IX. OCENJIVANJE PARAMETARA[1]

> *Neophodni matematički pojmovi za razumevanje teksta u ovoj glavi:*
>
> ***Osnovni pojmovi teorije verovatnoće . Posebno pojam kvantila.***

Kao što smo u prethodnoj glavi napomenuli, ocenjivanje parametara (engl. *Parameter estimation*) predstavlja jedan od glavnih segmenata statističkog zaključivanja. U psihologiji i srodnim oblastima ovom se segmentu do sada nije poklanjala dovoljna pažnja. Glavni akcent u statističkom zaključivanju bio je na testiranju statističkih hipoteza. S obzirom na probleme koji postoje u klasičnom testiranju statističkih hipoteza, tj. testiranju statističke značajnosti, verujemo da će ocenjivanje parametara sve više biti korišćeno u primeni statističkog zaključivanja u psihologiji i srodnim oblastima.

Statističko ocenjivanje parametara u psihologiji se do sada najčešće koristilo za definisanje normi za psihološke merne instrumente. Radi definisanja normi, psihološki merni instrument se zadaje određenom slučajnom uzorku ispitanika iz populacije kojoj je instrument namenjen. Takav uzorak se uobičajeno zove normativni uzorak. Na osnovu rezultata dobijenih na normativnom uzorku računaju se određene statističke mere (na primer aritmetička sredina, standardna devijacija, percentili) kojima se ocenjuju parametri populacije u kojoj će se dati psihološki merni instrument koristiti. Tako dobijene ocene parametara populacije predstavljaju norme za dati psihološki instrument. Poređenjem rezultata nekog budućeg ispitanika sa normama za dati test, rezultat tog ispitanika u tzv. normativnom merenju dobija svoje značenje.

Parametar predstavlja određenu kvantitativnu karakteristiku populacije ili procesa koji je generisao podatke i najčešće, tj. praktično nikad nam nije poznat. Na primer, parametar može biti proporcija muškaraca u nekoj populaciji, aritmetička sredina ili standardna devijacija inteligencije u nekoj populaciji, veličina efekta anksioznosti ili nespavanja na uspeh u obavljanju određene aktivnosti, veličina razlike između određenih naroda po prosečnoj visini i slično. Znanje vrednosti određenih parametara populacije ili procesa koji je generisao podatke od primarnog je interesa, kako u naučnim istraživanjima tako i u primeni statistike u rešavanju praktičnih problema. Međutim, najčešće nije moguće saznati direktno tačne vrednosti parametara. Čak i onda kada bi to bilo moguće, prikupljanje podataka radi računanja vrednosti parametara bilo bi preskupo, tj. neisplativo Stoga se parametri u principu ocenjuju na osnovu podataka prikupljenih na uzorcima. U istraživanjima sa prebrojavanjem za to se koriste probabilistički uzorci, dok se u analitičkim istraživanjima ocenjivanje parametara često izvodi i korišćenjem neprobabilističkih uzoraka. Budući da je statistička teorija ocenjivanja parametara razvijena pod pretpostavkom da su uzorci

---

[1] Deo teksta u ovoj glavi predstavlja radikalnu modifikaciju teksta koji je preuzet iz poglavlja VI.1. u Tenjović, 2020.

probabilistički i mi ćemo u ovoj glavi pretpostaviti da se ocenjivanje parametara izvodi na osnovu probabilističkih uzoraka. Zapravo, pošto je statističku teoriju ocenjivanja parametara najlakše razumeti ako se pretpostavi da se ocenjivanje parametara izvodi na osnovu jednostavnog slučajnog uzorka, pretpostavićemo da se parametri ocenjuju na osnovu jednostavnih slučajnih uzoraka.

Ocenjivanje parametara može biti tačkasto (engl. *Point estimation*) i intervalno (engl. *Interval estimation*).

1. Tačkasto ocenjivanje parametara

Pri tačkastom ocenjivanju parametar se ocenjuje jednom vrednošću, dok pri intervalnom ocenjivanju parametar ocenjujemo čitavim skupom vrednosti, tj. intervalom vrednosti. Kao ocenitelji parametara služe statistici. Na primer za ocenjivanje aritmetičke sredine populacije najčeše koristimo ocenitelj (engl. *Estimator*) koji zovemo aritmetičkom sredinom uzorka. Dakle, aritmetička sredina kao statistik, definisana opštim obrascem:

$$M = \frac{\sum_{i=1}^{n} X_i}{n}$$

predstavlja ocenitelj μ, tj. aritmetičke sredine populacije ili procesa koji je generisao podatke. Uočimo da u ovom obrascu za aritmetičku sredinu ne figuriraju konkretni podaci ispitanika već slučajne varijable X_i.. Dakle, u skladu sa teoremom koju smo izložili u Glavi III izlažući teoriju slučajnih varijabli, pošto je *M* funkcija slučajnih varijabli, onda i ona sama predstavlja slučajnu varijablu. Po istom principu, svaki statistik koji služi kao ocenitelj parametra posmatramo kao opažljivu slučajnu varijablu.

Kada u istraživanju na podacima sa jednog jednostavnog slučajnog uzorka računamo aritmetičku sredinu sledećim obrascem:

$$M = \frac{\sum_{i=1}^{n} x_i}{n}$$

tada aritmetička sredina kao ocenitelj uzima na slučaj jednu od svojih mogućih vrednosti. Ta vrednost koju aritmetička sredina kao statistik/ocenitelj, tj. kao slučajna varijabla uzima na jednom konkretnom slučajnom uzorku predstavlja *tačkastu ocenu* parametra μ, tj. aritmetičke sredine populacije ili procesa koji je generisao podatke.[2]

U opštem slučaju, ocenitelj (tačnije *tačkasti ocenitelj*) nekog parametra θ predstavlja funkciju G = G (X_1,…, X_n), tj. zavisi od opažljivih slučajnih varijabli, X_1,…, X_n, Kada prikupimo konkretne podatke na slučajnom uzorku, svaka slučajna varijabla X_1,…, X_n uzima na slučaj neku od svojih

[2] Zapravo bi aritmetičku sredinu kao ocenitelj i njenu konkretnu ocenu na slučajnom uzorku trebalo drugačije označavati. Mi to u ovom tekstu nećemo činiti iz čisto parktičnih razloga. Budući da su statistici/ocenitelji teorijski pojam iz konteksta će biti jasno da li se oznaka odnosi na ocenitelj ili na njegovu ocenu na konkretnom uzorku.

mogućih vrednosti. Vrednost koju statistik/ocenitelj uzme na osnovu podataka, tj G (x_1, …, x_n) predstavlja *tačkastu ocenu* parametra.

