizi stajališta skupine pojedinaca, rijetko je moguće prikupiti podatke od svake osobe u toj skupini. Umjesto toga pristupa se određivanju okvira uzoraka. Kako bi znanstvenici na temelju prikupljenih podataka u sklopu znanstvenih istraživanja donijeli valjane zaključke, moraju znati pravilno odabrati uzorak. Uzorak se može odabrati uz primjenu vjerojatnosti, a odnosi se na uzorak u kojemu svaki član populacije ima jednaku vjerojatnost biti dio uzorka i ta vjerojatnost točno se može odrediti. Postoji i uzorak u kojem se članovi odabiru bez primjene vjerojatnosti – vjerojatnost odabira ne može se točno odrediti, a članovi populacije nemaju jednaku vjerojatnost biti dio uzorka.

Umjesto da kontaktiraju sa svakom osobom u populaciji, istraživači mogu odgovoriti na većinu pitanja uzorkovanjem pojedinaca. Zapravo, uzorkovanje je ono što nacionalni zavodi za statistiku rade kako bi prikupili detaljne informacije o stanovništvu, kao što su prosječni prihod kućanstva, razina obrazovanja i vrsta posla koju analizirana skupina ispitanika obavlja. U svojoj srži, istraživački je uzorak kao i svaki drugi uzorak (Andrade, 2016). On je manji dio ili dio nečega što predstavlja veću cjelinu. Kako se odabiru sudionici (slučajnim uzorkovanjem), prvi je korak odrediti populaciju za koju se rezultati istraživanja mogu generalizirati (Andrade, 2019).

Taherdost (2016:19) navodi kako „pristranost uzorkovanja nastaje kada uzorak ne odražava karakteristike populacije“. Pogreške uzorkovanja nastaju kada se za odabir uzorka koristi pogrešna podpopulacija. Različite karakteristike jedinica koje se uzorkuju unutar populacije poput spola, rase ili ekonomskih čimbenika mogu dovesti do pogrešaka u zaključivanju. Sustavne pogreške nastaju kada se rezultati iz uzorka značajno razlikuju od rezultata populacije. Stoga je izuzetno važno koristiti se metodama pravilnog odabira uzoraka (Taherdost, 2016).

Istraživanje u sklopu rada uključuje analizu sustava za uzorkovanje, metode i vrste uzoraka. U drugom poglavlju opisana je metodologija istraživanja i ključne stavke koje naglašavaju važnost iste. Treće poglavlje prikazuje metode uzoraka. Četvrto poglavlje analizira vrste uzoraka i njihove specifičnosti. U raspravi je prikazana sinteza najvažnijih pojmova. U zaključku su sažeti najvažniji elementi rada.

2. Metodologija rada

Uzorci su dijelovi populacije, a izabiru se u skladu s postavljenim ciljem istraživanja, istraživačkim pitanjima te su usko povezani teorijskim i empirijskim odlikama istraživačkog problema. Vrste uzoraka impliciraju primjenu statističkih analiza, a vrste distribucija koje se očekuju na temelju provedenog izbora jedinica u uzorak utječu na primjenu statističkih procedura. Primjena uzoraka u ekonomskim istraživanjima ovisi o okviru izbora jedinica populacije u uzorak, kao i o metodama koje se u istraživanju namjeravaju primijeniti.

Statističke metode i metode uzorkovanja izravno utječu na sam proces i pravilnost istraživanja. Samo istraživanje provedeno u sklopu izrade rada oslanja se na recentnu literaturu iz područja matematike, statistike i ekonomije, a navedeni radovi i istraživanja okosnica su rada. Cilj je završnog rada obuhvatiti ključne pojmove povezane sa statističkim podacima, uzorcima i vrstama uzoraka. U radu su obuhvaćene poveznice među navedenim pojmovima, a to se odnosi na povezivanje definicija, značenja metoda i na načine kojima se provodi planiranje uzorkovanja.

Istraživanje rada zasnovano je na znanstvenim radovima i literaturi koja jasno povezuje važnost upravljanja uzorkovanjem i statistikom u formiranju kvalitetnog znanstvenog istraživanja. Prilikom pisanja rada korišteni su sekundarni podaci koji su uključivali relevantnu znanstvenu literaturu. U pretraživanju su korištene ključne riječi kao što su uzorci, statističke metode, vrste uzoraka, a proučavane su baze znanstvenih radova poput Hrčka i ResearchGate-a.

3. Metoda uzoraka

U poglavlju se opisuje pregled literature povezan s ključnim pojmovima u uzorkovanju. Opisane su definicije populacije (3.1.), podjela statističkih metoda (3.2.) i važnosti pravilnog odabira uzoraka (3.3).

3.1. Populacija

U statistici, populacija je skup pojedinaca iz kojih se odabire statistički uzorak za istraživanje. Za bilo koju selekciju pojedinaca grupiranih po zajedničkom obilježju može se reći da predstavlja populaciju (Taherdost, 2016). U većini upotreba riječ populacija podrazumijeva skupinu pojedinaca ili barem skupinu živih bića. Međutim, statističari populacijom označavaju one jedinice promatranja koje im se nalaze u istraživačkom fokusu. Horvat i Mijoč (2019) navode kako se jedinicama populacije mogu smatrati oni dijelovi statističkog skupa koji su predmet istraživanja. Iako se učestalo pod populacijom smatraju pojedinci, predmet statističke analize ipak mogu biti i objekti, organizacije, geografske ili teritorijalne jedinice, poduzeća, događanja itd. Populaciju je potrebno odrediti vremenski, prostorno i pojmovno. Tako primjeri za populaciju mogu biti: studenti koji su završili diplomski studij akademske godine 2021./22. u Republici Hrvatskoj, djelatnici IT sektora u Osječko-baranjskoj županiji 2021. godine, broj poreznih obveznika u gradu Osijeku na dan 31. 12. 2021. godine itd.

Statističari i istraživači radije proučavaju karakteristike svakog entiteta u populaciji kako bi mogli iznijeti što preciznije zaključke. Međutim, to je većinu vremena nemoguće ili nepraktično budući da su skupovi populacije prilično veliki, stoga se često oslanjaju na metode uzorkovanja. Političko anketiranje dobar je primjer uzorka koji pokazuje sve nedostatke i poteškoće u odabiru slučajnog uzorka stanovništva. Jedan od razloga zašto su posljednje dvije ankete za predsjedničke izbore bile pogrešne mogao bi biti taj što postoji određena skupina pojedinaca koja voljno odgovara na pitanja ankete, a oni možda ne predstavljaju slučajni uzorak populacije vjerojatnih birača (Omair, 2014).

Ipak, ankete mogu biti jedini učinkovit način da se identificiraju i potvrđuju problemi i trendovi koji utječu na širu populaciju. Uzorak je slučajni odabir članova populacije. Uzorak predstavlja manju skupinu jedinica iz populacije koja ima karakteristike cjelokupne populacije. „Zaključci o populaciji često se donose na temelju rezultata iz uzorka. Razlog uporabe uzorka umjesto cijele populacije nalazi se u financijskoj, vremenskoj i tehničkoj opravdanosti, odnosno u uštedi različitih resursa“ (Horvat i Mijoč, 2019:37). Podatke, koji se dobivaju iz statističkog uzorka „statističari koriste kako bi razvili hipoteze o populaciji. U statističkim jednadžbama

populacija se obično označava velikim slovom N, dok se uzorak obično označava malim slovom n“ (Omair, 2014:145).

