MODELI PREKIDNIH RASPOREDA VJEROVATNOĆE

Model rasporeda vjerovatnoće ćemo formulisati u vidu opšteg algebarskog izraza, tj. formule koja daje vezu između pojedinih vrijednosti koje uzima slučajna promjenljiva *X* i odgovarajućih vjerovatnoća.

Takva funkcionalna veza:

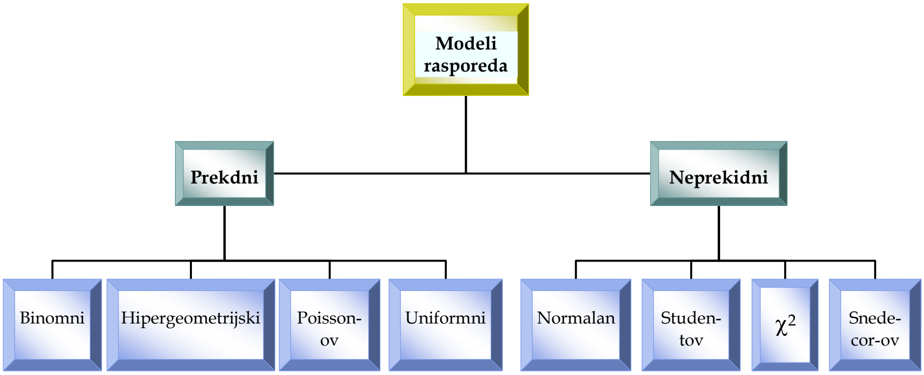
$$p_i = f\left(x_i\right) \qquad i = 1, 2, \ldots, n,$$

na kondenzovan način sadrži cijeli raspored vjerovatnoće i naziva se model rasporeda.

Osnovni značaj modela rasporeda je u tome što se na njima bazira cjelokupno statističko zaključivanje.

Najpoznatiji modeli prekidnih rasporeda vjerovatnoće su: binomni, hipergeometrijski, Poisson-ov i uniformni.

U okviru modela neprekidnih rasporeda vjerovatnoće, u statističkoj praksi najčešće se koriste: normalan, Student-ov, χ^2 (hi kvadrat) i Snedecor-ov (Fisher-ov) raspored.



Binomni raspored

Najčešće korišćeni prekidni raspored u primijenjenoj statistici jeste binomni raspored.

Posmatrajmo sukcesivno izvođenje nekog eksperimenta u kome se jedno izvođenje naziva opit. Pretpostavimo da su u svakom pojedinom opitu moguća samo dva, uzajamno isključiva, ishoda, koje ćemo arbitrarno nazvati uspjeh i neuspjeh. Ova dva termina (uspjeh i neuspjeh) se koriste samo radi razgraničavanja dva moguća ishoda i ne označavaju "kvalitet" nekog događaja.

Jacob Bernoulli (1654 -1705)

Takav opit koji može produkovati samo dva ishoda naziva se Bernoulli-jev opit, po švajcarskom matematičaru J. Bernoulli-u, koji je dao veliki doprinos teoriji vjerovatnoće. Nas, međutim, neće interesovati samo jedan Bernoulli-jev opit, već niz nezavisnih, ponovljenih Bernoulli-jevih opita.

Takav niz opita naziva se Bernoulli-jev proces, ukoliko su ispunjeni sljedeći uslovi:

(1) Svaki opit rezultira u jednom od dva moguća ishoda, koje tehnički klasifikujemo kao uspjeh (U) i neuspjeh (N).

(2) Vjerovatnoća uspjeha, $p = P(U)$, konstantna je od opita do opita.

Vjerovatnoća neuspjeha, $P(N) = 1 - p$, označava se sa q, tako da je $p + q = 1$.

(3) Opiti su nezavisni, odnosno bilo koji ishod da se realizuje u nekom opitu neće imati uticaja na vjerovatnoću ishoda u bilo kom drugom opitu.

Ukoliko izvršimo n ponovljenih Bernoulli-jevih opita, tada broj uspjeha može iznositi $0, 1, 2, \ldots, n$.

Cilj nam je da odredimo formulu za izračunavanje vjerovatnoće ostvarivanja svakog mogućeg broja uspjeha unutar n nezavisnih, uzastopno ponovljenih opita, koji formiraju Bernoulli-jev proces.

Raspored do koga dolazimo na ovakav način i koji nam daje željene vjerovatnoće naziva se binomni raspored.

Binomni raspored

$$P(x) = \binom{n}{x} p^x (1-p)^{n-x} \quad \text{za } x = 0, 1, 2, \ldots, n$$

n = broj opita

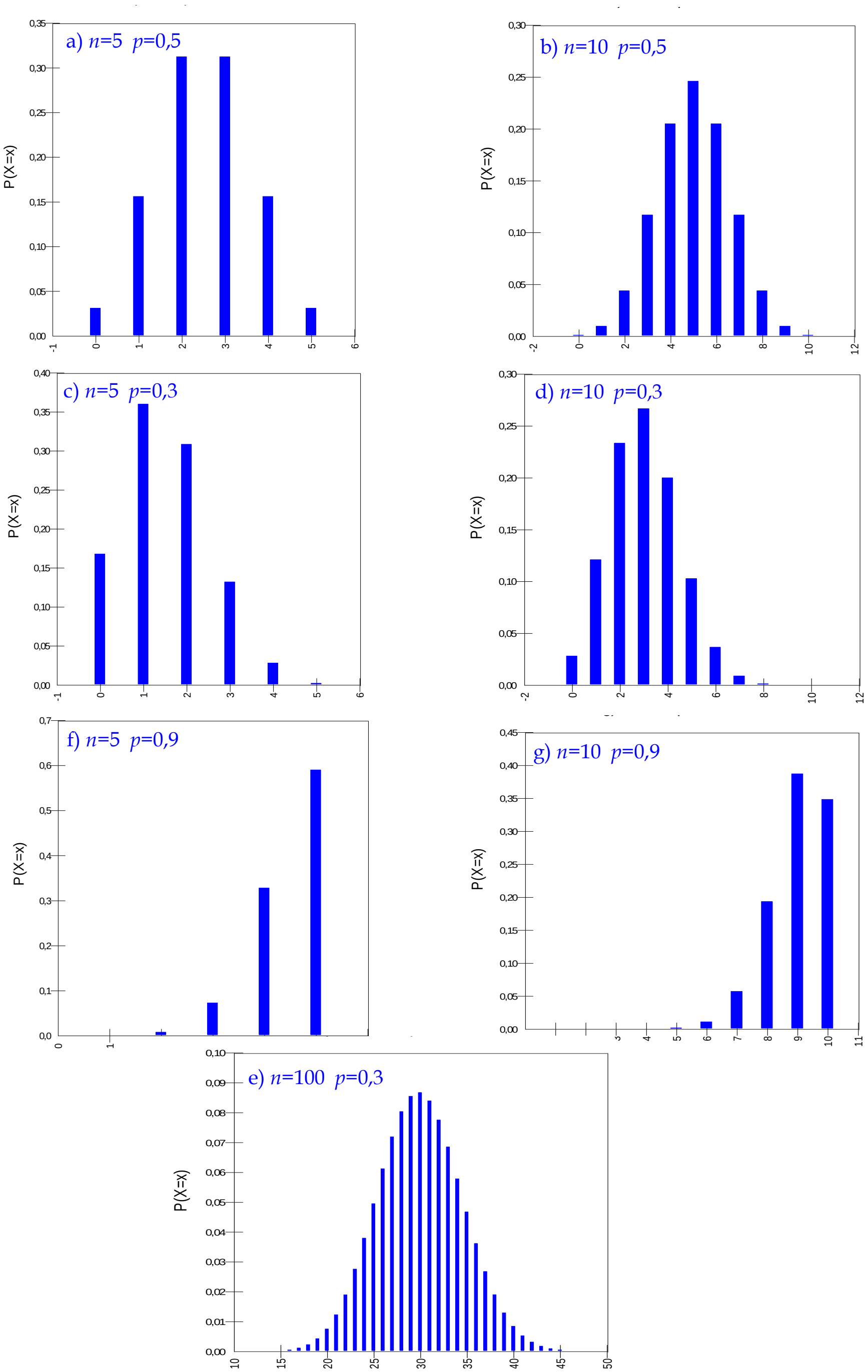
p = vjerovatnoća "uspjeha" u svakom opitu

Binomni raspored ima dva parametra, n i p. Znači, termin binomni raspored se odnosi na čitavu familiju binomnih rasporeda, od kojih svaki zavisi od konkretnih vrijednosti parametara n i p. Da bismo sagledali kako vrijednosti parametara utiču na oblik binomnog rasporeda, na Slici 4.6 grafički ćemo prikazati nekoliko različitih binomnih rasporeda.

Na osnovu Slike 4.6 zaključujemo da je **binomni raspored simetričan kada je p = 0,5, bez obzira na vrijednost n. Kada je p < 0,5 raspored je asimetričan udesno, a u slučaju kada je p > 0,5 raspored je asimetričan ulijevo.**

Izuzetno važna osobina binomnog rasporeda može se sagledati sa grafika c, d i e: ukoliko se parametar p ne razlikuje mnogo od 0,5, sa povećanjem n (tj. broja opita) raspored postaje sve više simetričan (dobija zvonasti oblik).

Međutim, kada se p znatno razlikuje od 0,5, kao na graficima f i g, takva tendencija se ne može uočiti.



Slika 4.6 Binomni rasporedi za odabrane vrijednosti p i n

Sa povećanjem vrijednosti n izračunavanje binomnih vjerovatnoća na osnovu formule 4.5 postaje zamorno. Zbog toga su, korišćenjem računara, statističari konstruisali specijalne tablice (kao što je Tablica 1 u prilogu knjige), koje nam daju pregled izračunatih vjerovatnoća za razne vrijednosti n, p i x.

Tako, na primjer, za $n = 5$, $p = 0{,}5$ i $x = 2$, u Tablici 1 možemo pronaći vrijednost 0,3125, tj. 0,3125, što predstavlja vjerovatnoću realizacije 2 uspjeha u eksperimentu koji se sastoji iz pet uzastopnih opita sa vjerovatnoćom uspjeha u svakom opitu 0,5.

Da bismo na sintetički način sagledali najvažnije aspekte binomnog rasporeda potrebno je odrediti srednju vrijednost i mjeru disperzije binomne slučajne promjenljive.

Aritmetička sredina (preciznije rečeno, očekivana vrijednost), varijansa i standardna devijacija binomnog rasporeda dobijaju se na osnovu sljedećih formula:

Aritmetička sredina $\mu_X = E(X) = np$

Varijansa $\sigma_X^2 = np(1-p) = npq$

Standardna devijacija $\sigma_X = \sqrt{npq}$

Poisson-ov raspored

Kao što smo vidjeli, binomni raspored pokazuje vjerovatnoće dobijanja x uspjeha prilikom n opita nekog eksperimenta, tj. Bernoulli-jevog procesa, pod uslovom da je vjerovatnoća uspjeha p konstantna iz opita u opit.

Ukoliko je vjerovatnoća uspjeha p veoma mala (najčešće se uzima da je $p \leq 0{,}05$) i kada je $n \geq 20$, umjesto binomnog modela možemo koristiti Poisson-ov model, kao zadovoljavajući način aproksimiranja vjerovatnoća.

Poisson-ov raspored nam ne služi samo za aproksimiranje binomnih vjerovatnoća. Pomoću njega možemo opisivati veliki broj pojava, bilo u vremenu ili prostoru.

Poisson-ov raspored se može koristiti da bi se odredila vjerovatnoća broja javljanja nekog događaja u jedinici vremena ili prostora.

Potrebno je, međutim, da su ispunjena **sljedeća tri uslova:**

1. Broj javljanja događaja je nezavisan od jedne do druge jedinice vremena ili prostora.
2. Vjerovatnoća javljanja nekog događaja je proporcionalna dužini određene jedinice vremena ili prostora.
3. Vjerovatnoća istovremenog javljanja dva ili više događaja u sasvim maloj jedinici vremena ili prostora je zanemarljivo mala.

Pod ovim uslovima, **raspored slučajne promjenljive *X*, koja pokazuje koliko puta se javio neki događaj u jedinici vremena ili prostora, ima oblik Poisson-ovog modela** (nazvan je po francuskom matematičaru S. D. Poisson-u, koji je 1837. objavio članak u kome se prvi put opisuje ovaj raspored).

Poisson-ov raspored

$$P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$x = 0, 1, 2, \ldots \qquad \lambda > 0$

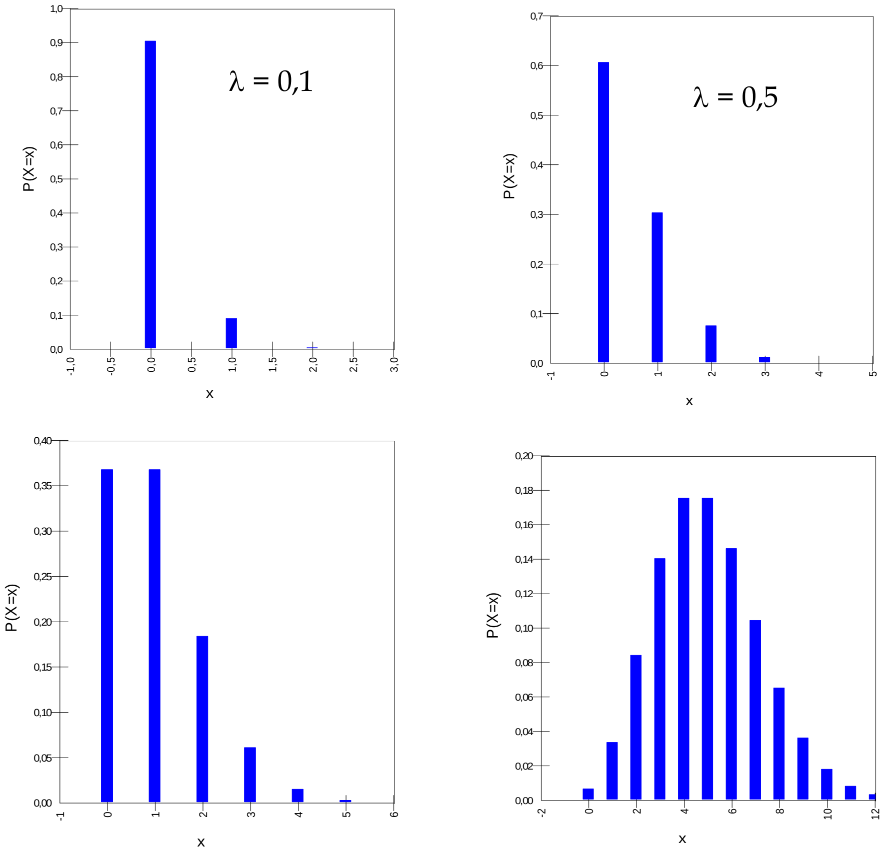
Simeon Poisson
(1781 -1840)

Ova formula reprezentuje vjerovatnoću da prekidna slučajna promjenljiva X uzme neku proizvoljnu vrijednost x. Broj e predstavlja bazu prirodnog logaritma i iznosi (zaokružen na pet decimala) 2,71828. Kao što vidimo, Poisson-ov raspored ima samo jedan parametar λ. Numerički, λ predstavlja prosječan broj javljanja nekog događaja u jedinici vremena ili prostora.