Metode tačkastog ocenjivanja parametara

Postoji veliki broj metoda tačkastog ocenjivanja parametara: metod momenata, Monte Carlo metod, metod najmanjih kvadrata, metod maksimalne verodostojnosti i drugi.

Opisaćemo samo vrlo grubo metod maksimalne verodostojnosti jer se on najčešće koristi.

Metod maksimalne verodostojnosti (engl. *Maximum likelihood*)

Metod se sastoji u izboru one ocene (kao vrednosti ocenjivanog parametra) koja maksimizuje funkciju L [(x_1, x_2, ...,x_n) | θ]. Dakle, biramo kao ocenu parametra onu **vrednost koja rezultate dobijene na slučajnom uzorku čini najverovatnijim za datu vrednost parametra**. Ocena aritmetičke sredine populacije dobijena ovom metodom definisana je sledećom formulom

$$M = \frac{\sum_{i=1}^{n} x_i}{n}$$

a ocena varijanse populacije formulom:

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - M)^2}{n}$$

Greška uzorka

Greška uzorka je razlika između stvarne vrednosti parametra i vrednosti statistika koju dobijemo na slučajnom uzorku kao ocenu parametra. Tako, ako μ predstavlja aritmetičku sredinu neke varijable u populaciji, a M_i njenu ocenu na osnovu slučajnog uzorka *i* veličine n, tada razlika $g_i = \mu - M_i$ predstavlja grešku tog uzorka *i*. Naravno, kako μ praktično nikada nije poznato, to nismo u stanju ni da odredimo grešku uzorka za neki određeni, konkretni uzorak. Međutim, smisleni zaključci se mogu donositi o verovatnoći pojave greške uzorka određene veličine.

Karakteristike ocenitelja parametra

Različiti statistici/ocenitelji sa različitim uspehom ocenjuju parametar. Dakle, neki od njih pod određenim uslovima imaju poželjnije osobine u pogledu ocenjivanja parametra od drugih.

Sledeće karakteristike ocenitelja parametra služe kao kriterijumi preferencije jednog ocenitelja nad drugim:

1. Nepristrasnost

Statistik/ocenitelj G je nepristrasni ocenitelj parametra ukoliko je očekivana vrednost statistika, $E(G)$, jednaka parametru θ čijem ocenjivanju je statistik namenjen :

$$E(G) = \theta$$

Dakle, statistik/ocenitelj je nepristrasan ocenitelj ako je aritmetička sredina njegove distribucije uzorkovanja jednaka vrednosti parametra θ koja se njime ocenjuje. To praktično znači da ukoliko bismo uzeli beskonačan broj uzoraka i za svaki uzorak izračunali vrednost ovog statistika, aritmetička sredina svih tih vrednosti statistika/ocenitelja bila bi jednaka vrednosti parametra koji se ovim statistikom ocenjuje. Na primer, aritmetička sredina je nepristrasan ocenitelj aritmetičke sredine populacije,[3] dok kod varijanse ova osobina zavisi od toga kako se ovaj ocenitelj varijanse populacije, σ^2, računa. Ako se varijansa kao ocenitelj definiše obrascem:

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - M)^2}{n}$$

tada ona ne predstavlja nepristrasan ocenitelj jer je njena očekivana vrednost, $E(S^2)$:

$$E(S^2) = \sigma^2 \frac{n-1}{n}, \text{ tj. manja od } \sigma^2.$$

Kako bi bila nepristrasni ocenitelj varijanse populacije, varijansa kao ocenitelj mora biti definisana na sledeći način:

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - M)^2}{n-1}$$

Da bi standardna devijacija bila nepristrasni ocenitelj standardne devijacije populacije trebalo bi je računati po obrascu:

$$S' = [1+1/(4n-4)]*S$$

[3] Dokaz za ovo može se videti u Tenjović, 2020, str. 83.

pri čemu je S koren iz varijanse koja je nepristrasna, tj. koja je dobijena po obrascu sa n - 1 u imeniocu.

Medijana je nepristrasni ocenitelj medijane populacije samo ako je distribucija u populaciji varijable za koju računamo medijanu normalna.

2. Konzistentnost

Ocenitelj parametra je konzistentan ako se verovatnoća njegovog približavanja parametru bliži jedinici sa povećanjem veličine uzorka. Formalno se ovo iskazuje na sledeći način:

$$n \to \infty \quad \Rightarrow \quad P(\,|G - \theta| < b\,) \to 1.$$

Dakle, sa povećanjem uzorka konzistentan ocenitelj se gotovo sigurno bliži parametru. Aritmetička sredina je konzistentan ocenitelj aritmetičke sredine populacije.[4] Varijansa je takođe konzistentan ocenitelj varijanse populacije. Medijana je konzistentan ocenitelj aritmetičke sredine populacije samo ako je distribucija varijable u populaciji simetrična. Ako distribucija varijable u populaciji nije simetrična, medijana je konzistentan ocenitelj medijane populacije.