Parametar je podatak izračunat za podatke iz populacije. Parametri poput aritmetičke sredine i standardne devijacije temeljem izračuna podataka „iz populacije nazivaju se parametrima populacije“ (Omair, 2014:146). Srednja vrijednost populacije i standardna devijacija populacije predstavljeni su grčkim slovima μ i σ . Valjana statistika može se izračunati ili iz uzorka populacije ili iz studije cijele populacije. Cilj je slučajnog uzorka izbjeći pristranost u rezultatima. Uzorak je slučajan ako svaki član cijele populacije ima jednaku vjerojatnost da bude odabran za sudjelovanje. Dok je parametar karakteristika populacije, pokazatelj uzorka karakteristika je uzorka. Inferencijalna statistika omogućuje istraživačima da postavite pretpostavku o parametru populacije na temelju statistike izračunate iz uzorka koji je nasumično odabran iz promatrane populacije. Pokazatelji uzorka primjenjuju se pri testiranju pretpostavki (hipoteza) o nepoznatim parametrima populacije te donošenju zaključaka o odbacivanju ili neodbacivanju postavljenih tvrdnji. Standardna devijacija varijacija je svih podataka u odnosu na prosječnu vrijednost populacije. Ako se „standardna devijacija podijeli s kvadratnim korijenom broja opažanja u uzorku, rezultat se naziva standardna pogreška srednje vrijednosti“ (Delgado-Rodriguez i Llorca, 2004:635).

3.2. Podjele statističkih metoda

Uzorkovanje je postupak odabira podskupine (unaprijed određeni broj promatranja) iz veće populacije. Ono predstavlja uobičajenu tehniku pomoću koje izvode eksperimenti ili istraživanja temeljem kojih se donose zaključci o populaciji, bez potrebe za proučavanjem cijele populacije. Postoje dvije vrste metoda uzorkovanja (Delgado i Llorca, 2004):

Uzorkovanje bez vjerojatnosti (neprobabilističko) – Ovdje se odabire uzorak na temelju neslučajnih kriterija i svaki član populacije nema jednaku mogućnost da bude uključen (Pavlič, 1997).

U kontekstu neprobabilističkog uzorkovanja, vjerojatnost odabira pojedinaca iz ciljne populacije jest nula. Ta vrsta uzorkovanja ne daje reprezentativan uzorak, stoga se opaženi rezultati obično ne mogu generalizirati na ciljnu populaciju. Ipak, nereprezentativni uzorci mogu biti korisni za neke specifične ciljeve istraživanja i mogu pomoći u odgovoru na određena istraživačka pitanja, kao i pridonijeti stvaranju novih hipoteza (Omair, 2014). Različite vrste neprobabilističkog uzorkovanja detaljno su opisane u sljedećem potpoglavlju.

Uzorkovanje uz primjenu vjerojatnosti (probabilističko) – U kontekstu probabilističkog uzorkovanja, sve jedinice ciljne populacije imaju ne-nultu vjerojatnost da će sudjelovati u istraživanju. Ako je jednaka vjerojatnost da će svi sudionici biti odabrani u studiji, upotrebljava se vjerojatnost uzorkovanja, a izgledi da budu odabrani od strane istraživačkog tima mogu se izraziti formulom: $P=1/N$, gdje je P jednaka vjerojatnosti sudjelovanja u studiji i N odgovara veličini ciljne populacije (Omair, 2014). Statističari se koriste različitim statističkim metodama za opis uzoraka, a najvažnije se prikazuju u nastavku.

3.2.1. Procjena aritmetičke sredine

Aritmetička sredina μ numeričke varijable X zbroj je umnoška vrijednosti svakog obilježja i njegove frekvencije podijeljen s opsegom populacije (Horvat i Mijoč, 2014). Izražava se formulom:

$$\mu = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

Populacija i uzorak dva su osnovna pojma unutar statistike. Stanovništvo se može okarakterizirati kao skup pojedinačnih osoba, jedinica ili objekata za koje se istražitelj primarno zanima tijekom svog istraživačkog problema. Andrade (2019:102) populaciju definira kao „skup svih jedinki ili stavki koje se razmatraju u statističkoj studiji“. Uzorak definira kao „dio populacije iz koje se podaci prikupljaju“ (Andrade, 2019:103).

Aritmetička sredina, koja se nekada naziva i srednja vrijednost ili prosjek, jedna je „od središnjih vrijednosti koje se koriste za računanje nekog skupa brojeva kao kvocijent zbroja članova i broja članova skupa“ (Newbold, 2010:63). Izraz "aritmetička sredina" pomaže u „razlikovanju od sličnih terminoloških pojmova poput geometrijske sredine i harmonijske sredine pa ga u statistici i matematici preferiraju“ (Newbold, 2010:63).

Osim u matematici i statistici, aritmetičkom sredinom često se koriste u mnogim različitim područjima kao što je ekonomija te se u određenoj mjeri upotrebljava u gotovo svim akademskim područjima. Na primjer, prihod po stanovniku aritmetički je prosječni dohodak stanovništva zemlje (Newbold, 2010).

Aritmetička sredina često se upotrebljava za izvještavanje o središtu distribucije. Nije robusna statistika, što znači da na nju uvelike utječu izdvojenice (vrijednosti koje se izrazito razlikuju od većine vrijednosti). Izdvojenice se razvrstavaju u dvije kategorije, one umjerene i sumnjive izdvojenice. Za iskrivljene raspodjele, poput raspodjele dohotka za koje su prihodi nekoliko pojedinaca znatno veći od većine, aritmetička sredina možda se ne podudara s nečijim pojmom

"sredine", a robusne statistike, poput medijana, mogu pružiti bolji opis središnje tendencije (Newbold, 2010).

Ako je skup podataka statistička populacija (tj. sastoji se od svih mogućih opažanja, a ne samo od njihovog podskupa), tada se sredina te populacije naziva srednja vrijednost populacije i označava se grčkim slovom μ . Ako je skup podataka statistički uzorak (podskup populacije), tada statistiku koja je rezultat izračuna nazivamo srednjom vrijednošću uzorka i zapisuje se malim slovom x potez (Newbold, 2010). Aritmetička sredina dodjeljuje ponder $1/n$ za svako promatranje, pod pretpostavkom da u zbirci ima n opažanja (Gogala i Šimićević, 2005).

Aritmetička sredina u statistici se može upotrebljavati kako bi se utvrdio središnji položaj unutar distribucije promatranih podataka. Međutim, ona nije uvijek idealan pokazatelj središta distribucije. Često se može dogoditi da se nađu promatrana opažanja koja su znatno veća ili manja od ostatka skupine te se tako ističu (Gogala i Šimićević, 2005). Ovakva odstupanja nisu reprezentativna za skupinu podataka, ali mogu značajno utjecati na aritmetičku sredinu. U pozitivno iskrivljenoj zbirci podataka iznimno velika isticanja povećavaju aritmetičku sredinu, a manje vrijednosti koje se ističu smanjuju vrijednost aritmetičke sredine (Gogala i Šimićević, 2005).

U situacijama kada se primjećuju vrijednosti koje znatno odstupaju od promatranih, mod ili medijan mogu bolje označiti središnju tendenciju skupa podataka od srednje vrijednosti. Mod je vrijednost koja se prikazuje s najvećom frekvencijom. Medijan je srednja točka koja točno razdvaja gornju polovicu i donju polovicu skupa podataka. Vrijednosti koje značajno odstupaju od promatranih imaju znatno manji utjecaj ta dva parametra (osobito mod). Stoga bi mod i medijan mogli biti reprezentativniji za skup podataka s iznimno velikim ili malim vrijednostima koje odstupaju od promatranih (s prisutnim izdvojenicama). U pozitivno asimetričnom skupu podataka medijan i mod manji su od aritmetičke sredine. U negativno asimetričnom skupu podataka medijan i mod veći su od aritmetičke sredine (Novoselac i Rimac, 2014).