Zanimljivo je da parametar λ istovremeno predstavlja aritmetičku sredinu i varijansu Poisson-ovog rasporeda, tj:

$$E(X) = \sigma_X^2 = \lambda .$$

Na osnovu Slike 4.8 **možemo sagledati da je Poisson-ov raspored asimetričan udesno, jer slučajna promjenljiva *X* ne može uzeti vrijednosti manje od nule, i da sa povećanjem vrijednosti λ raspored teži da zauzme simetričan oblik.**



Slika 4.8 Različiti oblici Poisson-ovog rasporeda u zavisnosti od parametra λ

λ = 1 λ = 5

Da bismo sagledali kako se pomoću Poisson-ovog rasporeda aproksimiraju binomne vjerovatnoće (kada je p blisko nuli, a $n \geq 20$), posmatrajmo proizvodni proces izvjesnih električnih komponenti i pretpostavimo da je 2% proizvoda neispravno. Ako se uzme uzorak od $n = 100$ elemenata, kolika je vjerovatnoća da će on sadržavati 3 neispravna proizvoda? Veza između binomnih i Poisson-ovih vjerovatnoća data je sa:

$$\binom{n}{x} p^x (1-p)^{n-x} \approx \frac{e^{-np}(np)^x}{x!}.$$

odnosno, kada se Poisson-ov raspored uzima kao aproksimacija binomnog, za vrijednost parametra λ uzima se np, pošto je np aritmetička sredina binomnog rasporeda. U našem primjeru je $n = 100$, $p = 0{,}02$ i $\lambda = np = 2$, pa Poisson-ova vjerovatnoća iznosi:

$$P(X=3) = \frac{e^{-2}2^3}{3!} = 0{,}1804.$$

Da smo direktno koristili formulu za binomni raspored dobili bismo:

$$P(X=3) = \binom{100}{3} 0{,}02^3 \cdot 0{,}989^7 = 0{,}1823,$$

što je veoma blisko aproksimativnoj vrijednosti, a znatno se teže izračunava.

Uniformni raspored

Jedan od najjednostavnijih prekidnih rasporeda, koji je posebnu primjenu našao u kompjuterskoj simulaciji i igrama na sreću, jeste uniformni raspored. Njegova osnovna karakteristika je da svaka vrijednost slučajne promjenljive ima jednaku vjerovatnoću da se ostvari.

Pretpostavimo da slučajna promjenljiva X može uzeti jednu od vrijednosti 1, 2, ... , n.

Ako je vjerovatnoća da X uzme bilo koju od ovih vrijednosti jednaka i iznosi $1/n$, i ako se ove vjerovatnoće ne mijenjaju od opita do opita, za X kažemo da ima prekidni uniformni raspored.

Uniformni raspored

$$P(X = x) = \frac{1}{n}$$

x = 1, 2, ... , n

n = broj različitih vrijednosti koje X može uzeti

Aritmetička sredina i varijansa uniformnog rasporeda date su sljedećim formulama:

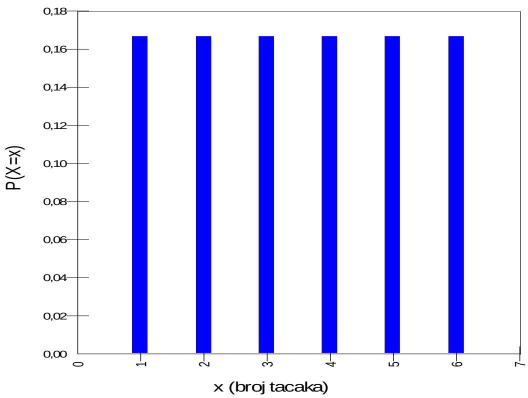
$$E(X) = \frac{n+1}{2}$$

$$\sigma_X^2 = \frac{n^2 - 1}{12}.$$

Uniformni raspored smo već imali prilike da upoznamo u eksperimentu sa bacanjem pravilne kocke (u Tabeli 4.5). Pošto svaka strana kocke ima podjednaku mogućnost pojavljivanja, slučajna promjenljiva, koja pokazuje broj tačaka na kocki, ima uniformni raspored:

$$P(X = x) = \frac{1}{6} \quad x = 1, 2, \ldots, 6.$$

Ovaj raspored možemo grafički ilustrovati Slikom 4.9.



Slika 4.9 Dijagram vjerovatnoće u eksperimentu sa bacanjem pravilne kocke

UZORKOVANJE

Razlozi za uzorkovanje
Okvir uzorka
Slučajno i neslučajno uzorkovanje
Uzorkovanje sa i bez ponavljanja
Jednostavno slučajno uzorkovanje
Stratifikovano slučajno uzorkovanje
Sistematsko uzorkovanje
Klaster uzorkovanje
Neslučajno – namjerno uzorkovanje
Uzoračka greška
Neuzoračke greške

CILJEVI POGLAVLJA

Nakon čitanja ovog poglavlja bićete u stanju da:

1. shvatite zašto je u mnogim situacijama uzorkovanje jedini podesan način da se dođe do informacije o parametru skupa
2. shvatite termin okvir uzorka
3. razumijete razliku između slučajnog i neslučajnog uzorkovanja
4. razumijete razliku između uzoraka sa i bez ponavljanja
5. odaberete slučajan uzorak korišćenjem Tablice slučajnih brojeva
6. razumijete razlike između različitih vrsta uzoraka

Vrijednosti parametara populacije određuju se u cilju spoznaje njenih karakteristika i kvantitativne i kvalitativne deskripcije.

Određivanje vrijednosti parametara osnovnih skupova može se izvesti u osnovi na dva načina:

- prvi podrazumijeva cjelovito posmatranje populacije (popis), odnosno mjerenje/određivanje vrijednosti posmatrane karakteristike za svaki elemenat osnovnog skupa,

- drugi način podrazumijeva da se vrijednost parametara osnovnog skupa u kvantitativnom smislu određuje na osnovu mjerenja samo jednog dijela osnovnog skupa, koji se naziva uzorak; postupak uzimanja uzorka naziva se uzorkovanje.

Svaki od pomenutih načina ima svoje prednosti i nedostatke.

RAZLOZI ZA UZORKOVANJE

Postoji nekoliko važnih razloga zbog kojih je analiza na osnovu uzoraka pogodnija u odnosu na cjelovito posmatranje populacije:

1. Uzorkovanje omogućuje da se **do informacija o parametrima osnovnog skupa dođe uz znatno niže troškove**.
2. **Uzorkovanjem se znatno brže dolazi do informacija o osnovnom skupu,** posebno u slučajevima velikih osnovnih skupova, kao i kod skupova sa beskonačno velikim brojem elemenata.
3. S obzirom na raspoložive izvore, kako u finansijskom, tako i u pogledu dostupnosti elemenata populacije, **uzorak može biti jedini način postizanja ciljeva istraživanja.**
4. Istraživački proces često podrazumijeva destrukciju, oštećenja ili uništavanje proizvoda, tako da se **uzorkovanjem mogu ostvariti značajne uštede.**
5. Ako je obuhvatanje cjelovitih populacija nemoguće (zbog toga jer su izuzetno velike ili, čak, imaju beskonačan broj elemenata, ili je, jednostavno, izuzetno teško njihovo obuhvatanje), tada je **zaključivanje na osnovu uzoraka jedino moguće rješenje.**

OKVIR UZORKA

Uzorak se bira iz liste populacije, nekog spiska, mape, direktorija ili drugih izvora koji predstavljaju populaciju. Ove liste, spiskovi, mape, direktoriji predstavljaju okvir uzorka.

U najvećem broju slučajeva uzorkovanje se bazira na okviru uzorka. **Teorijski, okvir uzorka i ciljna populacija su jednaki**.

Okvir uzorka za neke populacije ustanovljen je po prirodi stvari. **U slučaju postojanja okvira uzorka, proces uzorkovanja uveliko je olakšan, što podrazumijeva efikasnije odabiranje elemenata u uzorak, odnosno omogućava brži izbor elemenata u uzorak, uz niže troškove izbora.**

Ukoliko ne postoji okvir uzorka, tada je značajno otežano svako ozbiljnije istraživanje, a naročito istraživanja populacija sa velikom brojem elemenata.

SLUČAJNO I NESLUČAJNO UZORKOVANJE

Postoje dva osnovna tipa uzorkovanja: **slučajno i neslučajno**.

Kod slučajnog izbora svaka jedinica populacije ima poznatu vjerovatnoću, odnosno šansu da bude izabrana u uzorak. **Slučajno uzorkovanje podrazumijeva da se vjerovatnoća izbora uzima u obzir prilikom selekcije.**

U neslučajnom uzorkovanju jedinice populacije nemaju poznatu vjerovatnoću izbora u uzorak. **To znači da elementi uzorka nisu birani slučajno,** što podrazumijeva da namjera njihovog izbora u uzorak može biti posljedica različitih okolnosti.

Kod neslučajnog izbora nije moguće određivanje vjerovatnoće pojave određenih jedinica.

Statistički metodi i modeli, najvećim dijelom, bazirani su na pretpostavci o slučajnom uzorkovanju.

Postoje četiri osnovne tehnike slučajnog uzorkovanja:

- jednostavno slučajno uzorkovanje ili, kako je uobičajeno da se kaže, **izbor prostog slučajnog uzorka,**
- stratifikovano slučajno uzorkovanje,
- sistematsko slučajno uzorkovanje, i
- klaster/**area** uzorkovanje.

Svaki od pomenutih načina ima svoje prednosti i nedostatke.

UZORKOVANJE SA I BEZ PONAVLJANJA

Uzorkovanje može da se izvede sa i bez ponavljanja elemenata.

Uzorkovanje sa ponavljanjem **je procedura izbora elemenata u uzorak tako da se u svakom pojedinom opitu (pojedinačno izabiranje elementa u uzorak) ne mijenjaju uslovi izbora.**

Ovo je veoma značajno sa stanovišta vjerovatnoće izbora pojedinog elementa u uzorak i podrazumijeva da se prilikom svakog pojedinog izvlačenja ne mijenja vjerovatnoća izbora elemenata u uzorak.

Uzorkovanje bez ponavljanja je procedura izbora elemenata u uzorak u kojoj se jednom izabrani elemenat ne vraća ponovno u osnovni skup. To znači da se izbor svakog novog elementa u uzorak vrši u promijenjenim uslovima, što dalje implicira da se prilikom svakog pojedinog izvlačenja mijenja vjerovatnoća izbora elemenata u uzorak.

Što je osnovni skup sa konačnim brojem elemenata manji, to su ove razlike veće. Zbog toga se **u metodima statističkog zaključivanja vodi računa o tome da li je izbor sa ili bez ponavljanja.**

JEDNOSTAVNO SLUČAJNO UZORKOVANJE

Ovo je, svakako, najčešći način u okviru slučajnog uzorkovanja. Jednostavno slučajno uzorkovanje može se posmatrati i kao osnova za ostale tehnike uzorkovanja.

U jednostavnom slučajnom uzorkovanju svaka jedinica okvira uzorka numerisana je od 1 do *N* (sa *N* je označen broj elemenata populacije). U daljem postupku koristi se Tablica slučajnih brojeva ili neki generator slučajnih brojeva za izbor n elemenata u uzorak.

U praksi, **Tablica slučajnih brojeva primjenjuje se u uslovima kada su svi elementi populacije numerisani.**

Numeracija se izvodi u obliku dvocifrenih, trocifrenih i višecifrenih brojeva, u zavisnosti od veličine skupa.

Određivanje početnih slučajnih brojeva vrši se na različite načine. **Jedan, koji se predlaže**, sastoji se u tome **da se olovka zavrti na samoj tablici slučajnih brojeva** i tamo gdje se zaustavi vrh olovke, tamo je početak izbora.

Drugi način je izvlačenjem iz šešira, u kome su cifre od 0 do 9, tako da se slučajnim izborom određuje početni red i kolona. Postoje i druge mogućnosti, ali je važno zapaziti da se **određivanje početnih brojeva vrši na slučajan način.**

Izbor će se izvršiti uz pomoć tablice slučajnih brojeva, čiji je jedan dio – 500 slučajnih cifara, izabran slučajno:

```
20414 73116 53396 45514 41258 59982 90259 55178 08458 71774
00086 50389 28113 93952 82332 09778 50310 78893 84070 30868
57604 62449 59431 66212 57537 33199 84221 64024 14336 69938
05434 14191 15388 25842 93715 62231 07953 95147 86167 09776
08584 91135 36550 81581 56632 83391 98519 29472 71053 78592
12041 96013 29124 40859 49310 05481 84852 70980 74976 97914
58412 31398 25110 63108 06348 29941 37202 39923 86091 45797
89284 27659 19512 96681 29967 73323 24451 77388 18265 01821
60759 77025 43704 99855 40233 64418 46902 42143 21817 48469
59824 60305 79936 77690 08212 16697 87844 11462 27942 75866
```

STRATIFIKOVANO SLUČAJNO UZORKOVANJE

Drugi tip slučajnog uzorkovanja je **stratifikovano slučajno uzorkovanje.**

U stratifikovanom slučajnom uzorkovanju **populacija se dijeli na subpopulacije,** kao dijelove koji međusobno **nemaju presjeka –** zajedničkih elemenata i koji se nazivaju stratumi. Rječnikom teorije skupova, **to je jedna particija osnovnog skupa**. U daljem toku izbora istraživač primjenjuje jednostavno slučajno uzorkovanje za svaki pojedini dio populacije, odnosno **za svaki stratum**.

Osnovni razlog za korišćenje stratifikovanog slučajnog uzorkovanja je zbog toga što se na taj način otvara mogućnost za redukovanje uzoračkih grešaka.

Uzoračka greška pojavljuje se u slučajevima kada uzorak ne reprezentuje osnovni skup. S druge strane, **stratifikovano slučajno uzorkovanje obično je skuplje u odnosu na jednostavno slučajno uzorkovanje.**

Stratifikacija se najčešće vrši uz pomoć demografskih varijabli, kao što je pol, socioekonomska grupacija, geografska regija, religijska, nacionalna pripadnost i sl.

Na potrebu stratifikacije može da ukaže veliki broj primjera:

- **prilikom glasanja na nivou države** (opšti parlamentarni izbori);
- **analiza slušanosti ili gledanosti određenih emisija ili pisanja određenih medija** obavezno mora da uvaži stratifikaciju po polu ili prema starosnim grupama.
- **razlika urbanih i ruralnih dijelova neke zemlje**, a naročito u našim uslovima, nezaobilazan je vid stratifikacije u velikom broju istraživanja.
- **kod ispitivanja nekih aspekata privredne aktivnosti**, jedinice posmatranja najčešće su firme, kao pravna lica i nosioci

privredne aktivnosti. Razumljivo je da o istim pitanjima ne mogu da imaju stavove, **mala i velika preduzeća, preduzeća različitih djelatnosti, različitog oblika organizovanja** i sl.

U mnogim slučajevima potrebno je da se izvši stratifikacija u više nivoa, tako da se na ovaj način može govoriti o substratifikaciji ili o simultanoj stratifikaciji na više nivoa.

Stratifikovano slučajno uzorkovanje može biti proporcionalno ili disproporcionalno.

Proporcionalno stratifikovano slučajno uzorkovanje javlja se u slučaju kada je procenat uzorka uzet iz svakog stratuma, proporcionalan procentu sa kojim svaki stratum učestvuje u osnovnom skupu.

Ukoliko omjer u uzorku nije ovakav, tada je u pitanju disproporcionalno stratifikovano slučajno uzorkovanje.

SISTEMATSKO UZORKOVANJE

U sistematskom uzorkovanju bira se svaka *k*-ta jedinica populacije, veličine *N* jedinica, da bi se dobio uzorak od *n*-jedinica.

Sistematsko uzorkovanje ima **dvije osnovne karakteristike**:

- elementi populacije tretiraju se kao **uređen skup sekvenci**,
- **elementi se biraju u uzorak u konstantnom intervalu** iz sekvencijalno uređenog okvira uzorka.

Vrijednost *k* može se odrediti na sljedeći način:

$$k = N/n,$$

gdje je *n* veličina uzorka,

***N* veličina populacije,**

a *k* predstavlja veličinu intervala selekcije.

KLASTER UZORKOVANJE

Ovaj vid uzorkovanja podrazumijeva **dijeljenje populacije na dijelove koji nemaju presjeka,** odnosno koji nemaju zajedničkih elemenata. U odnosu na stratifikovano slučajno uzrokovanje, gdje su stratumi homogeni dijelovi populacije**, klaster uzorkovanje identifikuje klastere koji su sami za sebe heterogeni.**

U teoriji, svaki klaster sadrži širok varijetet elemenata, tako da **klaster predstavlja minijaturnu (ili mikrokosmos) populaciju, čiji je dio**. **Primjeri klastera su gradovi, kompanije, domaćinstva, dijelovi grada, geografske regije itd.** Vrlo često, klasteri su određeni prirodom stvari i identifikovani ciljevima, sadržajem i organizacijom statističkih istraživanja populacija.

