



1. Uvodna statistika

1.1 Uloga i važnost statistike u analizi podataka u lancima snabdevanja

Statistika igra ključnu ulogu u modernim lancima snabdevanja, gde su učinkovito upravljanje, planiranje i kontrola ključni. Statističke metode koriste se za prikupljanje, analizu i tumačenje podataka, omogućujući kompanijama da bolje razumeju i optimiziraju svoje lance snabdevanja.

Istaknimo neke od važnih uloga statistike u analizi lanca snabdevanja.

Deskriptivna statistika ključna je za opisivanje osnovnih svojstava podataka o lancu snabdevanja, kao što su srednja vrednost, standardna devijacija, medijana, kvartili i druge mere. Ovi alati nam pomažu razumeti distribuciju i karakteristike podataka kao što su prosečna vremena isporuke, količine na zalihama i prosečni troškovi, što doprinosi boljem razumevanju i upravljanju lancem snabdevanja.

Osim toga, statističke tehnike kao što su regresija, analiza vremenskih serija i analiza uzoraka koriste se za predviđanje budućih događaja i trendova u lancima snabdevanja. To uključuje predviđanje, tj. prognoziranje potražnje, zaliha i vremena isporuke, što omogućuje bolje planiranje i prilagođavanje ponude.

Statistika igra ključnu ulogu u prepoznavanju obrazaca u podacima, omogućujući bolje razumevanje ponašanja lanca snabdevanja, uključujući sezonske obrasce, trendove i cikluse potražnje.

Optimizacija zaliha je još jedno ključno područje u kojem statistika pomaže u određivanju optimalnih količina narudžbina koje minimiziraju troškove skladištenja i naručivanja, koristeći metode kao što je EOQ (engl. *Economic Order Quantity* - ekonomična količina narudžbine).

Osim toga, statistika se takođe koristi za procenu rizika lanca snabdevanja, kao što je verovatnost kašnjenja u isporukama, oštećenja tokom transporta i drugih potencijalnih problema.



Statističkim praćenjem i kontrolom procesa identifikujemo odstupanja od standarda, što nam omogućuje poboljšanje kvaliteta i učinkovitosti procesa lanca snabdevanja.

Osim toga, statistika se koristi za praćenje i poboljšanje kvaliteta proizvoda i usluga u lancu snabdevanja, uključujući kontrolu kvaliteta kod dobavljača.

Konačno, statistika je ključni alat za donošenje utemeljenih odluka o nabavci, zalihama, izboru dobavljača i drugim aspektima upravljanja snabdevanjem, doprinoseći učinkovitom i delotvornom radu celog lanca snabdevanja.

U analizi lanca snabdevanja statistika se koristi za optimizaciju procesa, smanjenje troškova, povećanje učinkovitosti i poboljšanje zadovoljstva kupaca. Omogućuje bolje razumevanje dinamike lanca snabdevanja i bolje upravljanje rizicima, što je ključno za uspešno poslovanje kompanija i organizacija u današnjem globalnom okruženju.

1.2 Osnovni pojmovi statistike

Varijable

Varijable su osnovni gradivni blokovi u statistici jer predstavljaju svojstva ili karakteristike koje se mere ili posmatraju u anketi, eksperimentu ili uzorku podataka. Varijable su ključne za razumevanje i analizu podataka jer omogućuju istraživačima, analitičarima i statističarima da opisuju, analiziraju i razumeju fenomene.



Važno je razumeti različite vrste varijabli i njihovu važnost u statistici.

Kvalitativne (deskriptivne, kategoričke) varijable su varijable koje predstavljaju kvalitativne karakteristike ili kategorije koje se ne mogu prebrojati ili klasifikovati prema matematičkom redu. Primeri uključuju pol (muški, ženski), boju očiju (plave, smeđe, zelene) ili vrstu automobila (limuzina, karavan, SUV). Kvalitativne varijable često su korisne za opisivanje demografskih karakteristika ili osobina.