Osim aritmetičke sredine, druge dvije vrste prosjeka koje se uobičajeno koriste u financijama jesu geometrijska sredina i harmonijska sredina. Aritmetičku sredinu treba upotrijebiti pri traženju prosjeka niza izvornih podataka, poput cijena dionica. Geometrijska sredina trebala bi se koristiti kada se radi o skupu postotaka koji su izvedeni iz izvornih vrijednosti, kao što je postotna promjena cijena dionica (Novoselac i Rimac, 2014). Ona je prosječna vrijednost ili sredina koja označava središnju tendenciju skupa brojeva (izračunava se uzimanjem korijena

umnoška njihovih vrijednosti). Na primjer: za dani skup od dva broja kao što su 8 i 1, geometrijska sredina jednaka je $\sqrt{(8 \times 1)} = \sqrt{8} = 2\sqrt{2}$ (Dizdar, 2006:78).

„Dakle, geometrijska sredina definira se kao n-ti korijen umnoška n brojeva“ (Dizdar, 2006:89).

Također, izračun geometrijske sredine uzima u obzir složeni učinak po razdobljima koji se ne može obuhvatiti aritmetičkom sredinom. Aritmetička sredina često se koristi za procjenu budućih izvedbi. Prilikom izračuna aritmetičke sredine vrijednosti podataka dodaju se i zatim dijele s ukupnim brojem vrijednosti. Geometrijska sredina može se pronaći množenjem svih brojeva u danom skupu podataka i uzimanjem n-tog korijena za dobiveni rezultat (Jelić, 2010).

$$\text{geometrijska sredina} = (a_1 \cdot a_2 \cdot a_3 \cdot \dots \cdot a_n)^{\frac{1}{n}}$$

Geometrijska sredina ima mnogo prednosti u odnosu na aritmetičku sredinu i koristi se u brojnim područjima (Tadić, 2017). Koristi se u dioničkim indeksima jer mnogi indeksi vrijednosti kojima se koriste financijski odjeli upotrebljavaju izračun geometrijske sredine za izračunavanje godišnjeg prinosa na investicijski portfelj, u financijama za određivanje prosječnih stopa rasta koje su također poznate kao složena godišnja stopa rasta (CAGR) ili, na primjer, u biološkim studijama kao što su dioba stanica i stopa rasta bakterija (Arslanagić, 2015).

Harmonijska sredina može se odnositi na razlomke s različitim nazivnicima. Stoga je ona najprikladniji pristup prosječnim omjerima, npr. u ekonomiji u omjerima P/E i EV/EBITDA. Aritmetička sredina omjera P/E pristrana je ako svi omjeri P/E u skupini ne pokazuju istu vrijednost za nazivnik (ista zarada po dionici), što je rijetko slučaj. Prednost je harmonijske sredine ta što svim podacima u skupini dodjeljuje jednake pondere, bez obzira na to jesu li nazivnici jednaki ili nisu (Jelić, 2010).

$$\text{harmonijska sredina} = \frac{n}{\frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3} + \dots + \frac{1}{a_n}} = \frac{n}{\sum_{i=1}^n \frac{1}{a_i}}$$

Kada se srednja vrijednost uzorka koristi kao procjena prosječne vrijednosti populacije, može se očekivati neka pogreška zbog činjenice da se uzorak ili podskup populacije koristi za izračunavanje procjene. Apsolutna vrijednost razlike između prosječne vrijednosti uzorka, \bar{x} , i populacije, μ , napisane $|\bar{x} - \mu|$, naziva se pogreška uzorkovanja. Intervalna procjena uključuje izjavu o vjerojatnosti o veličini pogreške uzorkovanja. Distribucija uzorka \bar{x} daje osnovu za takvu tvrdnju (Jelić, 2010).

Statističari su pokazali da je srednja vrijednost distribucije uzorka \bar{x} jednaka srednjoj vrijednosti populacije μ te da je standardna devijacija dana σ/\sqrt{n} korijen od \sqrt{n} , gdje je σ standardna devijacija populacije. Standardna devijacija distribucije uzorka naziva se standardna pogreška. Za velike veličine uzorka središnji granični teorem pokazuje da se distribucija uzorka \bar{x} može aproksimirati normalnom raspodjelom vjerojatnosti. Iz prakse statističari obično smatraju da su uzorci veličine 30 ili više dovoljno veliki (Pedišić i Dizdar, 2009).

Simbolički promatrano, \sqrt{n} je standardna pogreška, a 1,96 broj je standardnih pogrešaka od srednje potrebne za uključivanje 95% vrijednosti u normalnu distribuciju (Pedišić i Dizdar, 2009). Tumačenje intervala pouzdanosti od 95% glasi da će 95% intervala na ovaj način sadržavati prosjek populacije (Pedišić i Dizdar, 2009). Dakle, svaki interval izračunat na ovaj način ima 95% pouzdanost da sadrži prosjek populacije. Niža razina pouzdanosti dovodi do još užih intervala. U praksi se najčešće koristi interval pouzdanosti od 95% (Arslanagić, 2015). Veće veličine uzoraka dovode do manjih grešaka. To opažanje čini osnovu za postupke koji se koriste za odabir veličine uzorka. Veličine uzoraka mogu se odabrati tako da interval pouzdanosti zadovolji sve željene zahtjeve o veličini margine pogreške (Arslanagić, 2015).

Upravo opisani postupak za izradu intervalnih procjena prosječne populacije temelji se na upotrebi velikog uzorka. U slučaju malog uzorka, to jest, u slučajevima gdje je veličina uzorka n manja od 30 – t raspodjela se koristi pri specificiranju margine pogreške i izgradnji procjene intervala pouzdanosti (Arslanagić, 2015).

Za kvalitativne varijable udio stanovništva najčešći je parametar koji je od interesa istraživačima (Dizdar, 2006). Bodovna procjena udjela stanovništva izračunava se udjelom uzorka (Arslanagić, 2015). Postupci procjene mogu se proširiti na dvije populacije radi usporednih studija. Na primjer, pretpostavimo da se provodi studija kako bi se utvrdile razlike između plaća koje se isplaćuju populaciji muškaraca i populaciji žena. Dva nezavisna jednostavna slučajna uzorka, jedan iz populacije muškaraca i jedan iz populacije žena, osigurala bi dva prosječna uzorka, \bar{x}_1 i \bar{x}_2 . (Tadić, 2017).

„Aritmetička sredina može se procijeniti brojem i intervalom. Ako veličina uzorka teži prema beskonačno, distribucija aritmetičkih sredina teži normalnom obliku. Kod malih uzoraka distribucija aritmetičkih sredina ima oblik Studentove ili t -distribucije. Stoga vrijedi:

$n > 30 \Rightarrow$ koristi se normalna distribucija i

$n \leq 30 \Rightarrow$ koristi se Studentova ili t -distribucija“ (Dizdar, 2006:91).

Aritmetička sredina populacije μ „parametar je koji se procjenjuje brojem: na primjer ukoliko istraživače zanimaju prosječna primanja stanovnika nekog područja, oni mogu izabrati uzorak od n stanovnika tog područja, izračunati aritmetičku sredinu uzorka i zaključiti da su ona istovjetna prosječnim primanjima stanovnika cijelog područja. Prilikom provođenja istraživanja važno je formirati interval određene širine, ovisno o željenoj pouzdanosti (ili povjerenju) procjene – što je interval širi, procjena je pouzdanija, tj. veća je vjerojatnost da će se u njemu naći aritmetička sredina populacije“ (Pedešić i Dizdar, 2009:43).

3.2.2. Procjena totala osnovnog skupa

„Na temelju podataka iz uzorka procjenjuje se total za jednu populaciju, tj. statistički skup. Total osnovnog skupa može se procijeniti brojem i intervalom“ (Dizdar, 2006:60).