3. Efikasnost

Efikasnost ocenitelja određuje se na osnovu količnika minimalne srednje kvadratne greške i srednje kvadratne greške statistika G čija efikasnost se procenjuje.Srednja kvadratna greška ocenitelja predstavlja očekivanje kvadrirane razilke ocenitelja G i parametra θ koji se ocenjuje tim parametrom. Ako je statistik nepristrasan, **srednja kvadratna greška jednaka je varijansi distribucije uzorkovanja tog statistika**

Dakle, efikasnost ocenitelja G, u oznaci *Ef.(G)*, procenjuje se vrednošću sledećeg količnika:

$$Ef.(G) = \frac{E({G_0}^2 - \theta)^2}{E(G - \theta)^2}$$

pri čemu je G_o ocenitelj koja ima minimalnu srednju kvadratnu grešku.

Ako je *Ef.(G)* = 1, onda je G efikasan ocenitelj parametra θ. Od dva ocenitelja parametra efikasniji je onaj za koji je količnik *Ef.(G)* bliži jedinici.

Jednostavnije rečeno, ocenitelj je efikasan ako malo varira od uzorka do uzorka. Efikasnost ocenitelja u bliskoj je vezi sa preciznošću ocenjivanja parametra. Što je efikasnost ocenitelja bolja ocenjivanje parametra je preciznije.

4. Dovoljnost

Statistik je <u>dovoljna</u> ocena parametra kada koristi sve informacije koje su na raspolaganju u podacima. Na primer, medijana izračunata iz intervalnih podataka ne koristi informacije o veličini

---

[4] Dokaz ovoga može se videti u Tenjović, 2020, str. 83.

razlike među rezultatima, dok aritmetička sredina koristi sve informacije sadržane u takvim podacima. U ovom slučaju aritmetička sredina je dovoljna ocena, a medijana nije. Dakle, dovoljan ocenitelj je najbolji ocenitelj parametra.

Standardna greška ocenitelja

Standardna greška statistika/ocenitelja predstavlja **standardnu devijaciju** njegove **distribucije uzorkovanja**.

Standardna greške za aritmetičku sredinu:

a) Ako je populacija beskonačna ili se uzorkovanje vrši sa vraćanjem, standardna greška za aritmetičku sredinu definiše se na sledeći način:

$$\sigma_M = \frac{\sigma}{\sqrt{n}}$$

pri čemu je σ standardna devijacija populacije, a n veličina uzorka.

b) Ako je reč o konačnoj populaciji veličine N, ili se uzorkovanje odvija bez vraćanja standardna greška za aritmetičku sredinu definisana je sledećim obrascem:

$$\sigma_M = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

Uočimo da je i standardna greška parametar koji ocenjujemo na osnovu statistika, tj. standardne greške dobijene na uzorku. Dakle, ocenitelj standardne greške predstavlja ocenitelj standardne devijacije distribucije uzorkovanja određenog statistika.

Ocenitelj standardne greške za aritmetičku sredinu za beskonačnu populaciju ili pri uzorkovanju sa vraćanjem definiše se ovako:

$$SE_M = \frac{S}{\sqrt{n}}$$

pri čemu je S ocenitelj standardne devijacije populacije, tj. standardna devijacija uzorka a n veličina uzorka.

Za konačnu populaciju ili u slučaju kada uzorak čini frakciju populacije veću od 5% (n/N > 0.05) ocenitelj standardne greške za aritmetičku sredinu definiše se sledećim obrascem:

$$SE_M = \frac{S}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

Kada je n/N < 0.05, tj. kada je uzorak veoma mala **frakcija** populacije, izraz $\sqrt{\frac{N-n}{N-1}}$ u prethodnom obrascu se može zanemariti budući da je u tom slučaju praktično jednak jedinici.

Iz obrazaca za standardnu grešku aritmetičke sredine očigledno je da efikasnost aritmetičke sredine kao ocenitelja zavisi od varijabilnosti pojave kojom se bavimo i veličine uzorka. Pošto na varijabilnost pojave ne možemo da utičemo, efikasnost aritmetičke sredine možemo povećati korišenjem što većeg uzorka. Naime, što je uzorak veći standardna greška za aritmetičku sredinu će biti manja, tj. efikasnost aritmetičke sredine veća. Ali u tome treba biti razuman: ako je varijabilnost pojave iskazana standardnom devijacijom jednaka 30 a uzorak sa 900 ispitanika povećamo na 3600, standardna greška se neće smanjiti 4 puta, već samo 2 puta (u prvom slučaju će biti jednaka 1, a u drugom 0.5). Cena istraživanja na uzorku od 3600 je mnogo veća nego na uzorku od 900 ispitanika te dobitak u efikasnosti ocenitelja ne može da opravda tu cenu.

Jednostavnosti radi, za nekoliko statistika koji se veoma često koriste nećemo prikazati obrasce za njihove standardne greške već samo obrasce za ocenitelje njihovih standardnih grešaka.

Ocenitelj standardne greške za varijansu

Ako je distribucija varijable normalna sa parametrima μ i σ^2 i ako je uzorak veliki, ocenitelj standardne greške za varijansu, u oznaci SE_{S^2}, definiše se sledećom formulom (prema Ahn & Fessler, 2003):

$$SE_{S^2} = S^2 * \sqrt{\frac{2}{(n-1)}}$$

Ocenitelj standardne greške za standardnu devijaciju

Ako je distribucija varijable normalna sa sa parametrima μ i σ^2 i ako je uzorak veliki, standardna greška za standardnu devijaciju može se približno oceniti sledećom formulom (prema Ahn & Fessler, 2003):

$$SE_S \approx \frac{1}{\sqrt{2(n-1)}}$$

Ocenitelji standardnih grešaka za procenat i proporciju:

Ocenitelj standardne greške za procenat P, u oznaci SE_P, definiše se na sledeći način:

$$SE_P = \sqrt{\frac{P(1-P)}{n}}$$

a ocenitelj standardne greške za proporciju p, u oznaci SE_p sledećom formulom:

$$SE_p = \sqrt{\frac{p(1-p)}{n}}$$

Ocenitelj standardne greške za medijanu

Ocenitelj standardne greške za medijanu, u oznaci SE_{Mdn}, za n vrednosti koje su uzorkovane sa vraćanjem iz normalne raspodele sa parametrima μ i σ^2 definisana je sledećim izrazom:

$$SE_{Mdn} = \sqrt{\frac{\pi}{2}\frac{S^2}{n}}$$

Očigledno je da je $SE_{Mdn} = \sqrt{\frac{\pi}{2}\frac{S^2}{n}} = \sqrt{\frac{3.14}{2}} * \frac{S}{\sqrt{n}} = 1.253 * SE_M$.Prema tome, ocena standardne greške za medijanu je 1.25 puta veća od ocene standardne greške za aritmetičku sredinu ako se one ocenjuju na osnovu istog uzorka iz populacije u kojoj se varijabla normalno raspodeljuje.

U opštem slučaju, bez obzira na to kakva je distribucija varijable u populaciji, asimptotska standardna greška za medijanu definiše se sledećom formulom (prema Freund, 1962):

$$\sigma_{Mdn} = \sqrt{\frac{1}{8[f(\mu\delta\nu)]^2\, n}}$$

pri čemu je $\mu\delta\nu$ oznaka za medijanu populacije a $f(\mu\delta\nu)$ vrednost funkcije gustine varijable na medijani populacije. Dakle, standardna greška medijane, pored veličine uzorka, zavisi i od toga kakva je distribucija varijable u populaciji. Ako je raspodela varijable u populaciji normalna, onda je i distribucija uzorkovanja za medijanu normalna. Ali ako raspodela varijable u populaciji nije normalna, distribucija uzorkovanja medijane nije normalna i što je vrednost funkcije gustine na mestu medijane populacije u distribuciji varijable viša to je standardna greška medijane manja. Na primer, za simetričnu Laplasovu distribuciju standardna greška za medijanu je manja od standardne greške za aritmetičku sredinu. Dakle, za ovu distribuciju, medijana je efikasniji ocenitelj od aritmetičke sredine. Ovu formulu ima smisla koristiti za ocenu standardne greške medijane samo na velikim uzorcima.

Ocenitelj standardne greške za koeficijent linearne korelacije

Ako je zajednička distribucija dveju varijabli bivarijatna normalna distribucija, ocenitelj standardne greške za koeficijent linearne korelacije, u oznaci SE_r, definisan je sledećom formulom:

$$SE_r = \frac{1 - r^2}{\sqrt{n}}$$

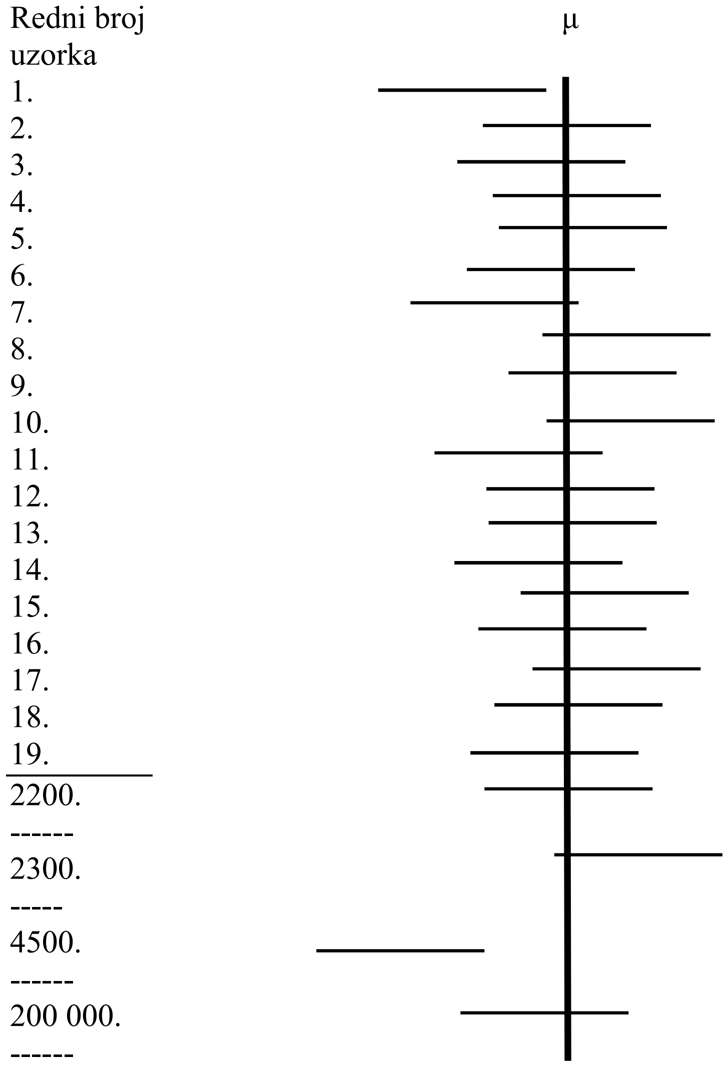
2. Intervalno ocenjivanje parametara – Intervali poverenja

Premda se u dosadašnjoj primeni statistike u psihologiji i srodnim oblastima statističko ocenjivanje parametara znatno manje koristilo od drugih segmenata inferencijalne statistike (testiranja statističke značajnosti ili testiranja statističkih hipoteza), u novije vreme se naglasak pri primeni statistike u ovim oblastima premešta na ocenjivanje parametara i to, pre svega, na tzv. intervalno ocenjivanje parametara. Još se 1999. Radna grupa Američke psihološke asocijacije za statističko zaključivanje svojski zalagala za mnogo veće korišćenje intervala poverenja pri prikazivanju rezultata psiholoških istraživanja i istraživanja u srodnim oblastima (Wilkinson, L., & Task Force on Statistical Inference, APA Board of Scientific Affairs, 1999). Najnoviji trend u ovim oblastima jeste ne samo prikazivanje intervala poverenja pri testiranju statističke značajnosti osnovnih efekata već i intervalno ocenjivanje veličine tih efekata (tome će biti posvećen deo teksta u Glavi **X**.