Kada se klasteri ustanove, primjenjuje se slučajno uzorkovanje elemenata u svakom pojedinom klasteru i tako se u konačnosti dobija uzorak ovim metodom. **U mnogim državama se vrši istraživanje tržišta za nove proizvode upravo na osnovu klaster uzorkovanja. U tim istraživanjima, kao klasteri se odrede gradovi u kojima će se istraživanje vršiti.**

Ponekada su klasteri veoma veliki, pa se primjenjuje formiranje drugog nivoa klastera unutar klastera prvog nivoa. Ova se tehnika označava kao uzorkovanje na dva nivoa.

Klaster ili area uzorkovanje ima nekoliko važnih prednosti.

Prije svega, **to su pogodnost u primjeni i relativno niski troškovi provođenja uzorkovanja.**

Klasteri se dobijaju relativno jednostavno i troškovi uzorkovanja iz cijele populacije se redukuju, jer se obuhvat istraživanja svodi na klastere.

Uz to, **administriranje istraživanja može da se pojednostavi**.

Ponekad je klaster ili area uzorkovanje jedini mogući pristup, kada okvir uzorka individualnih elemenata populacije nije dostupan, što onemogućava primjenu drugih tehnika uzorkovanja.

Klaster ili area uzorkovanje ima i nekoliko nedostataka.

Ako su elementi klastera slični, tada ovaj vid uzorkovanja, u statističkom smislu, može biti manje efikasan u odnosu na jednostavno uzorkovanje.

U ekstremnom slučaju – kada su elementi klastera identični – uzorkovanje iz klastera može biti isto kao i uzorkovanje samo jedne jedinice iz klastera.

I još više, **troškovi i problemi statističke analize sa klaster ili area uzorkovanjem su veći nego sa jednostavnim slučajnim uzorkovanjem**.

NESLUČAJNO – NAMJERNO UZORKOVANJE

U situacijama kada se ne koristi slučajni izbor elemenata u uzorak, radi se o neslučajnom ili namjernom uzorkovanju. **Ove tehnike uzorkovanja se označavaju kao neprobabilističke, zbog toga što nije poznata vjerovatnoća izbora jedinica populacije u uzorak**. Uz to, **ne može da se odredi uzoračka greška.**

Tri najčešće tehnike neslučajnog uzorkovanja su:

1. pogodno (podesno, prikladno) uzorkovanje – **elementi se u uzorak biraju na način koji pogoduje istraživaču.**

2. subjektivno – prosudbeno uzorkovanje **javlja se u onim slučajevima kada se elementi uzorka biraju na osnovu suda – procjene istraživača.**

3. kvota uzorkovanje – **unekoliko liči na stratifikovano uzorkovanje.** Kvota uzorkovanje počinje određivanjem stratuma u okviru posmatrane populacije prema nekom obilježju: starosnoj dobi, polu, geografskoj regiji i dr. U daljem postupku, umjesto slučajnog uzorkovanja u okviru dobijenih

stratuma vrši se neslučajno uzorkovanje, sve dok se ne ispuni željena kvota za formiranje uzorka.

Kvote se određuju na različite načine, a najčešće na osnovu proporcija pojedinih podgrupa elemenata u osnovnom skupu.

U ovom slučaju, koncept kvota uzorkovanja sličan je proporcionalnom stratifikovanom uzorkovanju. Razlika je u neslučajnom izboru kod kvota uzorkovanja.

UZORAČKA GREŠKA

Uzoračka greška javlja se kao posljedica fluktuacije uzorka i uvijek je prisutna prilikom uzorkovanja. **Ovo zbog toga što uzorak, po pravilu, a to znači u najvećem broju uzoraka izabranih iz neke populacije, nije savršeno reprezentativan.**

Iz neke populacije moguće je izabrati manji ili veći broj uzoraka, u zavisnosti od toga da li se izbor vrši sa ponavljanjem ili bez ponavljanja elemenata, da li uzorak sadrži dva, tri ili više elemenata, a posebno u zavisnosti od toga da li je broj elemenata populacije manji ili veći.

Iz sljedećeg pregleda to se jasno vidi:

Veličina populacije	Veličina uzorka	Broj mogućih uzoraka bez ponavljanja*
30	5	142506
300	5	19582837560

* Broj mogućih uzoraka bez ponavljanja izračunat je kao broj kombinacija od n elemenata r-te klase, prema izrazu $C_r^n = \binom{n}{r} = \frac{n!}{r!(n-r)!}$.

Zbog toga što najveći broj uzoraka nije savršeno reprezentativan, statistika koja se izračunava na osnovu vrijednosti iz uzorka nije potpuno precizna ocjena parametra populacije. To je posljedica

uzoračke greške.

Kada se koriste tehnike slučajnog uzorkovanja, uzoračka greška javlja se sa određenom vjerovatnoćom. Prilikom korišćenja slučajnih uzoraka, uzoračka greška može se izračunati i analizirati.

NEUZORAČKE GREŠKE

Sve ostale greške koje nisu uzoračke greške označavaju se kao neuzoračke.

Mnoštvo mogućih neuzoračkih grešaka uključuje nedostajuće podatke, greške u zapisivanju, greške prilikom unosa podataka u računar, kao i analitičke greške.

Uz to, **neuzoračke greške se pojavljuju i kao posljedica instrumenata mjerenja, greške nejasnog definisanja, nepodesnih i pogrešnih pitanja u upitnicima, kao i nejasnog koncepta istraživanja.**

Loše definisan okvir takođe uzrokuje greške ovog tipa. Greške u odgovorima uzrokuju neuzoračke greške u slučajevima kada ispitanici ne žele da odgovore, kada ne znaju odgovor ili pređu okvir dozvoljenog u odgovoru (od ponuđena tri odgovora zaokruže sva tri, bez obzira što je moguće dati samo jedan odgovor).

U raznim vrstama anketa, u kojima ispitanici šalju odgovore pojavljuju se neuzoračke greške koje se ogladaju u sljedećem: u nekim slučajevima ispitanici neke događaje pripisuju pogrešnom vremenskom periodu; u drugom slučaju ispitanik jednostavno zaboravi prošle događaje, i konačno, greška se javlja i kada se ispitanik ne sjeća tačno određenih događaja.

Nastojanje treba da ide u pravcu **da se pažljivim planiranjem i provođenjem istraživanja ove greške eliminišu ili svedu na najmanju moguću mjeru.**

INTERVALI POVJERENJA (INTERVALNO OCJENJIVANJE PARAMETARA POPULACIJE)

Tačkasto i intervalno ocjenjivanje

Interval povjerenja za aritmetičku sredinu populacije μ kada je poznata standardna devijacija populacije

Interval povjerenja za aritmetičku sredinu populacije μ kada nije poznata standardna devijacija populacije

Interval povjerenja za proporciju populacije π (na osnovu velikih uzoraka)

Određivanje velicine uzorka

CILJEVI POGLAVLJA

Nakon čitanja ovog poglavlja bićete u stanju da:

7. shvatite smisao statističkog zaključivanja
8. shvatite razliku između tačkastog i intervalnog ocjenjivanja
9. formirate i objasnite interval povjerenja za aritmetičku sredinu
10. navedete karakteristike Studentovog *t* rasporeda
11. shvatite značenje statističkog termina stepeni slobode
12. formirate i objasnite interval povjerenja za proporciju skupa
13. odredite veličinu uzorka prilikom ocjenjivanja aritmetičke sredine i proporcije

Jedan od osnovnih ciljeva statističke inferencije ogleda se u nastojanju da se na osnovu informacija dobijenih iz uzorka izvedu zaključci o vrijednostima parametara osnovnog skupa.

Postupak donošenja zaključaka o vrijednostima parametara osnovnog skupa, na osnovu vrijednosti statistika dobijenih iz uzorka, predstavlja statističko zaključivanje.

Statističko zaključivanje ima dva oblika: izvodi se kao **statističko ocjenjivanje** ili se provodi kao **testiranje statističkih hipoteza**. **Razlika u ova dva pristupa proizlazi zbog različitih informacija kojima se raspolaže prije izvođenja postupka.**

Statističko ocjenjivanje provodi se u slučajevima kada se ne raspolaže informacijama na osnovu kojih bi se mogla pretpostaviti vrijednost nekog parametra osnovnog skupa. To su, najčešće, aritmetička sredina, proporcija i standardna devijacija osnovnog skupa.

S druge strane, **često smo u situaciji da postoje određene informacije o vrijednosti parametra osnovnog skupa i u tim slučajevima se provodi procedura testiranja statističkih hipoteza. Postupak testiranja statističkih hipoteza sastoji se u provjeravanju održivosti postavljenih hipoteza o vrijednostima parametara osnovnog skupa.**

TAČKASTO I INTERVALNO OCJENJIVANJE

Postoje dva načina statističkog ocjenjivanja: **tačkasto ocjenjivanje i intervalno.**

Tačkasto ocjenjivanje **ili ocjenjivanje jednim brojem kao polazište ima statistiku uzorka koja se koristi za ocjenu stvarne vrijednosti parametra populacije.**

Tačkasto ocjenjivanje podrazumijeva da se aritmetička sredina uzorka $\overline{X}$ proglasi tačkastom ocjenom aritmetičke sredine populacije μ, odnosno da se bilo koja statistika uzorka (proporcija, varijansa i sl.) odredi kao ocjena vrijednosti parametra populacije.

Intervalno ocjenjivanje **podrazumijeva proceduru u kojoj se formira interval u kojem se sa određenom vjerovatnoćom očekuje vrijednost parametra osnovnog skupa**.

Interval koji se formira ima posebnu osobinu pouzdanosti (povjerenja) ili vjerovatnoću tačnog (korektnog) ocjenjivanja prave vrijednosti parametra populacije.

Interval ocjene sa odgovarajućom mjerom pouzdanosti (povjerenja) uobičajeno se naziva interval povjerenja.

Interval povjerenja

Interval povjerenja je raspon (područje) vrijednosti za koje se vjeruje da uključuju nepoznatu vrijednost parametra populacije. Uz ovaj interval ide i informacija o mjeri povjerenja sa kojom očekujemo da ovaj interval zaista sadrži vrijednost posmatranog parametra.

(Napomena: kada se bazira na istoj informaciji, interval sa višim stepenom pouzdanosti je širi u odnosu na interval sa nižim stepenom pouzdanosti).

Zaključak o nepoznatoj vrijednosti parametra populacije izvodi se na osnovu odgovarajuće statistike uzorka.

Parametri populacije i njihove ocjene mogu se predstaviti sljedećim pregledom:

	Parametar populacije (konstanta)	Ocjena parametra populacije (slučajna promjenljiva)	Ocijenjena vrijednost (vrijednost dobijena iz uzorka) (konstanta)
Aritmetička sredina	μ	$\overline{X}$	$\overline{x}$
Proporcija	π	P	p
Standardna devijacija	σ	S	s

Statistika uzorka na osnovu koje se ocjenjuje parametar populacije naziva se ocjena parametra populacije i predstavlja slučajnu promjenljivu.

Osobine ocjena

Poželjne osobine ocjena su: nepristrasnost, efikasnost, konzistentnost i dovoljnost.

Ocjena parametra populacije je nepristrasna ako je njena očekivana vrijednost (prosjek) jednaka parametru populacije.

Efikasnost ocjene podrazumijeva njenu karakteristiku varijabiliteta. Ocjena je utoliko efikasnija ukoliko ima manju varijansu, odnosno standardnu grešku, za uzorke iste veličine.

Konzistentnost, kao osobina ocjena, veže se za tendencije ocjene u odnosu na vrijednost parametra populacije u slučajevima kada se veličina uzorka mijenja. Za ocjenu se može konstatovati da je konzistentna ako, sa povećanjem veličine uzorka, ona teži vrijednosti parametra populacije.

Što je *n* veće, to se realizovane vrijednosti ocjene u uzorcima sve više koncentrišu oko prave vrijednosti parametra populacije, a za n $\rightarrow \infty$

izjednačavaju se sa njegovom vrijednosšću.

Ocjena je dovoljna ako uzima u obzir cjelovitu informaciju iz uzorka. Tako, na primjer, $\overline{X}$ je dovoljna ocjena, jer koristi cjelovitu informaciju iz uzorka.

$\overline{X}$ $\sigma/\sqrt{n}$

INTERVAL POVJERENJA ZA ARITMETIČKU SREDINU POPULACIJE μ KADA JE POZNATA STANDARDNA DEVIJACIJA POPULACIJE

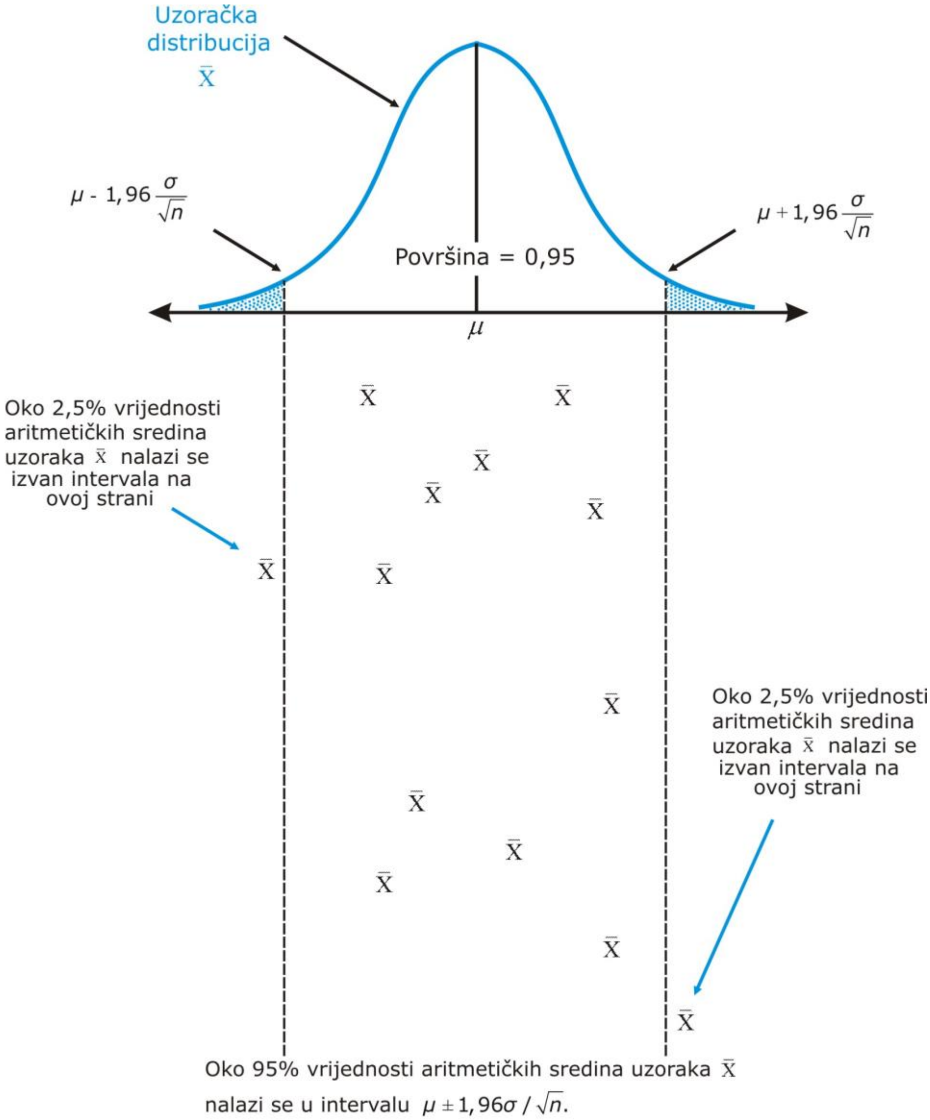
Prema Centralnoj Graničnoj Teoremi, kada se bira veliki slučajni uzorak (veći od 30) iz bilo koje populacije sa aritmetičkom sredinom μ i standardnom devijacijom σ, sredina uzorka je, bar približno, normalno raspoređena sa aritmetičkom sredinom μ i standardnom devijacijom.

Ako je populacija normalno raspoređena, tada je $\overline{X}$ takođe normalno raspoređena za bilo koju veličinu uzorka.