Postoji interval unutar kojega se „s određenom vjerojatnošću nalazi aritmetička sredina populacije“ (Dizdar, 2006:60). Zahvaljujući istom, sljedećom formulom može se procijeniti aritmetička sredina populacije:

$$\bar{x} - t_p \cdot s_x < \mu < \bar{x} + t_p \cdot s_x$$

Iz prethodnog izraza slijedi „da se u procjeni totala populacije može pogriješiti N puta onoliko koliko se griješi u procjeni njezine aritmetičke sredine“ (Dizdar, 2006:60).

„Za intervalnu procjenu totala potrebna je i standardna pogreška totala: $\sigma_T = N \times \sigma_\mu$.

Intervalna procjena totala osnovnog skupa na osnovi uzorka uz odgovarajuću razinu pouzdanosti procjene je“:

$$P\left\{\hat{T} - z_{\alpha/2}se(\hat{T}) < T < \hat{T} + z_{\alpha/2}se(\hat{T})\right\} = 1 - \alpha$$

(Dizdar, 2006:60).

3.2.3. Procjena proporcije osnovnog skupa

„Proporcija konačnog osnovnog skupa je parametar koji označava omjer članova skupa s određenim oblikom obilježja M i opsega skupa N “ (Dizdar, 2006:67), tj.

$$p = \frac{M}{N}$$

„Proporciju preostalih $N - M$ jedinica koje nemaju traženo obilježje označava se sa q :

$$q = \frac{N - M}{N}$$

Vrijedi: $p + q = 1$. Procjenitelj proporcije osnovnog skupa brojem je proporcija uzorka“ (Dizdar, 2006:67):

$$\hat{q} = \frac{n-m}{n}$$

Standardna pogreška može se izračunati preko dva izraza:

$$\sigma_{\hat{p}} = \sqrt{\frac{\hat{p} \cdot \hat{q}}{n - 1}}$$

$$\text{(ako je } f < 0.05), \sigma_{\hat{p}} = \sqrt{\frac{\hat{p} \cdot \hat{q}}{n-1} \cdot \frac{N-n}{N-1}}$$

(ako je $f \geq 0,05$)

Interval procjene proporcije za velike uzorke:

$$P \left\{ \hat{p} - z_{\frac{\gamma}{2}} \sigma_{\hat{p}} \leq p \leq \hat{p} + z_{\frac{\gamma}{2}} \sigma_{\hat{p}} \right\} = 1 - \gamma.$$

(Horvat i Mijoč, 2015).

3.3. Proces izabiranja uzoraka

Uzorak je „statistički značajan dio populacije, a ne cijela populacija“ (Pedišić i Dizdar, 2009:32). Iz tog razloga statistička analiza uzorka mora izvijestiti o približnoj standardnoj devijaciji, ili standardnoj grešci, njegovih rezultata iz cijele populacije. Samo analiza cijele populacije ne bi imala standardnu pogrešku.

Prvo, mora se razumjeti razlika između populacije i uzorka i odrediti ciljna populacija određenog istraživanja. „Stanovništvo je cijela skupina o kojoj se želi, hipotetski, iznijeti zaključak. Uzorak je specifična skupina pojedinaca od kojih se prikupljaju podatci“ (Newbold, 2010:45). Stanovništvo se može definirati u smislu spola, zemljopisnog položaja, starosti, dohotka i mnogih drugih karakteristika (Newbold, 2010). Kategorija po kojoj se definira populacija može biti vrlo široka ili određenija: možda se žele izvesti zaključci o cijeloj odrasloj populaciji zemlje; možda se istraživanje fokusira na kupce određene tvrtke, pacijente s

određenim zdravstvenim stanjem ili učenike u samoj školi. Važno je pažljivo definirati ciljanu populaciju u skladu sa svrhom i praktičnošću istraživanja ili projekta. Ako je populacija brojna, demografski pomiješana i geografski raspršena, postaje teže dobiti pristup reprezentativnom uzorku (Šošić, 2006).

4. Karakteristike uzoraka

Prilikom procesa uzorkovanja istraživači moraju provesti pripremu studije kako bi mogli adekvatno opisati željenu populaciju. U poglavlju je opisano kako se odabire veličina uzoraka (4.1.), reprezentativnost uzoraka (4.2.) i vrste uzoraka (4.3.).

4.1. Veličina uzoraka

Postoji nekoliko teorijskih i praktičnih razloga koji sprječavaju istraživače u provođenju velikih istraživanja, uključujući (Taherdoost, 2016):

- etička pitanja: neetično je uključiti veći broj pojedinaca nego što je to stvarno potrebno;
- proračunska ograničenja: visoki troškovi populacijskog istraživanja često ograničavaju njegovu upotrebu kao strategiju za odabir sudionika za istraživanje;
- logistika: popisi često nameću velike izazove u pogledu potrebnog osoblja, opreme itd. za provođenje studije;
- vremenska ograničenja: vrijeme potrebno za planiranje i provođenje ankete temeljene na popisu može biti pretjerano; i,
- nepoznata veličina ciljne populacije: nedostatak informacija o svim postojećim korisnicima onemogućuje provođenje studije temeljene na popisu stanovništva.

Svi ti razlozi objašnjavaju zašto se uzorci koriste prilikom provedbe znanstvenih istraživanja. Okvir uzorkovanja stvarni je popis elemenata (ljudi, životinja, predmeta) u promatranoj populaciji iz kojih se razmatra odabir uzorka. Veličinu uzorka određuje sam istraživač. Broj elemenata koje bi trebalo uključiti u uzorak ovisi o različitim čimbenicima, uključujući veličinu i varijabilnost populacije i sam dizajn istraživanja. Postoje različiti kalkulatori i formule za procjenu i izračun veličine uzorka ovisno o tome što se želi postići statističkom analizom (Šošić, 2006).

Okvir uzorka može predstavljati skupinu pojedinaca koja se odabire iz ciljane populacije s obzirom na proces uzorkovanja koji se koristi u istraživačkoj studiji. Na primjer, da bi identificirao slučajeve kožnog melanoma, istraživač može razmotriti za okvir uzorka

nacionalni sustav registra raka. S obzirom na to da uzorak može predstavljati samo dio ciljne populacije, istraživač treba pažljivo ispitati odgovara li odabrani okvir uzorka ciljevima ili hipotezama istraživanja, a posebno postoje li strategije za prevladavanje ograničenja okvira uzorka.

U idealnom slučaju, uzorak bi trebao biti nasumično odabran i reprezentativan za populaciju. Korištenje metode uz primjenu vjerojatnosti u uzorkovanju (kao što je jednostavno nasumično uzorkovanje ili stratificirano uzorkovanje) smanjuje rizik od pristranosti uzorkovanja i povećava unutarnju i vanjsku valjanost. Iz praktičnih razloga, istraživači se često koriste metodama neprobabilističkog uzorkovanja. Neprobabilistički uzorci biraju se za specifične kriterije; mogu biti prikladniji ili jeftiniji za pristup. Zbog nasumičnih metoda odabira bilo koji statistički zaključci o široj populaciji bit će slabije reprezentativni nego kod probabilističkih uzoraka (Malhotra i Birks, 2006).

Razlozi uzorkovanja prema Malhotri i Briksu mogu se sažeti u:

- Nužnost: ponekad jednostavno nije moguće proučiti cijelu populaciju zbog njezine veličine ili nepristupačnosti.
- Praktičnost: lakše je i učinkovitije prikupljati podatke iz uzorka.
- Isplativost: manje je uključenih troškova sudionika, laboratorija, opreme i istraživača.
- Upravlјivost: pohranjivanje i izvođenje statističkih analiza na manjim skupovima podataka lakše je i pouzdano.