U intervalnom ocenjivanju parametara, tj. pravljenju intervala poverenja za pojedine parametre, umesto jedne vrednosti, tj. tačkaste ocene vrednosti određenog parametra, parametar se ocenjuje čitavim skupom vrednosti u datom intervalu poverenja. Reč je o postupcima koji ***garantuju da će, ukoliko su ispunjeni određeni uslovi i kada se intervali poverenja prave na određeni način, određeni procenat tih intervala obuhvati parametar.*** Za koliki procenat intervala poverenja možemo da kažemo da će obuhvatiti parametar zavisi od odabranog nivoa rizika ili „nepoverenja", u oznaci α. Najčešće se uzima da je $\alpha = .05$ ili .01. Ukoliko je $\alpha = .05$, tada je reč o konstruisanju 95-postotnog intervala poverenja [95% CI], a ako je $\alpha = .01$, reč je o konstrukciji 99-postotnog intervala poverenja [99% CI].
Dakle, u opštem slučaju zavisno od odabranog nivoa α, konstruisaćemo 100%α α) interval poverenja. Interval poverenja 100α α)% predstavlja interval za nepoznati parametar koji je tako konstruisan na osnovu probabilističkog (slučajnog) uzorka da kada bismo ogroman (u krajnjoj

liniji beskonačan) broj tih intervala konstruisali na isti način na osnovu nezavisnih uzoraka, 100(1- α)% tih intervala bi obuhvatilo parametar, a 100* α% tih intervala ne bi obuhvatilo parametar. Shematski bismo to mogli prikazati kao na Shemi 9.1.

Shema 9.1
Intervali poverenja za ocenjivanje parametra μ na slučajnim uzorcima



Ako pogledamo samo prvih 19 uzoraka, uočićemo da 18 od svih 95-postotnih intervala poverenja (horizontalne podebljane linije) za ovih 19 uzoraka obuhvataju parametar μ (vrednost parametra je fiksna i nalazi se na podebljanoj vertikalnoj liniji). To je u skladu sa onim što predviđa teorija, jer je 18/19 ≈ 0.95. Jedan interval, i to iz prvog uzorka, ne obuhvata parametar. To je isto onako kako predviđa teorija jer je 1/19 ≈ 0.05. Naravno, statistička teorija o intervalima poverenja podrazumeva znatno veći broj uzoraka. Treba, takođe, imati na umu da mi ***realno pravimo interval poverenja samo na jednom jedinom slučajnom uzorku*** a ne na 300 000 uzoraka. Ali iz statističke teorije znamo da od svih 100% intervala koje napravimo u skladu sa određenim pravilima, ogromna većina tih intervala (100* (1 - α)% će obuhvatiti parametar. Stoga verujemo da će i

interval koji napravimo na jednom slučajnom uzorku obuhvatiti parametar, pri čemu smo svesni toga da postoji mogućnost da intervalom koji smo napravili nismo obuhvatili parametar.
Iz sheme je važno uočiti da je parametar μ (vertikalna podebljana linija) konstantna vrednost i da se ne menja od jednog slučajnog uzorka do drugog. Ono što se menja od uzorka do uzorka su donja i gornja granica intervala poverenja (levi i desni krajevi horizontalnih linija). Dakle, u ovom slučaju slučajne varijable su intervali poverenja, tj. njihove donje i gornje granice. Prema tome, kada interval poverenja pravimo samo na jednom konkretnom slučajnom uzorku, nema smisla govoriti o verovatnoći različitoj od 0 i 1 da je parametar obuhvaćen tim konkretnim intervalom. Parametar je u okviru tzv. frekventističke statistike koju ovde prikazujemo konstanta i ne može imati verovatnoću da bude u intervalu poverenja drugačiju od 0 i 1. Parametar jeste ili nije obuhvaćen konkretnim intervalom poverenja. Naime, iz verovatnoće od 95% da intervali velikog broja slučajnih uzoraka obuhvate parametar ne sledi verovatnoća od 95% da konkretni interval iz jednog slučajnog uzorka obuhvati parametar. Ipak, mi uobičajeno verujemo da je parametar obuhvaćen konkretnim intervalom poverenja koji smo napravili na jednom slučajnom uzorku jer je takvih intervala po teoriji mnogo više nego intervala koji ne obuhvataju parametar. Ne verujemo da smo tako „zlehude sreće“ da „naletimo“ baš na vrlo retke uzorke čiji interval poverenja ne obuhvata parametar. Ali treba uvek da budemo svesni da takva mogućnost postoji. Ukoliko baš želimo da tačno odredimo verovatnoću da parametar bude unutar intervala poverenja potrebno je napraviti, postupcima Bajesovske statistike, tzv. kredibilni interval. Naime, parametri u Bajesovskoj statistici su slučajne varijable a ne konstante (o ovome se može pročitati u Morey, Hoekstra, Rouder, Lee, & Wagenmakers, 2016).

Osnovne principe konstrukcije intervala poverenja objasnićemo na primeru pravljenja intervala poverenja za aritmetičku sredinu:

<u>Intervalna ocena aritmetičke sredine populacije - interval poverenja za aritmetičku sredinu:</u>

$100(1-\alpha)\%$ interval poverenja za aritmetičku sredinu konstruiše se tako što se od aritmetičke sredinu uzoka, tj. tačkastog ocenitelja parametra oduzme margina greške i na aritmetičku sredinu uzorka doda margina greške:

$$100(1-\alpha)\%\ CI\ (\mu):\ M \pm \text{margina greške}$$

Donja granica intervala poverenja dobija se oduzimanjem margine greške od aritmetičke sredine uzorka, a gornja granica intervala poverenja tako što se na aritmetičku sredinu uzorka doda margina greške.