Za standardizovanu normalnu promjenljivu Z vjerovatnoća je 0,95 da uzme vrijednost između –1,96 i +1,96. $\overline{X}$ $\overline{X}$

Transformacija Z u slučajnu promjenljivu omogućava da se prije uzorkovanja ustanovi vjerovatnoća od 0,95 da će biti u intervalu:

$$\mu \pm 1{,}96\frac{\sigma}{\sqrt{n}}.$$



Slika 7.1 Raspored aritmetičkih sredina uzoraka

Treba uočiti da je udaljenost između $\bar{x}$ i μ ista kao i udaljenost

između μ i $\bar{x}$.

To implicira da će se $\bar{x}$ naći u intervalu $\mu \pm 1{,}96\sigma/\sqrt{n}$ onda i samo onda ako se μ nalazi u intervalu $\bar{x} \pm 1{,}96\sigma/\sqrt{n}$. U velikom broju ponovljenih pokušaja, ovo će se desiti u 95% slučajeva.

Interval $\bar{x} \pm 1{,}96\sigma/\sqrt{n}$ nazivamo 95%-tni interval povjerenja za nepoznatu vrijednost aritmetičke sredine populacije μ.

Interval sa 95% pouzdanosti za μ, kada je σ poznata, a uzorkovanje izvedeno iz populacije koja je normalno raspoređena ili su korišćeni veliki uzorci, je:

$1{,}96\sigma/\sqrt{n}$ $$\bar{x} \pm 1{,}96\frac{\sigma}{\sqrt{n}}$$

Veličina često se naziva granica greške ili uzoračka greška.

U poslovanju i nekim drugim aplikacijama obično se koristi 95%-tni interval povjerenja. S druge strane, **postoji niz drugih nivoa povjerenja.**

Tako, na primjer, za nivo pouzdanosti 90% z vrijednost je 1,65, a u slučaju 99% z = 2,58.

$z_{\alpha/2}$ $z_{\alpha/2}$

Označimo sa onu vrijednost z koja odsijeca desni dio standardizovane normalne krive α/2.

Nivo pouzdanosti

Dio površine ispod standardizovane normalne krive $1 - \alpha$ označava se kao koeficijent pouzdanosti. α predstavlja vjerovatnoću greške, odnosno vjerovatnoću da interval neće obuhvatiti nepoznati parametar. Koeficijent pouzdanosti, ako se pomnoži sa 100, dobija izraz u procentima i naziva se nivo pouzdanosti.

Možemo da konstatujemo da $(1-\alpha)100\%$ interval povjerenja za aritmetičku sredinu populacije μ kada je poznata standardna devijacija populacije σ, a uzorkovanje se vrši iz normalno raspoređene populacije ili na osnovu velikih uzoraka, dat je izrazom:

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \, .$$

Tako za 95%-tni interval povjerenja za μ je:

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

$(1 - \alpha) = 0{,}95$

$\alpha = 0{,}05$

$\alpha/2 = 0{,}025.$

Iz Tablice Standardizovanog normalnog rasporeda iznalazi se da je $z_{\alpha/2} = 1{,}96$.

Odnos nivoa pouzdanosti i širine intervala povjerenja

Kada se uzorkovanje izvodi iz iste populacije, korišćenjem fiksne veličine uzorka, tada što je nivo pouzdanosti veći, to je interval povjerenja širi.

Jasno je da širi interval ima veću šansu da obuhvati stvarnu vrijednost aritmetičke sredine populacije. Može se zaključiti da ukoliko želimo interval sa višim stepenom pouzdanosti, tada će interval biti utoliko širi.

To znači da se ovdje radi o nastojanju da budemo u isto vrijeme i dovoljno tačni i dovoljno precizni.

Princip tačnosti zahtijeva što širi interval, a preciznost podrazumijeva što uži interval. Nastojanje statističke metodologije ide u smjeru dobijanja što preciznijeg, a u isto vrijeme što tačnijeg intervala.

Odnos veličine uzorka i širine intervala povjerenja

Kada se uzorkovanje izvodi iz iste populacije, korišćenjem fiksnog nivoa povjerenja, tada što je uzorak veći, to je interval povjerenja uži.

Nivo pouzdanosti			
100 (1 - α)	α	$\alpha/2$	$z_{\alpha/2}$
90%	0,10	0,05	1,65
95%	0,05	0,025	1,96
99%	0,01	0,005	2,58

INTERVAL POVJERENJA ZA ARITMETIČKU SREDINU POPULACIJE μ KADA NIJE POZNATA STANDARDNA DEVIJACIJA POPULACIJE

Prilikom formiranja intervala povjerenja za μ pretpostavka je bila da je populacija normalno raspoređena ili da je uzorak veliki (kao obezbjeđenje postizanja normalnosti, u smislu Centralne Granične Teoreme). Uz to, **pretpostavljeno je da je standardna devijacija populacije poznata.**

Ova pretpostavka bila je potrebna iz teorijskih razloga, čime je bilo omogućeno **korišćenje standardizovane normalne vjerovatnoće u konstrukciji intervala povjerenja.**

U stvarnim situacijama, odnosno u praktičnim uslovima uzorkovanja, poznavanje standardne devijacije populacije σ je veoma rijetko.

t – distribucija

U praktičnim procesima uzorkovanja standardna devijacija populacije nije poznata.

Da bi se postupak nastavio, uprkos činjenici da nije poznata standardna devijacija osnovnog skupa σ, kao zamjena za nju uzima se standardna devijacija uzorka *S*, kao najlogičnija vrijednost.

Ako je populacija normalno raspoređena, tada standardizovana statistika:

$$t = \frac{\overline{X} - \mu}{S / \sqrt{n}}$$

ima *t* raspored sa *n* – 1 stepeni slobode.

William Gosset - Student 1876 - 1937

Autor ove distribucije je W. S. Gossett, koji je kao istraživač radio u Guinness pivari u Dublinu u Irskoj i koji je ustanovio ovu distribuciju 1908. godine. S obzirom da uposlenicima ove kompanije nije bilo dozvoljeno da objavljuju radove pod svojim imenom, Gossett je publikovao svoj rad pod imenom Student.

Ova distribucija naziva se i Student-ova *t* distribucija. *t* distribucija karakteriše se parametrom *df*, tj. brojem stepeni slobode.

Za bilo koju cijelu vrijednost *df* = 1, 2, 3, ..., postoji odgovarajuća vrijednost *t*- distribucije.

Kao i standardizovana normalna distribucija, tako je i *t* distribucija simetrična i zvonastog oblika, s tim što *t* distribucija ima šire krajeve u odnosu na *Z* distribuciju.

Karakteristike *t* distribucije

Aritmetička sredina *t* distribucije je 0. Za *df* > 2, varijansa *t* distribucije jednaka je *df*/(*df* – 2).

Pojam stepeni slobode može se pojasniti na sljedeći način: Izračunavanje varijanse uzorka podrazumijeva izračunavanje izraza:

$$\sum_{i=1}^{n}(X_i - \overline{X})^2 .$$

Da bi ovaj izraz mogao da se izračuna, potrebno je da prethodno bude izračunata (poznata) vrijednost aritmetičke sredine $\overline{X}$. Imajući to na umu, razumljivo je da samo *n* – 1 vrijednosti iz uzorka može slobodno da varira, a da je jedna vrijednost određena sa tih *n* – 1 vrijednosti. Zbog tog ograničenja, u ovom slučaju bismo odredili da je broj stepeni slobode *n* – 1.

Uočljivo je da *t* i *Z* distribucija imaju jednake aritmetičke sredine, dok je varijansa *t* distribucije veća.

Stepeni slobode

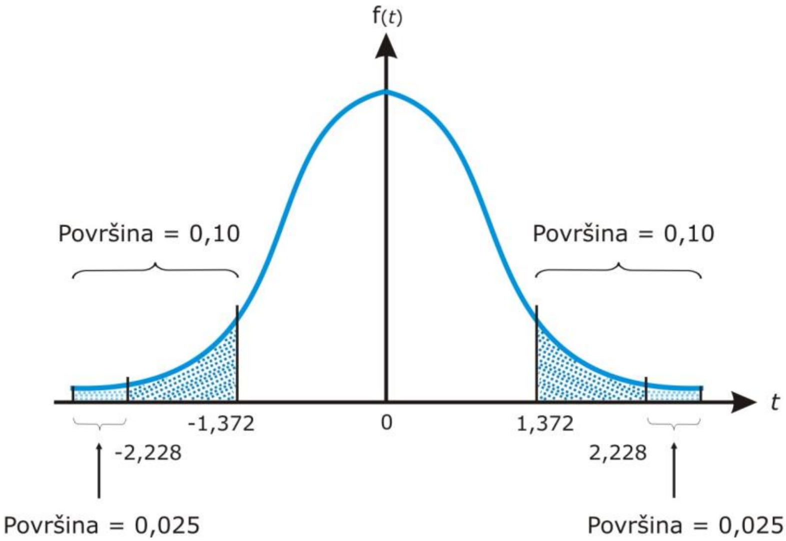
Broj stepeni slobode predstavlja broj nezavisnih vrijednosti u uzorku, koji se dobije kada se broj raspoloživih vrijednosti umanji za broj ograničenja koja se nameću ovim vrijednostima.

Ukoliko *df* raste, tada se vrijednost varijanse za *t* približava 1,00, što je vrijednost varijanse za *Z*. Veća varijansa i širi krajevi distribucije *t* prirodna su posljedica činjenice da se radi o uslovima veće neizvjesnosti.

Što je broj stepeni slobode *df* veći, to se *t* distribucija više približava Z distribuciji.

Vrijednosti *t* distribucije za odgovarajuće vjerovatnoće date su u Tablici Student-ovog *t* rasporeda.

Kao što je već rečeno, *t* distribucija se približava normalnoj ukoliko se broj stepeni slobode uvećava. Za beskonačno veliki broj stepeni ove dvije distribucije se izjednačavaju.



Slika 7.3 *t* – distribucija (*df* = 10)

(1 - α)100% interval povjerenja za aritmetičku sredinu populacije μ, u slučaju kada nije poznata standardna devijacija populacije σ, uz pretpostavku da je populacija normalno raspoređena, dat je sa:

$t_{\alpha/2}$

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

u kojem je vrijednost *t* distribucije sa *n* – 1 stepeni slobode i za vjerovatnoću α/2.

Za veliki broj stepeni slobode *t* distribucija može dobro da se aproksimira Z distribucijom. Bez obzira što standardna devijacija populacije σ nije poznata (uz pretpostavku da je populacija normalno raspoređena), prava distribucija koju treba koristiti je *t* distribucija sa *n* – 1 stepeni slobode.

Ovo znači da se uvijek mogu koristiti vrijednosti *t* distribucije, pod pretpostavkom da mogu da se pronađu u tablici. Pri tome je u velikom broju slučajeva, pa i primjera u ovoj knjizi, prisutna pretpostavka o (bar približno) normalnom rasporedu populacije.

Za velike uzorke ova pretpostavka manje je važna, zbog djelovanja Centralne Granične Teoreme.

(1 - α)100%-tni interval povjerenja za aritmetičku sredinu populacije μ, uz korišćenje velikih uzoraka, je:

$$\overline{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}.$$

INTERVAL POVJERENJA ZA PROPORCIJU POPULACIJE π (NA OSNOVU VELIKIH UZORAKA)

Vrlo često, interes u ispitivanju više je usmjeren na kvalitativni aspekt, nego na kvantitativni. To se dešava, prije svega, u slučajevima kada je od interesa ustanovljavanje relativne frekvencije pojavljivanja elemenata sa određenim karakteristikama u populaciji.

U ovakvim slučajevima potrebno je ocijeniti proporciju populacije π.

Ocjena proporcije populacije π je proporcija uzorka *P*.

Prosječna vrijednost uzoračke distribucije proporcije P je jednaka vrijednosti proporcije populacije π,

a standardna greška proporcije je $\sqrt{\frac{\pi(1-\pi)}{n}}$.

U slučaju velikih uzoraka, umjesto nepoznate vrijednosti parametra populacije koristi se realizovana vrijednost statistike uzorka p, tako da se standardna devijacija izračunava na osnovu izraza $\sqrt{\frac{p(1-p)}{n}}$.

Dovoljno velikim smatraju se oni uzorci kod kojih su zadovoljeni uslovi $n\pi > 5$ i $n(1 - \pi) > 5$. Za realizovanu vrijednost proporcije iz uzorka to podrazumijeva provjeru vrijednosti np i $n(1 - p)$.

(1 - α)100% interval povjerenja proporcije populacije π dat je izrazom:

$$p \pm z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}},$$

gdje proporcija uzorka p predstavlja učešće elemenata uzorka sa određenom osobinom (x) u uzorku veličine n, odnosno $p = x/n$.

Evidentno je da se povećavanjem ili smanjivanjem uzorka može očekivati drugačija vrijednost parametra (u ovom slučaju proporcije), što nije prepreka u daljem radu. Ipak, **logična je pretpostavka da je vrijednost parametra iz većih uzoraka bliža pravoj vrijednosti parametra populacije.**

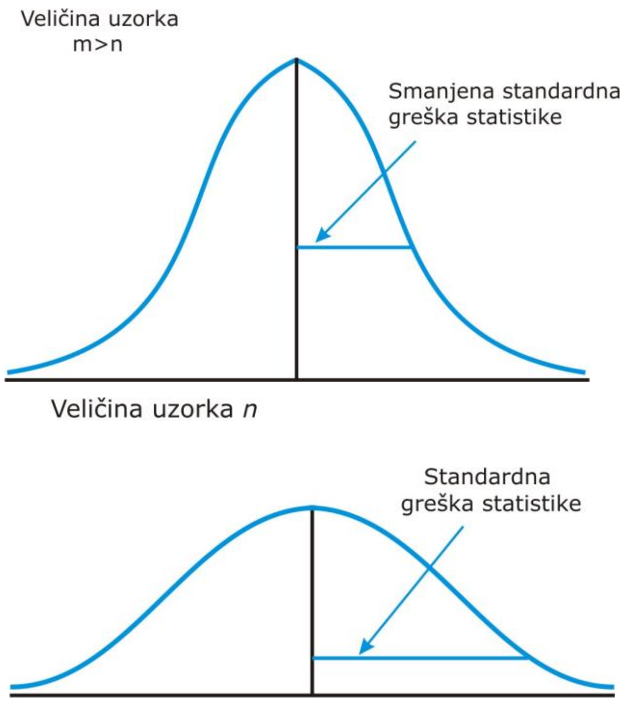
Ovo je prirodna posljedica smanjenja uzoračke greške, prilikom povećanja veličine uzorka. **To upućuje na zaključak da je povećanje uzorka uvijek poželjno u nastojanju da se poveća preciznost intervala povjerenja.**

ODREĐIVANJE VELIČINE UZORKA

Jedno od veoma važnih pitanja koje se postavlja prije svakog uzorkovanja jeste pitanje, **koliki uzorak treba da bismo dobili validne rezultate i zadovoljavajuće ocjene. Sa statističke tačke gledišta, potrebno je da je uzorak što veći.**

U ovakvim slučajevima moraju se dobiti odgovori (od dizajnera istraživanja) na sljedeća pitanja:

1. **Koliko blizu treba da je ocjena uzorka parametru populacije**? Odgovor na ovo pitanje označava se sa *G* (kao granica).
2. **Koliki treba da je interval povjerenja** tako da udaljenost između ocjene i parametra populacije bude manja ili jednaka *G*?
3. Zadnje, i pitanje koje se često ne razumije, jeste **pitanje koja je (kako je dobijena) ocjena varijanse (ili standardne devijacije) populacije.**



Kada je u pitanju ocjena varijanse, često se statističar upućuje na to da odredi varijansu.

Jedan od načina može biti i ovaj: ako je populacija približno normalnog oblika i ako se dobiju 95%-tne granice vrijednosti u populaciji, potrebno je podijeliti razliku između gornje i donje granice sa 4 i na taj način se dobije gruba ocjena standardne devijacije populacije σ.

Može da se uradi i ovako: **izvrši se malo i ne skupo pilot istraživanje, i σ se ocijeni kao standardna devijacija dobijenog uzorka. Nakon što se dobiju odgovori na postavljena pitanja, u daljem je potrebno uvrstiti vrijednosti u odgovarajući izraz.**

Minimalna veličina uzorka prilikom ocjenjivanja aritmetičke sredine populacije μ je:

$$n = \frac{z_{\alpha/2}^2 \sigma^2}{G^2}.$$

Minimalna veličina uzorka prilikom ocjenjivanja proporcije populacije π je:

$$n = \frac{z_{\alpha/2}^2 \pi(1-\pi)}{G^2}.$$

Prethodni izrazi izvedeni su iz odgovarajućih izraza za intervale povjerenja za ove parametre bazirane na normalnoj distribuciji.