4.2. Reprezentativnost uzoraka

Izrazi nepristran i preciznost dobili su posebna značenja u statistici. Kada se kaže da je mjerenje nepristrano, misli se da će prosjek velikog skupa nepristranih mjerenja biti blizu prave vrijednosti. Kada se kaže da je uzorak precizan, misli se da je ponovljiv (Malhotra i Briks, 2006). Ponovljena mjerenja mogu biti sličnih intervalnih vrijednosti, ali ne nužno blizu prave vrijednosti. Istraživač u idealnom slučaju želi mjerenje koje je točno i precizno. Neki autori poistovjećuju nepristranost s točnošću, ali to nije univerzalno, a drugi koriste izraz točnost za mjerenje koje je i nepristrano i precizno. Poznato je da je procjena parametra temeljem slučajnog uzorka nepristrana. Kako se veličina uzorka povećava, postaje preciznija (Malhotra i Birks, 2006).

Druga je mogućnost korištenje randomizacije, odnosno nasumične podjele u istraživanju. To osigurava da nema pristranosti u raspodjeli, a ispitanici su u svakoj skupini usporedivi u

poznatim i nepoznatim prognostičkim čimbenicima. Čak i ako se osigura da svaki član populacije ima poznatu i obično jednaku vjerojatnost da bude uključen u uzorak, iz toga ne slijedi da će niz uzoraka kreiranih iz jedne populacije ispunjavati svaki kriterij i biti identičan. Oni će pokazati slučajne varijacije od jedne do druge skupine, a varijacije mogu biti male ili znatne. Na primjer, niz uzoraka tjelesne temperature zdravih pojedinaca pokazao bi vrlo male varijacije od jednog do drugog, ali bi varijacije između uzoraka sistoličkog krvnog tlaka bile znatne. Stoga varijacije između uzoraka dijelom ovise o količini varijacija u populaciji iz koje su oni odabrani (Delgado-Rodriguez i Llorca, 2004).

Omair (2014) navodi da što je više članova populacije uključeno u uzorak, to uzorak ima veću vjerojatnost točnijeg predstavljanja populaciju, pod uvjetom da se za konstruiranje uzorka koristi slučajni proces. Posljedica toga je da, ako su dva ili više uzorka odabrana iz populacije, što su veći, veća je vjerojatnost da će nalikovati jedan drugome – opet pod uvjetom da se slijedi tehnika slučajnog odabira. Stoga varijacije između uzoraka dijelom ovise i o veličini uzorka. Obično istraživači nisu u poziciji uzeti slučajni uzorak, nego se koriste metodama uzorkovanja kako bi na praktičan način što točnije mogli prikazati promatranu populaciju. Dakle, u praksi istraživači prilikom odabira metode uzorkovanja moraju uzeti u obzir način i praktičnost provedbe studije u odnosu na faktore poput troškova i vremena. Da bi se napravile valjane generalizacije, istraživači trebaju utvrditi da je uzorak na neki način reprezentativan za populaciju u cjelini i iz tog razloga prva je faza u izvješću opisati uzorak, recimo prema dobi, spolu i statusu, tako da drugi čitatelji mogu odlučiti je li reprezentativan za vrstu uzorka s kojima se susreću (Omair, 2014).

Pogreška uzorkovanja razlika je između parametra populacije i statistike uzorka. U istraživanju pogreška uzorkovanja razlika je između srednje vrijednosti parametra populacije i promatrane srednje vrijednosti koja je dobivena u istraživanju. Pogreške uzorkovanja događaju se čak i kada istraživači koriste slučajno odabran uzorak. To je zato što slučajni uzorci nisu identični populaciji u smislu numeričkih mjera kao što su srednje vrijednosti i standardne devijacije. Budući da je cilj znanstvenog istraživanja generalizirati nalaze iz uzorka na populaciju, cilj je istraživača da pogreška uzorkovanja bude niska. Pogreška uzorkovanja može se smanjiti povećanjem veličine uzorka (Normain i sur., 2012).

4.3. Vrste uzoraka

4.3.1. Metode uzorkovanja uz primjenu vjerojatnosti

Uzorkovanje uz primjenu vjerojatnosti znači da svaki član populacije ima priliku biti izabran u uzorak. Uglavnom se koristi u kvantitativnim istraživanjima. Ako želimo rezultate koji su reprezentativni za cijelu populaciju, tehnike uzorkovanja prema vjerojatnosti najvaljaniji su izbor. Postoje četiri glavne vrste uzorka uz primjenu vjerojatnosti (Vrdoljak, 2007).

1. Jednostavno slučajno uzorkovanje – U jednostavnom slučajnom uzorku svaki član populacije ima jednaku vjerojatnost odabira. Okvir uzorkovanja trebao bi obuhvaćati cijelu populaciju. Za provođenje ove vrste uzorkovanja mogu se koristiti alati poput generatora slučajnih brojeva ili drugih tehnika koje se u potpunosti temelje na slučajnosti (Sarapa, 2002).



Slika 1. Primjer programa za generiranje slučajnog odabira random.org (Dostupno na: <https://www.random.org>, pristupljeno: 19. 9. 2022.)

2. Sistematično uzorkovanje – Sistematično uzorkovanje slično je jednostavnom slučajnom uzorkovanju, ali se jednostavnije provodi u praksi nego jednostavni slučajni uzorak. Ovo uzorkovanje uključuje slučajan odabir početnog mjesta, a zatim se odabere svaka n-ta promatrana jedinica individualno, ovisno o nacrtu istraživanja. Na primjer: uzorkuje se svaki 5. artikl na polici, odabire se svako 10. ime unutar promatrane populacije ili istraživač bira svakog trećeg kupca koji dođe u trgovinu (Vrdoljak, 2007).
3. Stratificirano (slojevito) uzorkovanje – Stratificirano uzorkovanje uključuje podjelu stanovništva na subpopulacije koje se mogu razlikovati na važne načine. Omogućuje formiranje preciznijih zaključaka osiguravajući da je svaka podskupina pravilno

zastupljena u uzorku. Da bi se koristila ova metoda uzorkovanja, stanovništvo se dijeli u podskupine (slojeve) na temelju relevantnih karakteristika (npr. spol, dob, kategorija prihoda, uloga posla). Na temelju ukupnih udjela stanovništva izračunava se koliko jedinica treba uzeti u uzorke iz svake podskupine. Zatim se koristi slučajno ili sustavno uzorkovanje za odabir uzorka iz svake podskupine (Šošić, 2006).

4. Klustersko uzorkovanje – Klustersko uzorkovanje također uključuje „podjelu populacije u podskupine, ali svaka podgrupa treba imati slične karakteristike kao i cijeli uzorak. Umjesto uzorkovanja jedinica iz svake podskupine, nasumično se odabiru cijele podgrupe“ (Delgado – Rdriguez i Llorca, 2004:635). Ako je to moguće, može se uključiti svaka jedinica iz svakog uzorkovanog klastera. Ako su sami klasteri veliki, također se mogu uzorkovati pojedinci iz svakog klastera koristeći jednu od navedenih tehnika. To se naziva uzorkovanje u više koraka. Ta je metoda dobra za rješavanje velikih i raštrkanih populacija, ali postoji veći rizik od pogreške u uzorku jer mogu postojati značajne razlike među klasterima. Teško je jamčiti da su uzorkovani klasteri doista reprezentativni za cijelu populaciju (Šošić, 2006). Podjela uzoraka u klastera najvažniji je dio procesa. Kvaliteta klastera i koliko dobro predstavljaju veću populaciju određuju valjanost rezultata. U idealnom slučaju, istraživači žele da odabrani klasteri ispunjavaju sljedeće kriterije. Populacija svakog klastera trebala bi biti što je moguće raznolikija. Također žele da svaka potencijalna karakteristika cijele populacije bude predstavljena u svakom klasteru. Svaki klaster treba imati sličnu distribuciju karakteristika kao distribucija populacije u cjelini. Uzeti zajedno, klasteri bi trebali pokriti cjelokupnu populaciju. Ne smije biti preklapanja među klasterima (tj. iste jedinice ne bi se smjele pojaviti u više od jednom klasteru). Međutim, u praksi klasteri često ne predstavljaju savršene karakteristike populacije, zbog čega ova metoda pruža manju statističku sigurnost od jednostavnog slučajnog uzorkovanja. Budući da su klasteri obično prirodne skupine, poput škola, gradova ili kućanstava, često su homogeniji od populacije u cjelini. Istraživači moraju biti svjesni toga kada provode svoja istraživanja jer navedeno utječe na valjanost istih. Ako je svaki klaster sam po sebi predstavnik veće populacije, slučajno odabiranje i uzorkovanje iz klastera omogućuje istraživačima da oponašaju jednostavno nasumično uzorkovanje, što zauzvrat podupire valjanost konačnih rezultata. Suprotno tome, ako klasteri nisu reprezentativni, tada će slučajno uzorkovanje omogućiti prikupljanje podataka o raznolikom nizu klastera, što bi istraživačima opet djelomično moglo pružiti pregled populacije kao cjeline. Klustersko uzorkovanje obično se koristi zbog svojih praktičnih