Margina greške predstavlja proizvod kvantila $t_{1-\alpha/2}$ iz Studentove raspodele sa n – 1 stepeni slobode i ocene standardne greške za aritmetičku sredinu, SE_M.

Dakle:

$$\mathbf{100(1-\alpha)\%\ CI\ (\mu) : M \pm t_{1-\alpha/2}*SE_M}$$

Izborom nivoa α (rizik da intervalom **ne obuhvatimo** parametar ili nivo „nepoverenja“) definišemo koji interval poverenja pravimo. Najčešće se biraju dva nivoa, .05 i .01:

a) $\alpha = .05$

$100(1 - \alpha)\% = 95\%$

Devedesetpetpostotni interval poverenja, u oznaci 95% CI (μ), konstruiše se na sledeći način:

$$\mathbf{95\%\ CI\ (\mu):\ M \pm t_{0.975} * SE_M}$$

Pri tome, $t_{0.975}$ predstavlja kvantil .975 iz Studentove raspodele sa n – 1 stepeni slobode.

b) $\alpha = 0.01$

$100(1 - \alpha)\%$ je 99%

Devedesetdevetpostotni interval poverenja, u oznaci 99% CI (μ), konstruiše se na sledeći način:

$$\mathbf{99\%\ CI\ (\mu):\ M \pm t_{0.995} * SE_M}$$

Pri tome, $t_{0.995}$ predstavlja kvantil .995 iz Studentove raspodele sa n – 1 stepeni slobode.

Zašto se interval poverenja za aritmetičku sredinu pravi na ovaj način?

1. Polazimo od definicije statistika t:

$t = (M - \mu)/SE_M$

pri čemu je M aritmetička sredina uzorka, μ aritmetička sredina populacije, a SE_M ocenitelj standardne greške za aritmetičku sredinu;

2. Poznato je da statistik t ima Studentovu distribuciju uzorkovanja sa n - 1 stepeni slobode, pri čemu je n veličina slučajnog uzorka.[5]

3. Na osnovu prethodnog, možemo odrediti t-kvantil $t_{1-\alpha/2}$ tako da je:

$P(-t_{1-\alpha/2} \leq t \leq t_{1-\alpha/2}) = 100(1 - \alpha)\%$.

Dakle, budući da je t slučajna varijabla sa Studentovom distribucijom uzorkovanja, kada odaberemo kvantil $t_{1-\alpha/2}$ znamo da je verovatnoća da t bude u intervalu od $-t_{1-\alpha/2}$ do $+ t_{1-\alpha/2}$ jednaka željenoj verovatnoći $100(1 - \alpha)\%$. Koji kvantil ćemo odabrati zavisi od toga kakav interval

---

[5] To sledi iz sledeće teoreme: Ako su M i S aritmetička sredina i standardna devijacija slučajnog uzorka veličine n iz populacije sa normalnom funkcijom gustine $\mathcal{N}(x;\ \mu,\sigma^2)$. tada slučajna varijabla $t = (M - \mu)/(S/\sqrt{n}) = (M - \mu)/SE_M$, ima Studentovu T funkciju gustine sa parametrom $\nu = n – 1$. (cf. Freund, str. 203).

poverenja hoćemo da konstruišemo. Uobičajeno je da se konstruišu 99-postotni ili 95-postotni intervali poverenja tako da je $\alpha = .01$ a $100(1 - \alpha)\% = 99\%$ u prvom slučaju, a $\alpha = .05$ i $100(1 - \alpha)\% = 95\%$ u drugom slučaju.

4. Umesto promenljive t u izrazu $P(-t_{1-\alpha} \leq t \leq t_{1-\alpha}) = 100(1 - \alpha)\%$ stavićemo $(M - \mu)/SE_M$ (na osnovu tačke 1). Sređivanjem ove nejednačine u zagradi dobijamo $100(1 - \alpha)\%$ interval poverenja:

$$P(M - t_{1-\alpha/2} *SE_M \leq \mu \leq M+ SE_M\, t_{1-\alpha/2}) = 100(1 - \alpha)\%.$$

Na osnovu prethodne formule očekujemo da interval poverenja uključi parametar za određeni procenat slučajnih uzoraka, $100(1 - \alpha)\%$, za koje je interval definisan na dati način. Dakle, na ovaj način izraz u zagradi postaje interval poverenja za aritmetičku sredinu. Ako je $\alpha = .05$, $t_{1-\alpha/2} =. t_{0.975}$, a $100(1 - \alpha)\% = 95\%$ te za 95% slučajnih uzoraka očekujemo da intervali poverenja konstruisani na prikazani način obuhvate parametar. Ako je $\alpha = .01$, $t_{1-\alpha/2} = t_{0.995}$, a $100(1 - \alpha)\% = 99\%$ te za 99% slučajnih uzoraka očekujemo da intervali poverenja konstruisani na prikazani način obuhvate parametar. Na osnovu ove opšte formule možemo zaključiti da je verovatnoća $100(1 - \alpha)\%$ da je parametar μ obuhvaćen intervalom poverenja. Ovo je jedini trenutak kada imamo pravo na ovakav zaključak jer su donje i gornje granice intervala poverenja slučajne varijable koje zavise od uzorka do uzorka.

Ključni uslov za primenu prikazanog postupka konstrukcije intervala poverenja za aritmetičku sredinu jeste da varijabla u populaciji iz koje biramo slučajni uzorak ima normalnu raspodelu. Ukoliko raspodela varijable u populaciji nije normalna, za određivanje margine greške na velikim uzorcima preporučuje se umesto kvantila $t_{1-\alpha/2}$ korišćenje kvantila $z_{1-\alpha/2}$ iz standardizovane normalne raspodele. Nije sasvim jasno da li je ovaj predlog neophodno rešenje problema nenormalnosti distribucije populacije za velike uzorke jer sa povećanjem veličine uzorka, tj. broja stepeni slobode Studentova distribucija se približava standardizovanoj normalnoj.