U slučaju aritmetičke sredine populacije, *G* je polovina od (1 - α)100% intervala povjerenja za aritmetičku sredinu populacije, tako da je:

$$G = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}.$$

Treba zapaziti da je *G* granica greške. Rješenje se ogleda u iznalaženju minimalne veličine uzorka za datu granicu greške.

Da bi se mogla koristiti jednačina za minimalnu veličinu uzorka prilikom ocjenjivanja proporcije populacije π potrebno je odrediti zamjenu za π – nepoznatu vrijednost proporcije uzorka. **Mora biti urađena neka a priori ocjena.**

Ako to nije moguće, može se izabrati pilot uzorak ili, u odsustvu bilo kakve informacije, koristi se vrijednost $\pi = 0.5$.

Ova vrijednost maksimizira izraz $\pi(1 - \pi)$ i na taj način se obezbjeđuje da minimalna veličina uzorka zadovoljava i sve ostale moguće vrijednosti za π.

TESTIRANJE STATISTIČKIH HIPOTEZA

Uvod

Kontroverze i ograničenja statističkog testiranja hipoteza

Postupak testiranja statističkih hipoteza

Testiranje hipoteze zasnovano na jednom uzorku

Testiranje hipoteze zasnovano na dva uzorka

Veza između testiranja hipoteza i intervala povjerenja

Česte greške u praksi i preporuke pri testiranju hipoteza

CILJEVI POGLAVLJA

Nakon čitanja ovoga poglavlja bićete u stanju da

1. shvatite značaj, logiku kao i ograničenja statističkog testiranja hipoteza
2. sprovedete testiranje na osnovu *p*-vrijednosti
3. primjenite najpoznatije parametarske testove u slučaju jednog i dva nezavisna uzorka
4. protumačite kompjuterski izlaz testiranja hipoteza za bilo koji statistički test

UVOD

Cilj inferencijalne statistike je da se na osnovu informacije iz uzorka dođe do što je moguće prezicnije informacije o statističkom skupu.

Dva osnovna metoda inferencijalne statistike (ili statističkog zaključivanja) su: ocjenjivanje i testiranje statističkih hipoteza.

Inferencijalna statistika → **Ocjenjivanje**

Inferencijalna statistika → **Testiranje hipoteza**

Tradicionalno, **ocjenjivanje se koristilo kada istraživač nije imao nikakvu prethodnu informaciju o posmatranom skupu.** Interval povjerenja sadrži dragocjenu informaciju o nepoznatom parametru osnovnog skupa. Na osnovu njega istraživač može da formuliše probabilistički stav i da sa odgovarajućom pouzdanošću tvrdi da se nepoznati parametar nalazi u dobijenom intervalu.

Druga oblast inferencijalne statistike, **testiranje statističkih hipoteza, tradicionalno se koristilo kada o skupu imamo neko prethodno saznanje**. Cilj testiranja je da se ispita prihvatljivost nekih tvrđenja ili pretpostavki koje se tiču osobina jednog ili više osnovnih skupova.

> Statistička hipoteza
>
> Statistička hipoteza je precizno formulisano tvrđenje ili pretpostavka o nekoj važnoj karakteristici jednog ili više skupova.
>
> Statistički metod kojim se empirijska evidencija uzorka koristi radi procjenjivanja njene prihvatljivosti naziva se testiranjem hipoteze, a procedura statističkim testom.

Na osnovu do sada navedenog možemo dati sledeću generalnu preporuku o tome kada se koristi ocjenjivanje a kada testiranje.

Preporuka za izbor između ocjenjivanja i testiranja

Ocjenjivanje se koristi kada je potrebno samo doći do informacije o nepoznatom parametru osnovnog skupa.

Testiranje hipoteza se primjenjuje ako su istovremeno ispunjena sledeća dva uslova

a) Na istraživačko pitanje (problem) postoje smo dva moguća odgovora, da ili ne

b) Problem se rješava korišćenjem uzorka.

Sve statističke testove možemo klasifikovati na dva načina. Po prvom posmatramo broj uzoraka (ili skupova) koje analiziramo. Shodno tome sve testove dijelimo u tri grupe.

- Statistički testovi
  - Testovi zasnovani na jednom uzorku
  - Testovi zasnovani na dva uzorka
  - Testovi zasnovani na tri i više uzoraka

Po drugoj klasifikaciji svi testovi se dijele u dvije grupe, s obzirom na vrstu i jačinu pretpostavki na kojima se zasnivaju.

- Statistički testovi
  - Parametarski (klasični) testovi
  - Neparametarski testovi

Važna napomena: naziv "neparametarski testovi" ne znači da se ovi testovi ne koriste pri testiranju parametara skupa.

Zbog toga i naglasak u modernom testiranju hipoteza više ne smije biti puka primjena nekog testa, već prije svega a) ispitivanje preduslova na kojima se test zasniva i b) interpretaciji rezultata.

	Parametarski	Neparametarski
Testovi zasnovani na jednom uzorku	Z test, t test Poglavlje 8	Test znakova, Wilcoxon-ov test ranga sa znakom Poglavlje 13
Testovi zasnovani na dva uzorka	Z test, t test Poglavlje 8	Test znakova, Wilcoxon-ov test ranga sa znakom, Test sume rangova
Testovi zasnovani na tri i više uzoraka	Analize varijanse Poglavlje 9	χ^2test Poglavlje 10 Kruskal-Wallice, Friedman (nisu

KONTROVERZE I OGRANIČENJA STATISTIČKOG TESTIRANJA HIPOTEZA

Statističke testove prvi je ozbiljnije razmatrao najpoznatiji statističar prošlog veka, Sir Ronald Fisher.

Ronald Fisher
(1890 - 1962)

Egon Pearson
(1895 - 1980)

Jerzy Newman
(1894 - 1981)

On je smatrao da bi procedura testiranja hipoteza trebalo da se zasniva na postavljanju samo jedne hipoteze i na tzv. *p*-vrijednosti, koju ćemo detaljno objasniti u ovom poglavlju. Jedan od nedostataka *p*-vrijednosti, međutim, je u tome što se jako teško izračunava.

Sa druge strane, **Jerzy Newman i Egon Pearson (sin Karla Pirsona) su imali drugačije viđenje testiranja hipoteza i uveli su različit način testiranja koji se zasniva na nultoj i alternativnoj hipotezi, na nivou značajnosti testa i na kritičnim oblastima testa. Sve ove pojmove upoznaćemo malo kasnije.** Skoro trideset godina (od 1930 do 1962) vodila se (često neumjerena) polemika između navedenih statističara. Nakon Fisherove smrti definitivno je prihvaćen Newman-Pearson-ov koncept testiranja i Fisherov napušten. Ovakva situacija je potrajala sve do prije desetak godina.

Sa revolucijom kompjuterske tehnologije izračunavanje *p*-vrijednosti je postalo krajnje jednostavno, naravno uz pomoć statističkog softvera. Svi statistički softveri bazirali su testiranje isključivo na osnovu *p*-vrijednosti. Ovo je, kao bumerang, imalo povratni uticaj na statistički praksu, i Fisherove ideje su krajem XX-og veka ponovo stupile na scenu. **Štaviše, s obzirom na veliku prednost u tumačenju rezultata testa pomoću *p*-vrijednosti, počele su da se pojavljuju knjige koje su čak potpuno ignorisale Newman-Pearson-ove kritične vrijednosti.**

Naš pristup zasnivaće se na integraciji navedena dva pravca mišljenja. Drugačije rečeno, od Newman-Pearson-ovog koncepta preuzećemo postojanje dvije hipoteze (nulte i alternativne) i nivoa značajnosti, a od Fisherove *p*-vrijednost. Ipak ćemo ukratko objasniti i značenje kritičnih vrijednosti i koristiti ih kod nekih testova (recimo χ^2 testa).

Drugi problem kod statističkih testova je u njihovoj često pogrešnoj primjeni, kao i pogrešnom tumačenju rezultata.

Treći i najvažniji, je u tome da je sama filozofija testiranja osporena u velikom broju radova koji su se pojavili u posljednjoj deceniji

prošlog vijeka. Suštinska zamjerka odnosi se na nultu i alternativnu hipotezu. Naime, **ističe se da je testiranje hipoteza besmisleno jer unaprijed znamo da je nulta hipoteza pogrešna.**

Dodatno, ovi autori ističu da će se sa povećavanjem uzorka svaka nulta hipoteza konačno odbaciti. Čemu onda testirati bilo šta ako se ishod testiranja unaprijed zna? Ovo ćemo nazvati Berksonov paradoks, po prvom statističaru koji je ukazao na ovaj problem.

Šta autori koji osporavaju testiranje hipoteza nude kao alternativu? Svi oni smatraju da bi trebalo koristiti samo intervale povjerenja jer na osnovu njih nikada ne možemo dolaziti do besmislenih i nepreciznih zaključaka kao kod testiranja.

POSTUPAK TESTIRANJA STATISTIČKIH HIPOTEZA

Odmah napomenimo da se u statistici nikada ne zaključuje da je nulta hipoteza stvarno istinita, već samo da nemamo dovoljno dokaza da je odbacimo.

Prilikom testiranja neke tvrdnje ili pretpostavke postavljaćemo dve suprostavljene hipoteze, nultu i alternativnu. Da bi se odbacila nulta hipoteza uzimaćemo uzorak i tražiti empirijske dokaze. Ako pronađemo dovoljno dokaza odbacićemo nultu hipotezu i usvojiti alternativnu.

Ako, pak, ne pronađemo, kazaćemo samo jedno: nismo našli dokaze da odbacimo nultu hipotezu. Potpuno je pogrešno tvrditi da je nulta hipoteza tačna.

Ovo važi uvijek: statističkim testiranjem je nemoguće dokazati da je nulta hipoteza istinita.

Postupak statističkog testiranja hipoteza

1. Na osnovu postavljenog problema formuliše se nulta i alternativna hipoteza i bira tzv. nivo značajnosti, α

↓

2. Uzima se slučajan uzorak, bira odgovarajući test i izračunava tzv. statistika testa

↓

3. Provjerava se da li su ispunjene pretpostavke na kojima se izabrani test zasniva. Ako nisu, prekida se proces testiranja i bira neki drugi test.

↓

4. Određuje se *p*-vrijednost (kod Newman-Pearsonovog pristupa određuju se tzv. kritične vrijednosti iz tablica odgovarajućeg rasporeda)

↓

5. Formuliše se pravilo odlučivanja i donosi odluka o odbacivanju ili neodbacivanju nulte hipoteze.

↓

6. Formuliše se zaključak u kontekstu postavljenog problema

Dijagram 8.2 Etape u testiranju statističkih hipoteza

Nulta i alternativna hipoteza

Nulta hipoteza

Nulta hipoteza (H_0) je postavka (iskaz) o vrijednosti nepoznatog parametra skupa. Testiranje hipoteza je tako koncipirano da se procjeni jačina dokaza protiv nulte hipoteze. Da bi se H_0 odbacila potrebno je pronaći dovoljno ubjedljive dokaze u uzorku. Konkretna nulta hipoteza varira od problema do problema, ali tipično se ona postavlja u vidu status quo, tj. da "nema razlike", "nema uticaja", ili da "nema promjene".

Nulta hipoteza može biti **prosta i složena**.

Nulta hipoteza je prosta ako se njom tvrdi da je parametar jednak tačno jednoj, unapred poznatoj numeričkoj vrijednosti, tzv. hipotetičnoj vrijednosti.

Nultu hipotezu, na primjer, možemo simbolički napisati $H_0 : \mu = 100$ gr i čitamo: "nulta hipoteza glasi da je aritmetička sredina skupa jednaka 100 grama".

Ako nulta hipoteza obuhvata veći broj mogućih vrijednosti, na primjer, $H_0 : \mu \leq 100$ gr, ona je složena. Bez obzira kako je postavljena, nulta hipoteza ona uvijek mora da sadrži znak jednakosti.

Svakoj nultoj hipotezi suprostavljamo alternativnu hipotezu i označavamo je sa H_1 ili H_a.

Alternativna hipoteza

Alternativna hipoteza je tvrđenje o parametru skupa koje će biti prihvaćeno samo kada se nađu dovoljno ubjedljivi dokazi u uzorku. Najčešće se postavlja u vidu da "ima razlike ", da "postoji uticaj" ili da je "došlo do promjene".

Alternativna hipoteza najčešće sadrži sve vrijednosti koje parametar može imati a koje nisu obuhvaćene nultom hipotezom. Zbog toga je po pravilu data u obliku složene hipoteze. **H_1 se u naučnim istraživanjima često naziva i istraživačkom hipotezom jer njom istraživač izražava mišljenje koje postupkom testiranja nastoji da potvrdi.** Usljed toga u praksi često najprije postavljamo alternativnu hipotezu, a tek zatim nultu.

Alternativnu hipotezu kojom moguća odstupanja parametra u odnosu na hipotetičku vrijednost detektujemo u oba smjera nazivamo dvosmjernom ili dvostranom hipotezom. Test koji se primjenjuje u ovakvoj situaciji nazivamo dvosmjernim ili dvostranim testom.

Hipoteze mogu biti postavljene kao jednosmjerne ili jednostrane, a testove kojima ispitujemo prihvatljivost H_0 nazivamo jednosmjernim ili jednostranim testovima. Primjetimo da alternativna hipoteza ukazuje na smijer odstupanja koji je kritičan.

> Preporuka za izbor između jednosmjernog i dvosmjernog testa
>
> Ako ništa unaprijed ne znamo o potencijalnom odstupanju parametra od njegove hipotetične vrijednosti uvijek ćemo koristiti dvosmjerne testove.

Greške pri testiranju i nivo značajnosti testa

Nulta hipoteza, kao tvrđenje o vrijednosti nepoznatog parametra osnovnog skupa, može u stvarnosti biti ili istinita ili neistinita. Sa druge strane, podaci slučajnog uzorka mogu biti ili saglasni sa H_0 ili joj protivrječiti.

Tabela 8.1 Ishodi testiranja hipoteze i njihove vjerovatnoće

		Stvarna situacija	
		H_0 istinita	H_0 neistinita
Odluka	H_0 ne odbacujemo	Pravilna odluka	Greška II vrste
	H_0 odbacujemo	Greška I vrste	Pravilna odluka

Na osnovu gornje tabele vidimo da ćemo **ispravno postupiti u sljedeća dva slučaja:**

1. Ako je H_0 istinita, a informacija iz uzorka saglasna sa njom, nultu hipotezu nećemo odbaciti i zaključak testiranja će biti ispravan.

2. Ispravno ćemo postupiti i ako odbacimo H_0 koja je neistinita.

Iz Tabele 8.1 vidimo da realno postoji mogućnost da napravimo dvije različite greške u odlučivanju. Prva je odbacimo istinitu nultu hipotezu. Takva greška se naziva greškom I vrste**.** Drugu grešku u testiranju činimo ako ne odbacimo netačnu nultu hipotezu; to je greška II vrste**.** Jasno je da u zaključku možemo da načinimo samo jednu grešku, a ne istovremeno obe.

Nivo značajnosti testa

Nivo značajnosti testa (naziva se još rizikom greške prve vrste) je vjerovatnoća da ćemo odbaciti istinitu nultu hipotezu. Obilježava se sa α.

Vjerovatnoća da nećemo odbaciti netačnu nultu hipotezu naziva se rizikom greške II vrste i obilježava se sa β.

Rizik greške nije isto što i greška. Rizik greške je vjerovatnoća da ćemo napraviti grešku. U zbiru, rizik greške prve vrste i rizik greške

druge vrste nisu jednaki 1.

Na koji način možemo istovremeno smanjiti oba rizika pri testiranju? Statistički rečeno, tražili biste da se poveća uzorak, što je ujedno i odgovor na postavljeno pitanje.

Postupak testiranja sprovodićemo tako što ćemo unapred fiksirati rizik greške I vrste, tj. nivo značajnosti α. Većina istraživača u svijetu koristi uglavnom samo dva nivoa značajnosti: 0,05 i 0,01. Ako odaberemo najčešće korišćeni nivo značajnosti α = 0,05, onda svjesno unaprijed prihvatamo da ćemo u proseku u 5% slučajeva napraviti grešku, odnosno odbaciti istinitu H_0.