prednosti, ali ima neke nedostatke u smislu statističke valjanosti. Klastersko uzorkovanje vremenski je i troškovno učinkovito, posebno za uzorke koji su geografski široko rasprostranjeni i koje bi bilo teško pravilno uzorkovati na drugi način. Budući da klastersko uzorkovanje koristi randomizaciju, ako je populacija pravilno grupirana, studija će imati visoku vanjsku valjanost jer će uzorak odražavati karakteristike veće populacije. Zaključno, klastersko uzorkovanje mnogo je složenije za planiranje od drugih oblika uzorkovanja.

5. Složeno ili višestupanjsko uzorkovanje - Ova metoda vjerojatnog uzorkovanja kombinira različite strategije u odabiru jedinica uzorka (Sarapa, 2002).

4.3.2. Metode neprobabilističkog uzorkovanja

U uzorku koji nije temeljen na vjerojatnosti pojedinci se biraju na temelju neslučajnih kriterija i svaki pojedinac nema jednaku vjerojatnost da bude uključen (Vrdoljak, 2007).

Toj vrsti uzorka lakše je i jeftinije pristupiti, ali postoji veći rizik od pristranosti uzorkovanja. To znači da su zaključci o populaciji slabiji nego s uzorcima vjerojatnosti, a zaključci samim time mogu biti ograničeniji. Ako se koristi uzorak koji nije baziran na vjerojatnosti, ipak bi se trebalo nastojati učiniti ga što reprezentativnijim za populaciju. Tehnike uzorkovanja bez vjerojatnosti često se koriste u istraživanjima (Zenzerović, 2004).

Postoje 4 glavne vrste uzorkovanja bez vjerojatnosti (Zenzerović, 2004).

1. Uzorkovanje prema pogodnosti ili praktično uzorkovanje - Uzorak pogodnosti jednostavno uključuje pojedince koji su istraživaču najpristupačniji. Ovo je jednostavan i jeftin način prikupljanja početnih podataka, ali ne postoji način da se utvrdi je li uzorak reprezentativan za populaciju pa ne može dati općenite rezultate. Obično je praktično uzorkovanje omiljena tehnika uzorkovanja među studentima jer je jeftina i laka opcija u usporedbi s drugim tehnikama. Praktično uzorkovanje često pomaže u prevladavanju mnogih ograničenja povezanih s istraživanjem. Na primjer, korištenje prijatelja ili obitelji kao dijela uzorka, lakše je od ciljanja nepoznatih pojedinaca (Šošić, 2006). Također, ova vrsta uzorka koristi se i kao početna, odnosno pilot-studija, kako bi istraživač imao jasniju sliku o jedinicama populacije koje planira izučavati.
2. Dobrovoljno uzorkovanje – Slično je uzorku pogodnosti, a uzorak dobrovoljnog odgovora uglavnom se temelji na jednostavnosti pristupa. Umjesto da je istraživač birao sudionike i izravno im se obraćao, sudionici istraživanja volontiraju (npr. odgovaranjem na javnu

internetsku anketu). Uzorci dobrovoljnih odgovora uvijek su barem donekle pristrani jer će neki pojedinci imati veću vjerojatnost volontiranja od drugih (Zenzerović, 2004).

3. Namjerno uzorkovanje - Ova vrsta uzorkovanja poznata je i kao uzorkovanje prosudbe. Uključuje istraživača koji koristi svoju stručnost za odabir uzorka koji je najkorisniji za svrhe istraživanja. Često se koristi u kvalitativnim istraživanjima gdje istraživač želi steći detaljna znanja o određenom fenomenu, a ne donijeti statističke zaključke, ili gdje je populacija vrlo mala i specifična. Učinkovit ciljani uzorak mora imati jasne kriterije i obrazloženje za uključivanje. Namjerno uzorkovanje strategija je u kojoj su određene postavke osobe ili događaja namjerno odabrani kako bi pružili važne informacije koje ne mogu biti dobivene iz drugih izbora. To je mjesto gdje istraživači uključuju slučajeve ili sudionike u uzorak jer vjeruju da opravdavaju uključivanje (Vrdoljak, 2007).
4. Uzorkovanje „poput grude snijega“ (eng. snowball sampling) - Uzorkovanje grude snijega nenasumična je metoda uzorkovanja koja upotrebljava nekoliko slučajeva kao pomoć kako bi potaknula druge slučajeve da sudjeluju u studiji, čime se povećava veličina uzorka. Ovaj pristup najprimjenjiviji je u malim populacijama koje su teško dostupne zbog svoje zatvorene prirode, na primjer zatvoreni društvo i nedostupne profesije (Vrdoljak, 2007).

U Tablici 1. ukratko su pregledno prikazane sažete snage i slabosti različitih tehnika uzorkovanja.

Tablica 1. Prednosti i slabosti tehnika uzorkovanja

Tehnike uzorkovanja	Snage	Slabosti
dobrovoljno uzorkovanje	najjeftinije, najmanje vremenski zahtjevno, najjednostavnije	selekcija podložna utjecaju istraživača, nije preporučljivo za deskriptivna istraživanja
namjerno uzorkovanje	uzorak se može kontrolirati prema određenim karakteristikama	ne dopušta generalizaciju, subjektivno

uzorkovanje prema pogodnosti	jeftino, jednostavno, korisno za eksplorativna istraživanja	selekcija ovisi o ispitivaču, nema pouzdanosti
uzorkovanje poput grude snijega	može procijeniti rijetke karakteristike	vremenski zahtjevno
jednostavno slučajno uzorkovanje	jednostavno za razumijevanje, rezultati laki za projekciju	teško konstruirati okvir za uzorkovanje, skupo, smanjena preciznost, nema pouzdanosti o reprezentativnosti uzorka
sistematično uzorkovanje	može povećati reprezentativnost, lako primjenjivo	može smanjiti reprezentativnost
slojevito uzorkovanje	uključuje sve važne subpopulacije, precizno	teško izabrati relevantne varijable za stratifikaciju, nije jednostavno stratificirati prema puno varijabla, skupo
klustersko uzorkovanje	jednostavno, jeftino	nije precizno, teško analizirati podatke

Izvor: prilagodba autora prema Malhotra i Birks (2006)

Prije početka istraživanja, istraživači moraju ovisno o ciljevima svoga istraživanja odrediti koju će metodu uzorkovanja koristiti. Prije odabira određene tehnike uzorkovanja, potrebno je upoznati prednosti i slabosti istih koje su prikazane u Tablici 1 (Malhotra i Birks, 2006).

4.3.3. Primjeri uzorkovanja na studentima EFOS-a

Važnost i metode odabira uzoraka u najvećoj mjeri ovise o vrsti i postavkama istraživanja. U ovom će se potpoglavlju prikazati na koji bi se način iz populacije studenata Ekonomskog fakulteta u Osijeku (EFOS) birali uzorci za provođenje istraživanja.