Kvantitativne (numeričke) varijable su varijable koje predstavljaju numeričke vrednosti koje se mogu prebrojati ili izmeriti i mogu se sortirati po nekom matematičkom redu. Primeri uključuju starost, visinu, temperaturu, prihod ili rezultate istraživanja. Kvantitativne varijable često se koriste za analizu i kvantitativno istraživanje fenomena.



Zavisne i nezavisne varijable. Zavisna varijabla je ona koju želimo istražiti, meriti ili predvideti, dok je nezavisna varijabla ona koja treba uticati na zavisnu varijablu. Na primer, ako želimo istražiti da li nivo obrazovanja utiče na dohodak, dohodak je zavisna varijabla, a nivo obrazovanja nezavisna varijabla.

Diskretne i kontinuirane varijable. Varijable se takođe mogu podeliti na diskretne i kontinuirane. Diskretne varijable imaju ograničen skup mogućih vrednosti i obično su predstavljene celim brojevima. Primer je broj dece u porodici, gde su moguće vrednosti 0, 1, 2 itd. Kontinuirane varijable, s druge strane, imaju beskonačan broj mogućih vrednosti i obično se mere pomoću decimalnih brojeva. Primer je visina osoba, gde je moguć beskonačan broj vrednosti unutar zadatog raspona.

Varijable su osnovni alati za istraživanje i analizu podataka. Razumevanje i pravilno definisanje varijabli ključno je za sprovođenje statističkih analiza i proučavanje fenomena u istraživanju. Varijable omogućuju istraživačima izražavanje i kvantifikaciju različitih aspekata stvarnosti, omogućujući bolje razumevanje fenomena, donošenje odluka i predviđanje budućih događaja. Takođe omogućuju korišćenje različitih statističkih tehnika za testiranje hipoteza, predviđanja i bolje razumevanja uzročno-posledičnih veza između varijabli.

1.3 Osnovni statistički koncepti s primerima

Prosek (srednja vrednost)

Srednja **vrednost** (engl. *mean*), takođe poznata kao **prosek**, jedna je od osnovnih statističkih mera. Srednja vrednost je aritmetički prosek svih vrednosti u skupu podataka. Izračunava se sabiranjem svih podataka, a zatim deljenjem s brojem podataka.



Izračunavanje proseka:

- Sabrati sve vrednosti u skupu podataka.
- Podelite zbir s brojem vrednosti u skupu.
- Jednačina za izračunavanje proseka ($\bar{x}$) je: $\bar{x} = (x_1 + x_2 + x_3 + \dots + x_n) / n$

Gde je $\bar{x}$ prosek. $x_1, x_2, x_3, \dots, x_n$ su vrednosti u skupu podataka. n je broj vrednosti u skupu podataka.



Primer:

Zamislite skup podataka koji predstavlja ocene učenika na ispitu iz matematike: 80, 85, 90, 75, 95. Da biste izračunali prosek, saberite sve ove vrednosti i podelite s brojem ocena, koji je u ovom slučaju 5:

$$\text{Prosek} = (80 + 85 + 90 + 75 + 95) / 5 = 425 / 5 = 85$$

Dakle, prosečni studentski rezultat je 85. Prosek je koristan za merenje centralne tendencije podataka i daje nam grubu ideju o tome šta možemo očekivati kao "tipičnu" vrednost u skupu podataka. Međutim, srednja vrednost može se značajno promeniti ako su u podacima prisutni ekstremi. Stoga je važno poznavati druge statističke mere kao što su medijana i mod kako bi se bolje razumela distribucija podataka.

Medijana

Medijana je statistički koncept koji se koristi za merenje položaja srednje vrednosti skupa podataka. To je vrednost koja deli uređene podatke na dve jednake polovine. To znači da polovina podataka ima vrednosti manje ili jednake medijani, a druga polovina ima vrednosti veće ili jednake medijani. Medijana je jedna od osnovnih mera centralne tendencije u statistici i koristi se za opisivanje distribucije podataka, posebno kada su podaci iskrivljeni ili sadrže ekstremne vrednosti.



Kako izračunati medijanu?