Zašto se najčešće koristi nivo značajnosti od 5%, a ne neki drugi broj? Odgovor možemo potražiti u uticaju Ronalda Fishera.

Fisher je smatrao da ako pri testiranju već moramo da se suočimo sa mogučnosću javljanja greške, onda je negdje potrebno da se povuče demarkaciona linija, a on je lično preferirao 5%.

Izbor testa i izračunavanje statistike testa

Umjesto da vodi računa o smanjivanju β, istraživač se najčešće rukovodi komplementarnom vjerovatnoćom, 1- β. **Ova vjerovatnoća naziva se jačinom testa.**

Jačina testa
Jačina testa je vjerovatnoća da se odbaci pogrešna nulta hipoteza. Obilježava se sa π.

Zaključujemo da je optimalan onaj test koji za fiksni nivo značajnosti α ima najveću jačinu. Međutim, jačina testa zavisi od identičnih faktora kao i rizik greške β i njen nivo je skoro nemoguće izračunati u konkretnom uzorku.

Parametarski testovi predstavljaju optimalni izbor istraživača pri rješavanju postavljenog problema, uz uslov da je raspored skupa (tj. skupova) normalan.

Istraživački problem	
Testiranje aritmetičke sredine jednog skupa	Z test (ako je poznata standardna devijacija skupa)
	t test (ako je nepoznata standardna devijacija skupa)
Testiranje proporcije jednog skupa	Z test (ako su ispunjeni uslovi $n\pi_0 > 5$ i $n(1-\pi_0) > 5$)
Testiranje jednakosti aritmetičkih sredina dva skupa	Z test (ako su poznate standardne devijacije oba skupa)
	t test (ako su jednake standardne devijacije skupa)
Testiranje jednakosti proporcija dva skupa	Z test (ako su ispunjeni uslovi $n_1p_1 > 5$ i $n_1(1-p_1) > 5$, $n_2p_2 > 5$ i $n_2(1-p_2) > 5$

Kriterijum testiranja u statistici se naziva statistikom testa.

Statistika testa

Statistika testa je kriterijum na osnovu kojeg vršimo testiranje i na osnovu kojeg u konceptu testiranja na osnovu kritičnih vrijednosti donosimo odluku o odbacivanju ili neodbacivanju nulte hipoteze. U našem "integralnom" konceptu testiranja statistika testa je samo međukorak ka izračunavanju p-vrijednosti.

Grubo rečeno statistika testa je konkretna formula koja se koristi kod pojedinačnog testa.

Logika statističkog testiranja hipoteza

Postavlja se pitanje kako da provjerimo da li je nulta hipoteza istinita, odnosno da li je parametar skupa zaista jednak nekoj pretpostavljenoj vrijednosti? Jasno je da moramo da formulišemo neki indikator koji bi mjerio razliku između naše hipotetične vrijednosti i empirijske evidencije iz uzorka.

Na taj način dolazimo do relativnog pokazatelja koji je nezavisan od mjernih jedinica.

Statistike testa najčešće imaju sljedeći oblik:

$$\text{Statistika testa} = \frac{\text{Ocjena parametra - Hipotetična vrijednost parametra}}{\text{Standardna greška ocjene}}$$

Vidimo da je značaj statistike testa u tome da ona predstavlja mjeru usaglašenosti između podataka i nulte hipoteze.

Statistika testa će uzimati različite vrijednosti od uzorka do uzorka i te vrijednosti je nemoguće predvidjeti. Ona je slučajna promjenljiva i kao takva ima svoj raspored vjerovatnoća. Ako nam je taj raspored poznat tada možemo izračunavati različite vjerovatnoće realizacije različitih statistika testa, kao i one dobijene u konkretnom uzorku.

Mala vrijednost statistike testa nam ne daje osnove da odbacimo nultu hipotezu. Nasuprot tome, njene velike vrijednosti nam sugerišu da nulta hipoteza nije tačna. Što je veća vrijednost statistike testa uvjerljiviji su i dokazi da je nulta hipoteza pogrešna.

Kako odrediti kolika treba da bude statistika testa da bi se nulta hipoteza odbacila?

Po Newman-Pearson-ovom konceptu realizovanu vrijednost statistike testa poredimo sa vrijednostima iz tablica onog rasporeda koji slijedi ta statistika testa. Takve tablične vrijednosti nazivaju se kritične vrijednosti (ili pragovi značajnosti), a oblasti koje one formiraju kritične oblasti.

Ako statistika testa ima normalan raspored koristićemo naravno tablice standardizovanog normalnog rasporeda (jer statistika testa je sama po sebi standardizovana veličina), ako slijedi *t* raspored, kritične vrijednosti ćemo potražiti u tablici *t* rasporeda.

Shodno tome pravilo odlučivanja ovdje se postavlja poređenjem realizovane vrijednosti statistike testa sa kritičnom vrijednošću.

Preciznije, u zavisnosti od oblika testa (jednosmjernog ili dvosmjernog) postoje čak tri pravila odlučivanja.

A. Ako je test dvosmjeran tada se posmatra apsolutna vrijednost statistike testa. **Ako je ona veća od tablične odbacuje se nulta hipoteza i usvaja alternativna.**

B. Ako je test jednosmjeran a alternativna hipoteza ljevostrana **tada statistika testa treba da bude manja od kritične vrijednosti** (pri svemu ovome treba voditi računa da se kritičnoj vrijednosti doda negativan znak).

C. U slučaju jednosmjernog, desnostranog testa, **statistika testa mora da bude veća od tablične vrijednosti** da bi se nulta hipoteza odbacila.

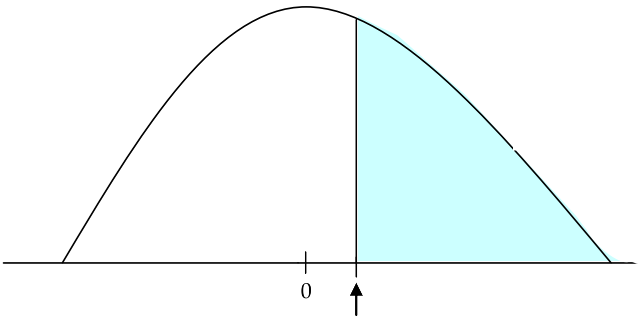
Zbog toga je bolje statistiku testa transformisati na odgovarajući način u neku drugu veličinu, tako da svako, bez gledanja u tablice, bez obzira na oblik testa, može odmah da donese odluku o nultoj hipotezi.

To se upravo postiže preko *p*-vrijednosti, koja konvertuje vrijednost

bilo koje statistike testa na skalu od 0 do 1. Razumijevanje *p*-vrijednosti studentima pruža life-long (doživotno) sredstvo da shvate kompjuterski izlaz svakog statističkog testa i da po krajnje jednostavnom pravilu donesu odluku o tačnosti nulte hipoteze.

p-vrijednost

Da bismo formulisali takvu mjeru moramo da poznajemo raspored vjerovatnoće javljanja svih mogućih vrijednosti aritmetičke sredine uzorka, odnosno u generalnom slučaju statistike testa. Do njega bi se došlo ako bismo, uz pretpostavku da je nulta hipoteza tačna, izvukli sve moguće uzorke iz skupa i formirali raspored tako dobijenih statistika. Takav raspored u statistici se naziva nulti raspored statistike testa. Jedan takav raspored prikazan je na Slici 8.1.



Slika 8.1 Nulti raspored proizvoljne statistike testa

Da sumiramo: što je ta površina manja sve je manje vjerovatno da je nulta hipoteza tačna, i obrnuto. Ta površina, preciznije, vjerovatnoća, u stvari, je *p*-vrijednost.

p-vrijednost

p-vrijednost je vjerovatnoća, izračunata pod pretpostavkom da je nulta hipoteza tačna, da statistika testa uzme vrijednost koja je toliko ekstremna ili još ekstremnija od one upravo realizovane. Što je manja p-vrijednost jači su dokazi protiv nulte hipoteze.

Što je *p*-vrijednost manja, podaci su manje konzistentni sa nultom hipotezom. U tom smislu kažemo da *p*-vrijednost mjeri jačinu dokaza protiv nulte hipoteze.

Pravilo odlučivanja na osnovu *p*-vrijednosti glasi:

Ako je *p*-vrijednost manja od nivoa značajnosti α nulta hipoteza se odbacuje. U suprotnom, kazaćemo da nemamo dovoljno argumenata da odbacimo H_0.

Zbog čestih grešaka u praksi u tumačenju *p*-vrijednosti, potrebno je voditi računa o sljedećem:

a) p-vrijednost nije vjerovatnoća da je nulta hipoteza tačna.

b) p-vrijednost ništa ne govori o veličini razlike između hipotetične vrijednosti i dobijenog rezultata na osnovu uzorka.

Ako je p-vrijednost manja od nivoa značajnosti α kažemo da je dobijeni rezultat statistički značajan ili signifikantan. U suprotnom, kažemo da rezultat nije statistički značajan.

Statistička značajnost

Ako je *p*-vrijednost jednaka ili manja od nivoa značajnosti α kažemo da je rezultat statistički značajan na nivou α.

Značajnost na nivou od, recimo, 0,01 obično se izražava izjavom "rezultati su statistički značajni ($p < 0{,}01$)". Praktična značajnost, međutim, zavisi od veličine razlike.

Jedan broj statističara smatra da bi postupak testiranja trebalo da se završi sa izračunavanjem *p*-vrijednosti. Po njima, umesto da se donosi odluka o odbacivanju ili neodbacivanju nulte hipoteze poređenjem *p*-vrijednosti sa nivoom značajnosti, treba jednostavno iskazati dobijenu *p*-vrijednost i prepustiti svakome da je lično protumači po svom nahođenju.

Tabela 8.2 Interpretacija jačine argumenata protiv nulte hipoteze

p-vrijednost > 0,05	Nemamo dovoljno dokaza protiv nulte hipoteze
p-vrijednost < 0,05	Jaki dokazi da je H_0 pogrešna
p-vrijednost < 0,01	Veoma jaki dokazi da je H_0 pogrešna
p-vrijednost < 0,001	Izuzetno jaki dokazi da je H_0 pogrešna

TESTIRANJE HIPOTEZE ZASNOVANO NA JEDNOM UZORKU

Testiranje hipoteze o aritmetičkoj sredini osnovnog skupa

U ovoj situaciji mogu se koristiti dva parametarska testa, *Z* i *t*. Međutim, *Z* test zahtijeva poznavanje standardne devijacije osnovnog skupa.

Iz poglavlja o deskriptivnim mjerama nam je jasno da je nemoguće izračunati standardnu devijaciju skupa ako nemamo vrijednost aritmetičke sredine skupa.

Da li to znači da *Z* test ima samo akademsku, i nikakvu praktičnu vrijednost?

Smatramo da Z test ipak treba prikazati jer ćemo se kasnije susresti sa nekoliko primjera u kojima je potrebno poznavanje logike ovog testa. Takođe, Z test ćemo iskoristiti da pokažemo kako se izračunava p-vrijednost. Kod svih ostalih testova *p*-vrijednost ćemo samo određivati na osnovu tablica. Naravno u praksi *p*-vrijednost nikada nećemo izračunavati već konsultovati kompjuterski izlaz.

Z test

Osnovne informacije vezane za primjenu Z testa u slučaju testiranja aritmetičke sredine jednog skupa prikazane su u Tabeli 8.3.

Tabela 8.3 Testiranje hipoteze o aritmetičkoj sredini skupa primjenom Z testa

Uslovi za primjenu Z testa: 1) Osnovni skup ima normalan raspored ili Može se primjeniti Centralna Granična Teorema ($n > 30$) 2) Standardna devijacija skupa, σ, je poznata

Izbor oblika nulte i alternativne hipoteze**:**

A) $H_0 : \mu = \mu_0$, $H_1 : \mu \neq \mu_0$

B) $H_0 : \mu \leq \mu_0$, $H_1 : \mu > \mu_0$

C) $H_0 : \mu \geq \mu_0$, $H_1 : \mu < \mu_0$

gdje je μ_0 hipotetična vrijednost aritmetičke sredine skupa.

Statistika testa**:** $Z = \dfrac{\bar{X} - \mu_0}{\sigma_{\bar{X}}}$

pod uslovom da je H_0 istinita ima standardizovan normalan raspored, $Z : N(0, 1)$.

RJEŠENJE PROBLEMA:

Etapa 1 Postavljanje nulte, alternativne hipoteze i određivanje nivoa značajnosti.

Etapa 2 Uzima se slučajan uzorak, bira odgovarajući test i izračunava statistika testa

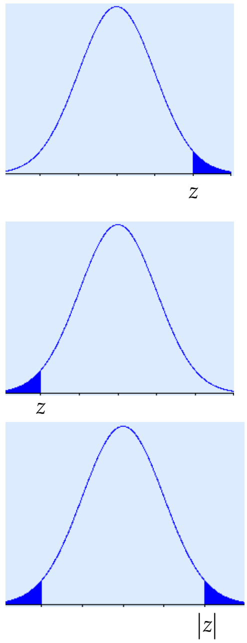
Etapa 3 Provjeravanje ispunjenosti uslova izabranog testa

Z test se zasniva na dva uslova. Provjeravamo da li su ispunjeni.

1) Osnovni skup ima normalan raspored ili
   Može se primjeniti Centralna Granična Teorema ($n > 30$)
2) Standardna devijacija skupa, σ, je poznata

Etapa 4 Izračunavanje p-vrijednosti.

Alternativna hipoteza	**p-vrijednost**
$H_1 : \mu > \mu_0$	$P(Z \geq z)$
$H_1 : \mu < \mu_0$	$P(Z \leq z)$
$H_1 : \mu \neq \mu_0$	$2P(Z \geq \lvert z \rvert)$



Slika 8.2 Izračunavanje *p*-vrijednosti kod Z testa

Ako je, npr. Z= - 2,324, p-vrijednost ćemo izračunati na sljedeći način:

$$2P(Z \geq |z|) = 2P(Z \geq |-2,324|) = 2P(Z \geq 2,324) = 2[1 - P(Z < 2,324)]$$
$$= 2[1 - 0,9898] = 0,0204$$

Etapa 5 Postavljanje pravila odlučivanja i donošenje odluke o nultoj hipotezi.

Pravilo odlučivanja kod korišćenja p-vrijednosti glasi:

Ako je *p*-vrijednost < nivoa značajnosti, H_0 se odbacuje, u suprotnom se ne odbacuje.

Etapa 6 Donošenje zaključka

Studentov *t* test

Osnovne informacije vezane za primjenu *t* testa u slučaju testiranja aritmetičke sredine jednog skupa prikazane su u Tabeli 8.4.

Tabela 8.4 Testiranje hipoteze o aritmetičkoj sredini skupa primjenom t testa

Uslovi za primjenu *t* testa:

1) Osnovni skup je normalno raspoređen ili
2) Može se primjeniti Centralna granična teorema (uzorak ima više od 30 elemenata)

Izbor oblika nulte i alternativne hipoteze:

A) $H_0: \mu = \mu_0$, $H_1: \mu \neq \mu_0$

B) $H_0: \mu \leq \mu_0$, $H_1: \mu > \mu_0$

C) $H_0: \mu \geq \mu_0$, $H_1: \mu < \mu_0$

gdje je μ_0 hipotetična vrijednost aritmetičke sredine skupa

Statistika testa: $$t = \frac{\bar{X} - \mu_0}{S_{\bar{X}}} = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

pod uslovom da je H_0 istinita ima Studentov raspored sa n - 1 stepeni slobode.

U praksi nam standardna devijacija populacije nikada nije poznata. Usljed toga ne možemo ni koristiti statistiku Z testa jer u formuli za standardnu grešku, $\sigma_{\bar{X}} = \sigma/\sqrt{n}$, figuriše nepoznata standardna devijacija skupa σ.

Da bi postupak testiranja bio moguć, standardnu devijaciju σ ćemo ocjeniti na osnovu standardne devijacije uzorka *S*, a ovu ocjenu ćemo zatim upotrijebiti za ocjenjivanje standardne greške.