Konstrukciju intervala poverenja za aritmetičku sredinu prikazaćemo na jednom primeru:

U istraživanju Tenjovića, Petrovića i Srne o blagostanju mladih u doba pandemije virusa SARS-CoV-2, koje je izvedeno 2021. godine na uzorku od 1921 srednjoškolca u Srbiji, dobijeni su sledeći rezultati za varijablu sreća koja je merena supskalom H upitnika EPOCH (Kern et al., 2016):

$M = 4.02$, $S = .93$, $SE_M = 0.02$. Skala za merenje sreće bila je u rasponu od 1 do 5.

Budući da je pretpostavka normalnosti raspodele varijable sreće u populaciji srednjoškolaca neplauzibilna, intervale poverenja smo konstruisali i na način koji smo podrobno opisali u prethodnom tekstu, tj. korišćenjem kvantila iz Studentove raspodele sa 1920 stepeni slobode ali i korišćenjem kvantila iz standardizovane normalne raspodele.

1. Korišćenjem kvantila $t_{0.975}$ (za 95%CI), odnosno $t_{0.995}$ (za 99%CI) iz Studentove raspodele dobijeni su sledeći intervali:

   95%CI: $4.02 \pm 1.9612 * 0.02 \rightarrow$ [95% CI: 3.98 ; 4.06]
   99%CI: $4.02 \pm 2.5784 * 0.02 \rightarrow$ [99% CI: 3.97 ; 4.07]

2. Korišćenjem kvantila $z_{0.975}$ (za 95%CI), odnosno $z_{0.995}$ (za 99%CI) iz standardizovane normalne raspodele dobijeni su sledeći intervali:
   95%CI: $4.02 \pm 1.96 * 0.02 \rightarrow$ [95% CI: 3.98 ; 4.06]
   99%CI: $4.02 \pm 2.5758 * 0.02 \rightarrow$ [99% CI: 3.97 ; 4.07]

Dakle, oba postupka daju praktično iste rezultate, što se moglo i predvideti s obzirom na veličinu uzorka. Šta na osnovu intervala poverenja napravljenog na jednom slučajnom uzorku možemo zaključiti? Tačkasta ocena parametra je 4.02. Intervali poverenja su veoma uski i verujemo da smo njima obuhvatili parametar. Naša vera u to je jača za 99%CI nego za 95%CI, ali je prvi interval malko širi. Naravno, pri tome smo svesni i mogućnosti da ovim intervalima nismo obuhvatili parametar. Ono što nikako ne možemo zaključiti jeste da je verovatnoća 95% (u slučaju 95%CI), odnosno 99% (u slučaju 99%CI) da se parametar nalazi u intervalu. Naime, ***kada napravimo konkretni interval poverenja onda parametar jeste ili nije u tom intervalu, nema nikakvog smisla govoriti o verovatnoći da je parametar u tom konkretnom intervalu*** jer brojke koje dobijemo su konstante a ne slučajne varijable. O verovatnoći da interval poverenja obuhvati parametar ima smisla da govorimo samo kada su donja i gornja granica intervala slučajne varijable, tj. u obliku $M - t_{1-\alpha/2} * SE_M$ i $M + SE_M\, t_{1-\alpha/2}$ jer su u tom slučaju i M i SE_M slučajne varijable. Interval poverenja je zapravo dobra mera preciznosti istraživanja: što je interval poverenja uži veća je preciznost ocenjivanja parametra, tj. samog istraživanja. Intervali poverenja u prikazanom primeru su veoma uzani i govore u prilog velikoj preciznosti ocenjivanja parametra, tj. aritmetičke sredine sreće srednjoškolaca u Srbiji.

Na sličan način kao za aritmetičku sredinu, mogu se konstruisati intervali poverenja za mnoge druge parametre. Navešćemo ovde samo nekoliko primera:

<u>Interval poverenja za proporciju (π) za velike uzorke</u>:

$$100(1-\alpha)\%\mathrm{CI}\ (\pi)\colon p \pm z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$$

Pri tome, π je proporcija u populaciji, p je ocenitelj proporcije u populaciji, tj. proporcija uzorka, n je veličina uzorka a $z_{1-\alpha/2}$ je kvantil $_{1-\alpha/2}$ iz standardizovane normalne raspodele.

Interval poverenja za procenat (Π) za velike uzorke:

$$100(1-\alpha)\%\text{CI }(\Pi): P \pm z_{1-\frac{\alpha}{2}}\sqrt{\frac{P(1-P)}{n}}$$

Pri tome, Π je procenat u populaciji, P je ocenitelj procenta u populaciji, tj. procenat uzorka, n je veličina uzorka a $z_{1-\alpha/2}$ je kvantil $_{1-\alpha/2}$ iz standardizovane normalne raspodele.

Aproximativni interval poverenja za standardnu devijaciju ako je distribucija varijable normalna:

$$100(1-\alpha)\%CI(\sigma): Donja\ granica: \frac{\sqrt{(n-1)\,S^2}}{\sqrt{\chi^2_{\alpha/2}}}\ ; Gornja\ granica: \frac{\sqrt{(n-1)\,S^2}}{\sqrt{\chi^2_{1-\alpha/2}}}$$

Pri tome, S^2 je varijansa uzorka, $\chi^2_{\alpha/2}$ je kvantil α/2 iz hi-kvadrat raspodele sa n-1 stepeni slobode, a $\chi^2_{1-\alpha/2}$ je kvantil 1-α/2 iz hi-kvadrat raspodele sa n-1 stepeni slobode.