Jednostavno slučajno uzorkovanje – U jednostavnom slučajnom uzorku svaki član populacije ima jednaku vjerojatnost za odabir. Okvir uzorkovanja trebao bi obuhvaćati cijelu populaciju studenata EFOS-a. Ako istraživači žele istražiti prosječnu visinu studenata EFOS-a, koristili bi se tom metodom na način da upotrijebe alate poput generatora slučajnih brojeva ili drugih tehnika koje se u potpunosti temelje na slučajnosti kako bi odabrali studente koji će predstavljati cijeli fakultet.

Sustavno uzorkovanje – Svaki član populacije naveden je s brojem, ali umjesto nasumičnog generiranja brojeva, pojedinci se biraju u redovitim intervalima. Ako se koristi ta tehnika, važno je osigurati da na popisu nema skrivenog uzorka koji bi mogao iskriviti odabrani uzorak populacije. Primjerice, ako grupiramo studente prema prosjeku, a zbog pretpostavljene Gaussove krivulje rasporeda ocjena, postoji rizik da interval preskoči ljude u rubnim kategorijama, što rezultira uzorkom koji je asimetričan prema studentima koji polažu ispite s ocjenom dobar (3).

Stratificirano (slojevito) uzorkovanje uključuje podjelu studenata na subpopulacije koje se mogu razlikovati na važne načine. Omogućuje formiranje preciznijih zaključaka osiguravajući da je svaka podskupina pravilno zastupljena u uzorku. Da bi se koristila ova metoda uzorkovanja, studenti se dijeli u podskupine (nazvane slojevi) na temelju relevantnih karakteristika (npr. spol, dob, studijski smjer, godina na kojoj studiraju). Na temelju ukupnih udjela studenata izračunava se koliko ispitanika treba uzeti u uzorke iz svake podskupine. Zatim se koristi slučajno ili sustavno uzorkovanje za odabir uzorka iz svake podskupine.

Složeno ili višestupanjsko uzorkovanje kombinira različite strategije u odabiru jedinica uzorka. U takvim se istraživanjima najčešće koristi više prethodno navedenih metoda. Na primjer, uzorkovanje se provodi u više faza. Ako istraživači žele proučiti učestalost korištenja instrukcija za polaganje ispita kod studenata starijih od 20 godina na fakultetu, mogu koristiti sljedeću metodu. Podijelit će studente EFOS-a po obiteljskim primanjima te tako dobiti prvu listu. Iz te liste izabrat će se sistematski 100 studenata. Zatim će se iz grupe od 100 studenata izabrati samo studenti iznad 20 godina. Prilikom statističke analize moraju se dobro poručiti svi stupnjevi uzorkovanja kako bi se utvrdile točne procjene. Na primjer, istraživanje na

Ekonomskom fakultetu želi ispitati sve svoje studente. Diješe studente prema godini na kojoj studiraju (na više klastera). Zatim istraživači upotrebljavaju generator slučajnih brojeva za odabir nekih klastera. Nakon toga anketiraju sve studente unutar svakog odabranog klastera. Moguće je i napraviti istraživanje unutar populacije studenata da bi se analizirao postotak studenata zaraženih virusom COVID – 19. Umjesto da se odaberu svi studenti i provode testiranje na svima, mogu se podijeliti na klastere (npr. prema studijskom programu, godini studija, dobi), a zatim se nasumično odaberu klasteri koji će se testirati. Istraživači će zabilježiti koliki je broj studenata zaražen u navedenom klasteru. U praksi stratificirano i klaster uzorkovanje često izjednačavaju. U stratificiranom uzorkovanju istraživači se koriste svim grupama i nekima od članova u svakoj grupi, a u klaster uzorku koriste se nekim od grupa i svim članovima u svakoj grupi.

Uzorak pogodnosti uključuje pojedince koji su istraživaču najpristupačniji. Na primjer, na populaciji studenata EFOS-a, zaposleni vanjski suradnik kao predavač Tjelesne i zdravstvene kulture mogao bi provesti istraživanje na studentima kojima predaje kako bi procijenio u kojoj se mjeri studenti EFOS-a bave sportom.

Uzorak dobrovoljnog odgovora uglavnom se temelji na jednostavnosti pristupa. Umjesto da istraživač bira sudionike i izravno im se obraća, pojedinci dobrovoljno pristupaju istraživanju (na primjer odgovaranjem na javnu internetsku anketu). Recimo da na stranici fakulteta istraživač zamoli studente da ispune anketu vezanu za njegov doktorski rad kako bi dobio podatke o promatranj populaciji. Ipak, neki će studenti radije ispunjavati ovakve ankete od drugih te stoga može doći do pogrešaka u predstavljanju cijele populacije studenata EFOS-a.

Namjerno uzorkovanje uključuje istraživača koji koristi svoju stručnost za odabir uzorka koji je najkorisniji za svrhe istraživanja. Učinkovit ciljani uzorak mora imati jasne kriterije i obrazloženje za uključivanje. Na primjer, ako istraživač želi istražiti zašto su studenti, koji nisu Osječani, upisali Ekonomski fakultet u Osijeku, za svoje istraživanje pozvat će samo studente EFOS-a koji nisu rođeni, ne žive i nisu odrasli u Osijeku.

5. Rasprava

Način na koji se iz populacije izabire uzorak, koji će se koristiti u istraživačkoj studiji, može značajno utjecati na to koliko su točni rezultati istraživačke studije. Postoji mnogo načina i strategija za odabir jedinica u uzora. Neki od njih, ovisno o dizajnu studije, rezultiraju odabirom prikladnijih uzoraka od drugih. Svaki puta kada se uzima uzorak unutar populacije pronalaze se različite vrijednosti podataka.

Nakon što dobiju svoje podatke, bilo iz populacije ili uzorka, istraživači moraju odlučiti kako sažeti i analizirati svoje podatke. Razumijevanje kako statistički analizirati i uzorkovati podatke nužno je definirati prije same provedbe istraživanja. Međutim, uz razlike koje mogu biti prisutne između populacije i uzorka, mogu se pojaviti pogreške uzorka. Ako istraživanje nije provedeno na ispravan način, odabrana je kriva metoda uzorkovanja. Stoga je ključno upotrijebiti najrelevantniju i najkorisniju metodu uzorkovanja u danoj istraživačkoj situaciji. Svakako, prilikom odabiranja metode uzorkovanja za istraživanje potrebno je eliminirati učinke zbunjujućih varijabli kontrolirajući sve čimbenike koji se mogu kontrolirati. Uzorak može biti loše prikupljen ili zbog pristranosti ili zbog pogreške u mjerenju.

Provedba istraživanja u koju su uključene sve jedinice promatrane populacije rijetko se provodi u društvenim znanostima. Umjesto toga istraživači proučavaju uzorke odabrane iz populacija. Metode uzorkovanja tehnike su za odabir uzoraka iz populacija. Okvir uzorka može, na primjer biti registar poslovnih subjekata iz kojega je uzorak izvučen, skupina jedinki ili proizvoda koji se analiziraju. Tehnike uzorkovanja koje se oslanjaju na vjerojatnosti su metode kod kojih je vjerojatnost da će jedinica populacije biti odabrana za uključivanje u uzorak uvijek poznata. Slučajno uzorkovanje postupak je za odabir sudionika pri čemu svaki sudionik ima jednaku vjerojatnost da bude uključen u uzorku. Na istraživačkom projektu resursi su ograničeni, a vrijeme, novac i ljudi ograničeni su resursi. Iz tog razloga većina istraživačkih projekata ima za cilj prikupljanje podataka temeljem uzoraka, a ne iz cjelokupne populacije (popis stanovništva jedna je od rijetkih iznimaka). Uzorkovanje ima mnoge prednosti za istraživače.