Dakle, kao kod postupka ocjenjivanja nepoznate vrijednosti μ, i ovde ćemo umjesto nepoznate $\sigma_{\bar{X}}$ koristiti njenu ocjenu, $S_{\bar{X}}$.

Tako dobijamo novu statistiku testa, tzv. statistiku t testa:

Statistika t testa

$$t = \frac{\overline{X} - \mu_0}{S_{\overline{X}}} = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}$$

Ako je nulta hipoteza istinita, statistika t ima Studentov raspored sa n - 1 stepeni slobode.

Određivanje p-vrijednosti:

Vjerovatnoća (površina na desnom kraju t rasporeda)	0,10	0,05	0,025	0,01	0,005
Kritična t vrijednost	1,341	1,7531	2,1315	2,602	2,9467

Vjerovatnoće u zaglavlju pokazuju kolika je površina ispod krive t rasporeda udesno od kritične vrijednosti. Ovo će nam direktno omogućiti, da bez ikakvih izračunavanja odredimo p-vrijednost, na osnovu njene definicije. Podsjetimo se da p-vrijednost pokazuje vjerovatnoću da se realizuje statistika testa koja je ekstremna ili više ekstremna od dobijene, uz uslov da je H_0 tačna.

Kako se konkretno određuje *p*-vrijednost?

Najprije određujemo red u tabeli koji se odnosi na traženi broj stepeni slobode. Nakon toga upoređujemo našu statistiku testa sa tabličnim vrijednostima da bismo locirali između kojih vrijednosti je ona "pala".

Neka statistika testa iznosi, npr. 4,6875, pa je veća od najveće

vrijednosti, 2,9467.

Da bismo pronašli *p*-vrijednost potražimo koja je vjerovatnoća u zaglavlju asocirana našoj statistici testa. Pažljivom inspekcijom vrijednosti u zaglavlju vidimo da se one smanjuju sa lijeva udesno, što je i logično jer se sa povećanjem statistike testa smanjuje i vjerovatnoća njene realizacije.

Koliko onda iznosi *p*-vrijednost, da li *p*-vrijednost < 0,005? Ne, jer je test dvosmjeran, dobijenu vjerovatnoću moramo (kao kod Z testa) pomnožiti sa 2.

Dakle, konačno rješenje je: *p*-vrijednost < 0,01.

Testiranje hipoteze o proporciji osnovnog skupa

Osnovne informacije vezane za primjenu Z testa u slučaju testiranja proporcije jednog skupa prikazane su u Tabeli 8.5.

Tabela 8.5 Testiranje hipoteze o proporciji skupa primjenom *Z* testa

Uslovi za primjenu *Z* testa:

$n \cdot P_0 \geq 5$, $n \cdot (1 - P_0) \geq 5$ i $n \geq 30$

Izbor oblika nulte i alternativne hipoteze:

A) $H_0 : \pi = \pi_0$ $H_1 : \pi \neq \pi_0$

B) $H_0 : \pi \leq \pi_0$ $H_0 : \pi > \pi_0$

C) $H_0 : \pi \geq \pi_0$ $H_0 : \pi < \pi_0$

gdje je P_0 hipotetična proporcija skupa.

Statistika testa: $Z = \dfrac{p - \pi_0}{S_p}$ $S_p = \sqrt{\dfrac{\pi_0 (1 - \pi_0)}{n}}$,

pod uslovom da je H_0 istinita ima standardizovan normalan raspored, $Z : N(0, 1)$.

Logika statistike testa o proporciji skupa je analogna onoj kod aritmetičke sredine. Razlika je u tome što raspored proporcije kod malih uzoraka nema normalan raspored.

Preciznije rečeno proporcija uzorka ima binomni raspored (kod uzoraka sa ponavljanjem) ili hipergeometrijski raspored (kod uzoraka bez ponavljanja). Navedeni rasporedi mogu se koristiti kada se testiranje vrši pomoću malih uzoraka.

Međutim, mi ćemo **testirati proporcije samo kod velikih uzoraka**. Tada se raspored vjerovatnoće proporcije *p*, ako su ispunjeni potrebni uslovi, može aproksimirati normalnim rasporedom, *p*: približno N(π_0, σ_p).

Generalna formula za statistiku testa sugeriše i oblik koji će statistika testa imati:

$$Z = \frac{\text{Proporcija uzorka - Hipotetična vrijednost proporcije}}{\text{Standardna greška proporcije}} = \frac{p - \pi_0}{\sigma_p}$$

Statistika, pod uslovom da je H_0 istinita, ima standardizovan normalan raspored, *Z*: *N*(0, 1). Ovo zbog toga što svaka normalna promjenljiva (poput *p*) kada se standardizuje preko izraza, transformiše u standardizovanu normalnu promjenljivu *Z*.

Primjetimo da **u formuli umjesto nepoznate standardne greške proporcije (σ_p) koristimo njenu ocjenu (S_p).**

Važno je da napomenemo da **ocjenu standardne greške proporcije**

kod testiranja izračunavamo primjenom hipotetične vrijednosti parametra (π_0), a ne kao kod ocjenjivanja na osnovu realizovane vrijednosti proporcije (p) u uzorku:

Standardna greška proporcije	$S_p = \sqrt{\dfrac{\pi_0\ (1\ -\ \pi_0)}{n}}$

HI KVADRAT (χ^2) TEST

Testiranje prilagođenosti
Analiza tabela kontingencije
Test jednakosti (razlike) proporcija više skupova
Uslovi za primjenu χ^2 testa

CILJEVI POGLAVLJA

Nakon čitanja ovog poglavlja bićete u stanju da:

14. navedete osnovne karakteristike χ^2 rasporeda
15. testirate da li osnovni skup ima uniformni raspored
16. testirate da li osnovni skup ima normalan raspored
17. na objektivan način ispitate da li su dva obilježja sa modalitetima datim na nominalnoj skali međusobno nezavisna
18. izmjerite jačinu veze između dva obilježja (varijable) sa modalitetima datim u nivou nominalne mjerne skale
19. testirate da li tri ili više populacija imaju jednake proporcije
20. shvatite ograničenja i uslove za primjenu χ^2 testa

χ^2 test zasnovan je na χ^2 distribuciji i koristi se tipski za rješavanje nekoliko problema:

- za testiranje značajnosti razlike između opaženih i teorijskih f rekvencija različitih rasporeda vjerovatnoće – test prilagođenosti;
- kod takozvanih tabela kontingencije i odnosi se na testiranje međusobne povezanosti različitih obilježja posmatrane pojave.
- za testiranje jednakosti (ili razlike) proporcija tri i više skupova, što se označava i kao testiranje homogenosti posmatrane pojave.

U sva tri navedena slučaja, **χ^2 test ima odlike neparametarskog testa.**

Osnovni koraci prilikom testiranja χ^2 testom mogu se svesti na sljedeće:

1. Ustanovljava se nulta i alternativna hipoteza.
2. Izračunavaju se teorijske frekvencije pojavljivanja neke osobine kod posmatrane populacije u skladu sa postavljenom nultom hipotezom.
3. Kod tabela kontingencije, različite opservacije – frekvencije razmještaju se u različite ćelije.
4. Određuje se razlika između opaženog i očekivanog, tako da se izračunava vrijednost χ^2 statistike testa, date izrazom:

$$\chi^2 = \sum_{i=1}^{r} \frac{(f_i - f_i^*)^2}{f_i^*}$$

   U prethodnom izrazu sa f_i su označene opažene – empirijske frekvencije, a sa f_i^* očekivane ili teorijske frekevencije. r predstavlja broj grupa frekvencija.

5. Ustanovljava se odgovarajuća p-vrijednost, odnosno poredi se izračunata vrijednost statistike testa sa kritičnim (tabličnim) vrijednostima iz χ^2 distribucije, uz odgovarajući broj stepeni slobode, i izvodi zaključak.

Broj stepeni slobode χ^2 distribucije određuje se posebno za svaki slučaj primjene.

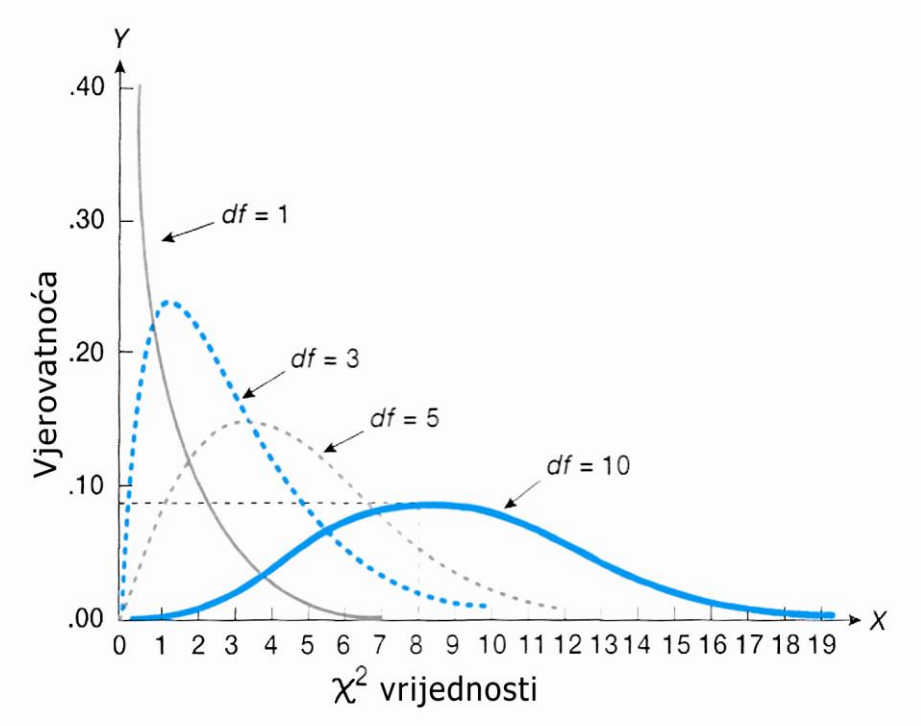
χ^2 distribucija

Osnovne karakteristike χ^2 distribucije su:

1. Izračunata vrijednost χ^2 testa uvijek je pozitivna, zbog kvadratnog izraza $(f_i - f_i^*)^2$.

2. Postoji familija χ^2 distribucija u zavisnosti od broja stepeni slobode. Broj stepeni slobode u većini slučajeva zavisi od broja grupa frekvencija ili od broja ćelija u tabeli kontingencije, a ne od broja elemenata u uzorku. Zbog toga i oblik χ^2 distribucija ne zavisi od broja elemenata u uzorku.

3. χ^2 distribucija je pozitivno asimetrična. Povećavanjem broja stepeni slobode ova distribucija se približava normalnoj, tako da već za 10 stepeni slobode uzima oblik približno normalne distribucije.



TEST JEDNAKOSTI (RAZLIKE) PROPORCIJA VIŠE SKUPOVA

Proporcije posmatranih populacija mogu biti međusobno jednake, a najčešće se razlikuju, u manjoj ili većoj mjeri. Zbog toga ima smisla govoriti **o testiranju i razlike i jednakosti proporcija, tako da nije greška ako se koristi jedan ili drugi termin.**

Primjenom χ^2 testa mogu se dobiti odgovori na prethodno postavljena pitanja. Test jednakosti (razlike) proporcija za više populacija naziva se i test homogenosti.

Postupak testiranja u osnovi je potpuno isti kao u prethodnim slučajevima. Za svaku pojedinu populaciju uzima se u obzir opažena (empirijska) frekvencija kao broj (pojavljivanja) elemenata sa određenom osobinom u toj populaciji.

Zatim se izračunavaju očekivane (teorijske) frekvencije, pod pretpostavkom da nema značajne razlike (da postoji jednakost) između posmatranih populacija prema učešću elemenata sa određenom osobinom. U daljem toku analize primjena χ^2 testa je potpuno ista kao u prethodno izloženim postupcima.

Prilikom poređenja *c* populacija (ili *r* populacija, ako su uređene kao redovi tabele) nulta i alternativna hipoteza glase:

$H_0 : \pi_1 = \pi_2 = . . . = \pi_c$

H_1 : sve proporcije π_i , $i = 1, 2, ..., c$ nisu međusobno jednake

Testiranje ovako postavljenih hipoteza predstavlja generalizaciju testiranja jednakosti proporcija dva skupa u slučaju kada je broj populacija veći od 2.

U opštem slučaju, teorijske frekvencije dobiju se tako da se ustanovi

opšta (generalna proporcija) kao odnos zbira broja elemenata iz svih uzoraka koji imaju određenu osobinu ili su na neki način kategorizovani i zbira broja elementa svih uzoraka.

Zatim se tako dobijena proporcija primjenjuje (jednostavno se pomnoži) na svaki pojedini uzorak i tako se izračunavaju teorijske frekvencije, čiji zbir je jednak zbiru originalnih (opaženih) frekvencija.

U ovom domenu primjene tablična vrijednost određuje se iz tablice χ^2 rasporeda uz $df = (r - 1) = (8 - 1) = 7$ stepeni slobode i odgovarajući nivo rizika.

PROSTA LINEARNA KORELACIONA I REGRESIONA ANALIZA

- Funkcionalna i stohastička veza

- Razlika između regresione i korelacione analize

- Dijagram raspršenosti

- Prosta linearna korelacija

- Prosta linearna regresija

Izučavanje međuzavisnost dvije ili više kvantitativnih promjenljivih sprovodimo pomoću dva, vjerovatno najvažnija statistička metoda, korelacije i regresije.

Logika regresione i korelacione analize je identična onoj kod ocjenjivanja i testiranja hipoteza: iz osnovnog skupa uzima se reprezentativan uzorak i na osnovu uzorka donosi se zaključak o parametrima osnovnog skupa.

Prilikom istraživanja međusobnih veza dvije promjenljive primjenjuju se metodi proste regresione i korelacione analize, a u slučaju posmatranja više promjenljivih, metodi višestruke regresije i korelacije.

FUNKCIONALNA I STOHASTIČKA VEZA

Međusobne veze između pojava (promjenljivih) možemo podijeliti u dvije grupe: funkcionalne i stohastičke.

Funkcionalna (naziva se još i deterministička ili egzaktna) veza javlja se u slučaju kada jednoj vrijednosti nezavisne promjenljive *X* odgovara samo jedna, tačno određena vrijednost zavisne promjenljive Y.

Slabija od funkcionalne veze je stohastička (ili probabilistička) veza. Kod stohastičkih veza jednoj vrijednosti nezavisne promjenljive odgovara čitav niz mogućih vrijednosti zavisne promjenljive.

Generalno, stohastički model, odnosno veza, može se prikazati na sljedeći način:

Generalna forma stohastičkog (probabilističkog) modela

Y = Deterministički član + stohastički član

Budući da stohastički član u različitim situacijama djeluje na slučajan način, nekada tako što utiče pozitivno na Y, nekada negativno, pretpostavićemo da je u prosjeku njegov uticaj jednak nuli. Zaključujemo da je i *Y* slučajna promjenljiva.

Kako je prosjek stohastičkog člana jednak nuli, zaključujemo da je:

Prosjek Y = Deterministički član

Ako posmatramo samo dvije pojave, tada se 11.4 može napisati na sljedeći način:

Prosjek $Y = f(X)$

Suština stohastičke veze jeste da između pojedinih vrijednosti nezavisne promjenljive *X* i prosječnih vrijednosti zavisne promjenljive *Y* (preciznije, očekivanih vrijednosti) postoji čvrsta, odnosno funkcionalna veza.

Veze kod kojih porastû (opadanju) vrijednosti nezavisne promjenljive *X* istovremeno odgovara porast (opadanje) zavisne promjenljive *Y* nazivamo direktnim vezama.

Sa druge strane, ako porastû jedne promjenljive odgovara opadanje druge, radi se o inverznim vezama (na primjer, sa porastom cijene avionskih karata opada broj putnika, uz konstantni realni dohodak).

U stvarnosti, između dvije ili više pojava moguće je postojanje najrazličitijih oblika veza, počev od onih koje se matematički mogu iskazati jednostavnom formulom, pa do onih veoma kompleksnih. Najjednostavniji oblik veze između pojava je linearna veza.