Napomena: za statistike za koje nije analitički izvedena distribucija uzorkovanja i standardna greška, do distribucije uzorkovanja, standardne greške i intervala poverenja može se doći postupkom samouzorkovanja (engl. *Bootstrapping*). Ovaj postupak izvodi se tako što se iz jednog slučajnog uzorka kao reprezenta populacije kompjuterski, uzorkovanjem sa vraćanjem, formira veliki broj uzoraka (obično je dovoljno 5000 do 10000 uzoraka), na osnovu kojih se ocenjuju distribucija uzorkovanja određenog statistika, njegova standardna greška i intervali poverenja za parametar koji se tim statistikom ocenjuje.

Zašto je za ocenjivanje parametara bitna veličina uzorka: zakon velikih brojeva

Zakon velikih brojeva predstavlja jednu od osnovnih teorema teorije verovatnoće koji su prvo formulisali, dokazali i opisali čuveni matematičari Jakob Bernuli i Simeon Poason. U najbazičnijem obliku ovaj zakon tvrdi da se relativna učestalost slučajnog događaja približava verovatnoći tog događaja sa povećanjem broja nezavisnih ponavljanja slučajnog eksperimenta. U vezi sa temom ovog poglavlja zakon velikih brojeva govori o tome da će se verovatnoća razlikovanja ocene aritmetičke sredine populacije dobijene na osnovu uzorka od parametra, tj. aritmetičke sredine populacije težiti nuli sa povećanjem veličine uzorka. Dakle, ako malo slobodnije tumačimo stroge formulacije zakona velikih brojeva rekli bismo da za ocenjivanje parametara važi jednostavno pravilo: što je veći uzorak koji koristimo za ocenjivanje parametara to bolje. Ali u primeni ovog pravila treba voditi računa o tome da dalji dobitak u preciznosti ocenitelja kada uzorak već dostigne određenu veličinu nije toliko veliki da bi opravdao trošenje novčanih i drugih resursa.

Apriorna analiza preciznosti ocenitelja parametra – određivanje neophodne veličine uzorka

U novije vreme razvijaju se postupci za određivanje neophodne veličine uzorka zavisno od toga koliko istraživač želi da ocena parametra bude blizu parametru koji se ocenjuje. Ovaj problem rešava se pod imenom „preciznost uzorkovanja" (engl. Sampling precision). Dakle, istraživač u rešavanju problema neophodne veličine uzorka za ocenjivanje nekog parametra, treba ***pre nego što krene u istraživanje*** da definiše dve stvari: 1. koliko želi da vrednost statistika-ocenitelja bude blizu parametru u jedinicama standardne devijacije populacije, i 2. koliko poverenja želi da ima u to da će vrednost statistika-ocenitelja biti toliko blizu parametru. Kada se ove dve stvari definišu postoji formula koja je razvijena u okviru ovog postupka na osnovu koje možemo odrediti potrebnu veličinu uzorka (Trafimov & MacDonald, 2017, Trafimov, 2018):

Na primer, ako je reč o ocenjivanju aritmetičke sredine populacije a istraživač želi da aritmetička sredina uzorka bude blizu parametru za određeni deo standardne devijacije populacije može da primeni sledeću formulu kako bi sračunao neophodnu veličinu uzorka, u oznaci n_p:

$$n_p = \left(\frac{z_c}{f}\right)^2$$

Pri tome z_c je kvantil standardizovane normalne raspodele koji odgovara verovatnoći koju je istraživač odabrao zavisno od toga koliko poverenja želi da ima u to da će vrednost aritmetičke sredine uzorka biti blizu parametru koji se ocenjuje, tj. u ovom slučaju aritmetičke sredine populacije, a f je frakcija, tj. deo standardne devijacije populacije, zavisno od toga koliko blizu istraživač želi da aritmetička sredina uzorka bude blizu aritmetičkoj sredini populacije.

Na primer, ako istraživač želi da aritmetička sredina uzorka bude za 1 deseti deo standardne devijacije populacije blizu aritmetičkoj sredini populacije sa verovatnoćom od 99% tada je:

$$n_p = \left(\frac{2.58}{0.1}\right)^2 = 665.64 \approx 666$$

Dakle, istraživaču je u ovom slučaju potreban uzorak od najmanje 666 ispitanika.

Premda postupak podrazumeva da je raspodela u populaciji normalna, Trafimov je simulacijom pokazao da je postupak prilično robustan na neispunjenost ovog uslova. Vrlo slični rezultati dobijaju se i kada je distribucija u populaciji normalna i kada nije normalna.

Reference:

Ahn, S., & Hessler, J. A. (2003). Standard Errors of Mean, Variance, and Standard Deviation Estimators. Skinuto 27.12. 2009. sa https://web.eecs.umich.edu/~fessler/papers/files/tr/stderr.pdf

Freund, J. F. (1963). *Mathematical statistics*. Englewood Cliffs, N.J.: Prentice-Hall, Inc.

Kern, M. L., Benson, L., Steinberg, E. A., & Steinberg, L. (2016). The EPOCH Measure of Adolescent Well-Being. *Psychological Assessment*, *28*(5), 586–597. http://dx.doi.org/10.1037/pas0000201

Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., & Wagenmakers, E.-J. (2016). The fallacy of placing confidence in confidence intervals. *Psychonomic Bulletin & Review*, *23*(1), 103–123. http://dx.doi.org/10.3758/s13423-015-0947-8.

Trafimow, D., & MacDonald, J. A. (2017). Performing inferential statistics prior to data collection. *Educational and Psychological Measurement*, *77*, 204–219. https://doi.org/10.1177/0013164416659745

Trafimow, D. (2018). Confidence intervals, precision and confounding. *New Ideas in Psychology*, *50*, 48–53. https://doi.org/10.1016/j.newideapsych.2018.04.005

Wilkinson, L., & Task Force on Statistical Inference, APA Board of Scientific Affairs. (1999). Statistical methods in psychology journals. *American Psychologist*, *54*, 594–604.