Za kontaktiranje sa svim jedinicama populacije potrebno je vrijeme, a do pojedinih jedinica nije moguće pristupiti ili odbijaju kontakt za uključivanje u istraživanje. Slučajno uzorkovanje unutar populacije može se brže provesti od anketiranja svih u populaciji te se može reći kako uzorkovanje štedi istraživačima vrijeme. Broj uzoraka o kojima istraživač prikuplja informacije izravno je povezan s cijenom studije. Uzorkovanje štedi novac dopuštajući istraživačima da prikupe iste odgovore iz uzorka koje bi dobili od populacije.

Nenasumično uzorkovanje znatno je jeftinije od slučajnog uzorkovanja jer smanjuje troškove povezane s pronalaženjem ispitanika i prikupljanjem podataka. Uzorkovanje omogućuje istraživačima da postavljaju sudionicima više pitanja i prikupe veći obujam podataka nego kontaktiranjem sa svima u populaciji.

Neovisno o nacrtu uzorka istraživači podatke mogu prikupiti putem interneta što je olakšalo provedbu nekih znanstvenih istraživanja online putem. Konačno, najbolja metoda uzorkovanja uvijek je ona koja bi mogla najbolje odgovoriti na istraživačko pitanje, a istovremeno dopušta drugima da iskoriste rezultate studije (općenitost rezultata).

6. Zaključak

Statistika je znanstvena disciplina s brojnim primjenama u različitim znanstvenim poljima. Općenito, može se reći da je statistika znanost o prikupljanju, analiziranju, tumačenju i izvođenju zaključaka temeljem podataka. Drugim riječima, statistika je metodologija koju su znanstvenici i matematičari razvili za tumačenje i izvođenje zaključaka iz prikupljenih podataka. U najvećem broju istraživačkih problema nemoguće je ili nepraktično ispitati cijelu populaciju. Osim toga takvi istraživački pristupi mogu biti preskupi za provođenje s obzirom na potrebno vrijeme ili novac. U analizi podataka znanstvenici biraju pažljivo pristupe odabiru uzorka, a s ciljem da je uzorak što sličniji populaciji. Nijedna vrsta uzorkovanja nije prikladna za sva istraživanja. Istraživači odabirom uzorka nemaju jamstvo da će prikupiti reprezentativni uzorak te je navedeno jedno od primarnih ograničenja uzorkovanja.

Pravilno odabrane metode uzorkovanja mogu rezultirati uzorcima koji daju približne rezultate kao i populacija. Statističko zaključivanje proces je korištenja podataka iz uzorka za izradu procjena ili provjeru hipoteza o populaciji. Pogreška uzorkovanja može se definirati kao „razlika između parametra populacije i statistike uzorka koja se koristi za njegovu procjenu. Na primjer, razlika između prosječne vrijednosti populacije i prosječne vrijednosti uzorka, pogreška je uzorkovanja“ (Andrade, 2019:99). Do pogreške uzorkovanja dolazi jer se anketira dio, a ne cijela populacija. Metode uzorkovanja prema vjerojatnosti, gdje je poznata vjerojatnost pojavljivanja svake jedinice u uzorku, omogućuju statističarima da daju izjave o vjerojatnosti o veličini pogreške uzorkovanja. Neprobabilističke metode uzorkovanja, koje se temelje na pogodnosti ili prosudbi, a ne na vjerojatnosti, često se koriste radi prednosti u troškovima i vremenu. Međutim, treba biti krajnje oprezan pri donošenju zaključaka iz uzorka koji se ne temelje na vjerojatnosti, a hoće li uzorak biti i reprezentativan ovisi o prosudbi istraživača koji je dizajnirao istraživački postupak. Osim toga ne postoji objektivna osnova za utvrđivanje granica pogreške uzorkovanja kada je upotrijebljen uzorak koji se ne bazira na vjerojatnosti. Stoga istraživači moraju dobro poznavati metode uzorkovanja kako bi pravilno odgovorili na sve izazove s kojima se susreću u istraživačkoj praksi.

Literatura

1. Andrade, C. (2013). Signal-to-noise ratio, variability, and their relevance in clinical trials. *The Journal of clinical psychiatry*, 74(5), str. 479–481.
2. Andrade, C. (2019). Multiple Testing and Protection Against a Type 1 (False Positive) Error Using the Bonferroni and Hochberg Corrections. *Indian journal of psychological medicine*, 41(1), str. 99–100.
3. Andrade, C. (2020). Sample Size and its Importance in Research. *Indian journal of psychological medicine*, 42(1), str. 102–103.
4. Arslanagić, Š. (2015). Jedna zanimljiva primjena nejednakosti između aritmetičke i geometrijske sredine. *Osječki matematički list*, 15(1), str. 51-59.
5. Delgado-Rodríguez, M. i Llorca, J. (2004). Bias. *Journal of Epidemiology and Community Health*, 58(8), str. 635.
6. Dizdar, D. (2006). *Kvantitativne metode*. Zagreb: Kineziološki fakultet.
7. Gogala, Z. i Šimičević, V. (2005). Korištenje statističkih metoda u hrvatskim poduzećima. *Zbornik Ekonomskog fakulteta u Zagrebu*, 3(1), str. 321-338.
8. Horvat, J. i Mijoč, J (2014). *Osnove statistike*. 2. izdanje, Zagreb: Naklada Ljevak.
9. Horvat, J. i Mijoč, J (2019). *Istraživački SpaSS*. Zagreb: Naknada Ljevak.
10. Jelić, S. (2010). Primjena karakterističnih funkcija u statistici. *Osječki matematički list*, 10(1), str. 85-94.
11. Malhotra, N. K. i Birks, D. F. (2006). *Marketing Research: An Applied Approach*. Harlow: FT/Prentice Hall.
12. Newbold, P. (2010). *Statistika za poslovanje i ekonomiju*. Zagreb: Mate.
13. Norman, G., Monteiro, S. i Salama, S. (2012). Sample size calculations: should the emperor's clothes be off the peg or made to measure?. *BMJ*, 345, e5278.
14. Novoselac, V., i Rimac-Drlje, S. (2014). Svojstva i primjena medijana i aritmetičke sredine, *Osječki matematički list*, 14(1), str. 51-67.
15. Omair, A. (2014). Sample size estimation and sampling techniques for selecting a representative sample. *Saudi Commission Journal of Health Specialties*, 2, str. 142-147.
16. Pavlić, I. (1997). *Statistička teorija i primjena*. Zagreb: Tehnička knjiga.
17. Pedišić, Ž. i Dizdar, D. (2009). *Priručnik za kvantitativne metode*. Zagreb: Kineziološki fakultet.
18. Sarapa, N. (2002). *Teorija vjerojatnosti*. Zagreb: Školska knjiga.

19. Šošić, I. (2006.) *Primijenjena statistika*. Zagreb: Školska knjiga.
20. Tadić, T. (2017). Aritmetička sredina i standardna devijacija, *Poučak*, 18(69), str. 10-24.
21. Taherdoost, H. (2016). Sampling Methods in Research Methodology; How to Choose a Sampling Technique for Research. *International Journal of Academic Research in Management*, 5, 18-27.
22. Vrdoljak, B. (2007). *Vjerojatnost i statistika*. Split: Građevinsko – arhitektonski fakultet Sveučilišta u Splitu.
23. Zenzerović, Z. (2004). *Statistički priručnik*. Rijeka: Pomorski fakultet Sveučilišta Rijeka.

Popis priloga

Slika 1. Primjer programa za generiranje slučajnog odabira random.org (Dostupno na: https://www.random.org , pristupljeno: 19. 9. 2022.)	15
Tablica 1. Prednosti i slabosti tehnika uzorkovanja	18