RAZLIKA IZMEĐU REGRESIONE I KORELACIONE ANALIZE

Prilikom istraživanja međuzavisnosti varijacija dvije ili više numeričkih varijabli, u statistici se primjenjuju metodi regresione i korelacione analize. Iako su ovi statistički metodi u bliskoj vezi i međusobno se dopunjuju, između njih postoje i značajne razlike.

Kod korelacije pri analizi dvije pojave svejedno je koja se označava kao nezavisno, a koja kao zavisna promjenljiva - dobija se identičan rezultat. Međutim, kod ispitivanja korelacione veze između tri ili više pojava mora se prethodno jedna od njih definisati kao zavisna promjenljiva, dok ostale dobijaju status nezavisnih promjenljivih.

Kod regresione analize nužno je unaprijed identifikovati koja pojava će imati ulogu zavisne promjenljive, a koja nezavisne promjenljive.

Da budemo precizniji, u statistici se kod regresije ne koristi termin "nezavisna promjenljiva", već objašnjavajuća promjenljiva ili regresor.

Cilj korelacione analize

Cilj korelacione analize jeste da se ispita da li između varijacija posmatranih pojava postoji kvantitativno slaganje i, ako postoji, u kom stepenu.

Svrha regresije jeste da se utvrdi oblik veze, odnosno zavisnosti između posmatranih pojava. To se postiže pomoću odgovarajućeg regresionog modela.

Regresioni model je takav stohastički model koji kroz matematičku formulu i niz odgovarajućih pretpostavki najbolje opisuje kvantitativnu zavisnost između varijacija posmatranih pojava u

realnosti. Regresioni model pokazuje prosječno slaganje varijacija ispitivanih pojava.

Regresioni model nije sam po sebi cilj, već samo sredstvo pomoću kojeg smo u stanju da ocijenimo i predvidimo vrijednosti zavisne promjenljive za željene vrijednosti objašnjavajuće promjenljive.

Cilj regresione analize

Cilj regresione analize je da se odredi regresioni model koji najbolje opisuje vezu između pojava i da se na osnovu toga modela ocijene i predvide vrijednosti zavisne promjenljive Y za odabrane vrijednosti objašnjavajuće promjenljive X.

Važno je napomenuti da pomoću regresije i korelacije nismo u stanju da otkrijemo da li između pojava postoji uzročno-posljedična veza, u smislu da je jedna pojava uzrok, a druga posljedica.

DIJAGRAM RASPRŠENOSTI

Nakon identifikacije promjenljivih, kao prvi sljedeći korak u analizi međuzavisnosti dvije pojave, podatke ćemo prikazati grafički, pomoću specijalnog dijagrama koji se naziva dijagram raspršenosti. Preporučujemo da se obavezno, prije bilo kakve kvantitativne analize, podaci prikažu na ovom dijagramu.

Na osnovu svega navedenog možemo zaključiti da **dijagramom raspršenosti grafički prikazujemo varijacije dvije pojave u cilju sagledavanja:**

1) **da li između njih postoji kvantitativno slaganje,**
2) **ako slaganje postoji, koji je njegov oblik (linearni ili krivolinijski),**
3) **koji je smjer slaganja (direktni ili inverzni), i**
4) koja je jačina slaganja.

PROSTA LINEARNA KORELACIJA

Podsjetimo se da je svrha korelacione analize da se utvrdi da li između varijacija posmatranih pojava postoji kvantitativno slaganje (korelaciona veza) i, ako postoji, u kom stepenu. Ako se pri tome posmatraju dvije pojave, govori se o prostoj korelaciji, a prilikom analize više pojava o višestrukoj korelaciji.

Kod proste korelacije moguće je ispitivati da li između pojava postoji linearna, krivolinijska ili monotona veza. Mi ćemo se ograničiti samo na linearne veze i govorićemo o prostoj linearnoj (pravolinijskoj) korelaciji.

Koeficijent proste linearne korelacije uzorka, *r*

Za razliku od regresione analize, kod proste linearne korelacije se ne pravi razlika između zavisne i nezavisne promjenljive - obje pojave dobijaju jednaki status.

Preciznije rečeno, obje posmatrane pojave tretiraju se kao slučajne promjenljive. Dakle, potpuno je svejedno koju pojavu ćemo označiti kao X, a koju kao Y, pošto se dobijaju identični rezultati.

Zadatak proste linearne korelacije jeste mnogo jednostavniji nego kod regresije: da pokaže da li između varijacija dvije pojave postoji pravolinijska veza. Zbog toga se kao cilj ispitivanja neće postaviti istraživanje jedne pojave u funkciji druge.

Karl Pearson
(1857 - 1936)

Kao mjera jačine proste linearne korelacione veze u uzorku koristi se relativna mjera, koja se naziva Pearson-ov koeficijent proste linearne korelacije, ili koeficijent proste linearne korelacije, ili često samo koeficijent korelacije. Formulisao ga je Karl Pearson 1896. godine.

Ovaj koeficijent pokazuje stepen pravolinijskog kvantitativnog slaganja dvije pojave. Označava se sa *r* i izračunava po formuli:

Koeficijent proste linearne korelacije u uzorku

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{n\sum x^2 - (\sum x)^2}\sqrt{n\sum y^2 - (\sum y)^2}}$$

Primjećujemo da je navedena formula simetrična u odnosu na promjenljive *X* i *Y*. Dakle, potpuno je svejedno koju smo promjenljivu označili sa *X*, a koju sa *Y*.

Pearson-ov koeficijent proste linearne korelacije, *r*, pokazuje stepen linearnog (pravolinijskog) kvantitativnog slaganja varijacija između dvije numeričke promjenljive (obilježja).

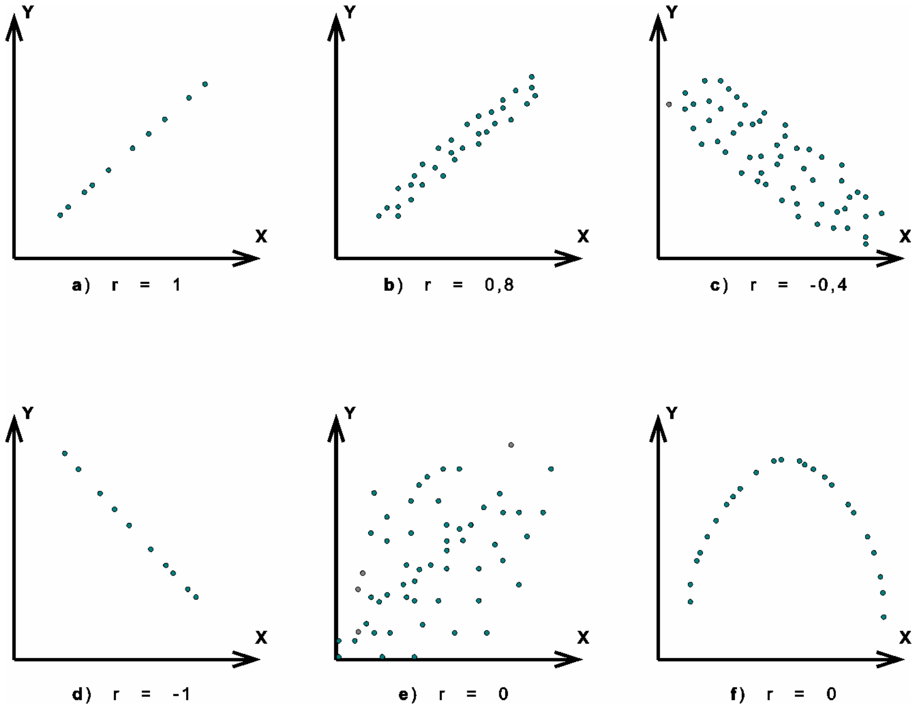
Koeficijent proste linearne korelacije, kao relativna mjera, uzima vrijednosti od -1 do +1. Ukoliko uzima pozitivne vrijednosti, korelacija između pojava je direktna ili pozitivna (obje pojave pokazuju istosmjerne varijacije). U slučaju kada je $r < 0$, veza je inverzna ili negativna (kada jedna pojava raste druga opada, i obrnuto).

Ako između posmatranih pojava postoji funkcionalna veza (sve empirijske tačke se nalaze tačno na pravoj liniji), govorimo o savršenoj (perfektnoj) korelaciji. Tada koeficijent korelacije uzima vrijednost -1 (ako je veza inverzna) ili +1 (ako je veza direktna). Što je koeficijent korelacije po apsolutnoj vrijednosti bliži jedinici, sve je jača korelaciona veza između pojava.

Nasuprot tome, što je bliži nuli linearna veza je slabija.

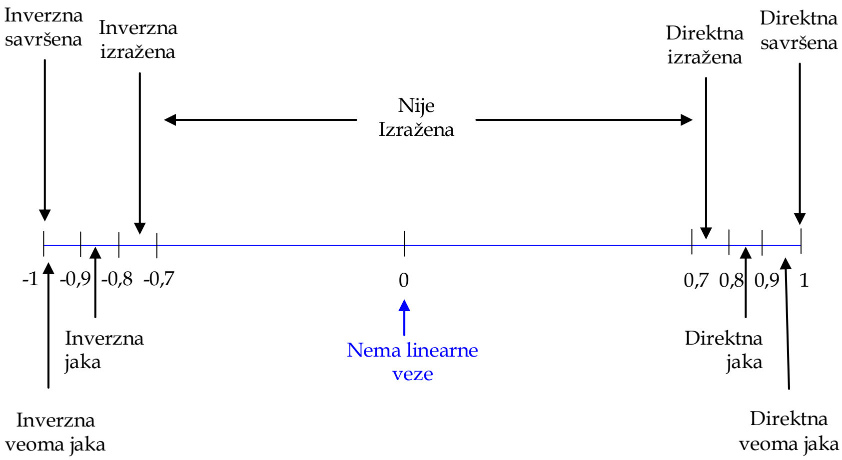
U ekstremnoj situaciji, kada koeficijent korelacije uzme vrijednost jednaku nuli, zaključuje se da nema linearne veze između pojava.

Obrnuto, ne važi. Dakle, kada se na osnovu uzorka dobije koeficijent korelacije jednak nuli, pogrešno je zaključiti da između dvije pojave ne postoji kvantitativno slaganje.



Slika 11.4 Raspršenost tačaka i odgovarajuće vrijednosti r

U statističkoj literaturi ne postoji potpuno slaganje u pogledu tumačenja značenja pojedinih mogućih vrijednosti koeficijenta proste linearne korelacije.



Slika 11.5 Tumačenje vrijednosti *r*

Pošto se prilikom izračunavanja koeficijenta korelacije *r* koriste podaci slučajnog uzorka, važno je shvatiti da *r* ukazuje samo na postojanje korelacije u uzorku.

Međutim, nas interesuje da li u osnovnom skupu iz koga potiče uzorak postoji korelaciona veza? Lako je zaključiti da *r* predstavlja ocjenu nepoznatog koeficijenta korelacije u osnovnom skupu. Stoga je potrebno testirati značajnost dobijene ocjene.

Testiranje značajnosti ocjene koeficijenta proste linearne korelacije

Koeficijent proste linearne korelacije u osnovnom skupu označava se sa grčkim slovom ρ (čitaj ro). On pokazuje jačinu pravolinijske veze između dvije posmatrane pojave u populaciji.

Budući da je on numerički pokazatelj skupa, jasno nam je da se radi o parametru. Njegove pojedinačne vrijednosti se, naravno, tumače istovjetno kao i vrijednosti koeficijenta korelacije u uzorku *r*.

Pretpostavka je da je raspored promjenljive *X* i *Y* normalan. Zbog toga je jasno da ćemo primijeniti parametarski test.

Nultu hipotezu postavićemo u obliku:

$$H_0 : \rho = 0,$$

odnosno, da u osnovnom skupu ne postoji *linearna* korelacija, ili, što je isto, da ocjena, *r*, *nije* statistički značajna. Ograničićemo se na dvosmjernu alternativnu hipotezu:

$$H_1 : \rho \neq 0.$$

Dakle, alternativna hipoteza ukazuje samo na to da u skupu postoji linearna veza, a ne govori ništa o jačini veze.

Odaberimo nivo značajnosti α = 0,05.

Postupak testiranja se svodi najprije na formulisanje odgovarajuće statistike testa.

Da bismo lakše shvatili formule za različite statistike testa, koje ćemo koristiti u ovom i sljedećem poglavlju, navedimo da se veoma često statistika testa može napisati u vidu:

$$\boxed{\text{Statistika testa} = \frac{\text{Ocjena - Hipotetična vrijednost parametra}}{\text{Standardna greška ocjene}}}$$

U ovom konkretnom primjeru ocjena parametra je *r*, hipotetična vrijednost je 0, a standardnu grešku od *r* ćemo označiti sa s_r. U statistici je pokazano da ta statistika testa slijedi Studentov, *t*, raspored sa *n* – 2 stepeni slobode.

Prema tome, izraz za statistiku testa glasi:

$$t = \frac{r}{s_r}$$

gdje je s_r standardna greška ocjene koeficijenta proste linearne korelacije.

Pri njenom izračunavanju koristi se formula:

Standardna greška ocjene koeficijenta proste linearne korelacije	$s_r = \sqrt{\dfrac{1-r^2}{n-2}}$

Šta pokazuje ova standardna greška? Podsjetimo se: svaka standardna greška u statistici pokazuje prosjek odstupanja ocjene od parametra.

Dakle, s_r pokazuje koliko u prosjeku koeficijent korelacije uzorka odstupa od koeficijenta korelacije skupa.

Interpretacija koeficijenta proste linearne korelacije

Pravilna interpretacija koeficijenta proste linearne korelacije zahtijeva dopunska objašnjenja i ograničenja, naročito u pogledu eventualne uzročne veze posmatranih pojava.

Ovo posebno napominjemo zbog činjenice da je **u praksi koeficijent korelacije, uz aritmetičku sredinu, statistički pokazatelj koji se često pogrešno tumači.**

1. Koeficijent proste korelacije r govori samo to da li u uzorku postoji korelacija.
2. ***r* zahtijeva numeričke podatke**; preciznije, da se promjenljive mogu mjeriti bar na intervalnoj mjernoj skali.
3. ***r* je relativna mjera**, a to znači da nije iskazan u mjernim jedinicama originalnih pojava.
4. Postojanje korelacije ukazuje samo na opšte slaganje varijacija dvije pojave i nikako ne važi za sve pojedinačne slučajeve.
5. **Nekorektna primjena** korelacione analize je u slučaju kada dobijeni koeficijent korelacije upućuje na pogrešni smjer veze (na primjer, direktnu vezu umjesto inverzne). To se može desiti ako

primijenimo korelacionu analizu na dvije vremenske serije.

6. **Statistika testa se smije koristiti** samo pri testiranju nulte hipoteze da u skupu nema korelacije. Da bi se testirala hipoteza da koeficijent korelacije u skupu ima neku određenu vrijednost (npr. $H_0 : \rho = 0,8$) mora da se koristi drugačija statistika testa, ali to je izvan okvira ove knjige.
7. Važno je naznačiti da se na osnovu postojanja linearne korelacione veze dvije pojave, X i Y, **ne smije zaključivati da je X uzrok, a Y posljedica,** ili obrnuto.

Između dvije pojave postoji iskrivljena (apsurdna) korelacija kada je koeficijent korelacije različit od nule, a nemamo nikakvog razloga da vjerujemo da su one međusobno povezane.

Primjeri iskrivljene (apsurdne) korelacije:

1. **Posmatrajmo štete od požara u jednom gradu i broj vatrogasnih kola tokom posljednjih 20 godina.** Najvjerovatnije ćemo dobiti visoku pozitivnu korelaciju, iz prostog razloga što štete pokazuju rastuću tendenciju, zbog inflacije. Da li to znači da ako smanjimo broj vatrogasnih kola, da će se vrijednost šteta smanjiti? Ili ako se korelacija tumači kao uzročnost, da li vatrogasna kola uzrokuju požare?

2. **Neka istraživanja su pokazala da postoji visoka direktna korelacija između dužine ruku osnovaca i stepena njihovog logičkog rezonovanja, odnosno da osnovci sa dužim rukama bolje rezonuju.** Ali ovo je besmisleno jer je iz analize izostavljena treća važna varijabla, a to su godine starosti. Dakle, osnovci sa dužim rukama zaista rezonuju bolje, ali zato što su stariji!