- Prvo morate sortirati skup podataka od najmanje do najveće vrednosti.
- Ako je broj podataka paran (n), tada je medijana prosek dve srednje vrednosti. To znači da je medijana jednaka proseku vrednosti na poziciji $n/2$ i $(n/2 + 1)$ kada su podaci poređani uzlaznim redosledom.
- Ako je broj podataka neparan, tada je vrednost medijane na sredini.

Primer:

Zamislite sledeći skup podataka koji predstavlja broj sati sna koje su ljudi imali u određenom razdoblju: 7, 6, 5, 8, 6, 9, 7

Prvo posložite podatke uzlaznim redoslijedom: 5, 6, 6, 7, 7, 8, 9



Budući da je broj podataka neparan (7), medijana će biti vrednost na srednjoj poziciji, što je četvrta vrednost u poređanom skupu podataka. Dakle, medijana je u ovom slučaju jednaka 7 sati. To znači da polovina ljudi u ovom skupu podataka spava 7 ili manje sati, dok druga polovina spava 7 ili više sati.

Mod

Mod je jedna od osnovnih statističkih metrika koja se koristi za merenje centralne tendencije skupa podataka. Mod predstavlja vrednost koja se najčešće pojavljuje u skupu podataka. To je vrednost koja ima najveću učestalost pojavljivanja među svim vrednostima u skupu podataka.

Modus je koristan za identifikovanje najčešće vrednosti u skupu podataka i posebno je koristan pri analizi kvalitativnih (kategorijskih) varijabli gde su vrednosti nenumeričke.

Ako postoji više modova u skupu podataka (više vrednosti koje se javljaju sa sličnom maksimalnom učestalošću), govorimo o višemodalnoj distribuciji. Ako svi podaci imaju istu učestalost pojavljivanja, tada skup podataka nema mod.

Primer: zamislite skup podataka koji predstavlja boje automobila na parkingu:

Crvena, Plava, Crvena, Zelena, Plava, Plava, Plava, Crvena

U ovom slučaju mod je "Plava", jer se ova vrednost najčešće pojavljuje (četiri puta), dok se "Plava" i "Zelena" pojavljuju ređe.

Mod je jednostavan za izračunavanje jer jednostavno identifikuje vrednost s najvećom učestalošću pojavljivanja u skupu podataka. Mod se koristi za opisivanje karakterističnih vrednosti u podacima i može biti koristan u razumevanju koja je vrednost najkarakterističnija za određenu situaciju ili grupu.

Raspon varijacije

Razlika između maksimalne i minimalne vrednosti u skupu podataka je statistički koncept koji se naziva raspon. Time se meri kolika je razlika između maksimalne i minimalne vrednosti u skupu podataka. Raspon je jednostavan način za procenu raspona vrednosti u skupu podataka i merenje varijabilnosti između minimalnih i maksimalnih vrednosti.



Izračunavanje raspona varijacije je jednostavno:

- Najpre pronađite minimalnu vrednost (min) i maksimalnu vrednost (max) u skupu podataka.
- Zatim izračunajte razliku između maksimalne i minimalne vrednosti (max - min).

Primer: zamislite skup podataka koji predstavlja starost učesnika događaja: 20, 25, 30, 35, 40. Da biste izračunali raspon varijacije, prvo pronađite minimalnu vrednost (20) i maksimalnu vrednost (40) u skupu podataka. Zatim izračunajte razliku između maksimalne i minimalne vrednosti: $VR = 40 - 20 = 20$

Dakle, raspon varijacije u ovom slučaju je 20 godina. To znači da je razlika između najstarijeg i najmlađeg učesnika 20 godina.

Dekompozicija varijacije je korisna za procenu raspona vrednosti u skupu podataka, ali je prilično jednostavna i ne uzima u obzir sve vrednosti u skupu podataka. Za detaljniju analizu varijabilnosti i disperzije podataka obično se koriste druge statističke mere kao što su varijansa ili kvartili.

Varijansa i standardna devijacija

Varijansa je prosečni zbir kvadrata odstupanja od srednje vrednosti. To je kvadrat standardne devijacije. **Standardna devijacija** je statistička mera koja se koristi za merenje disperzije ili varijabilnosti u skupu podataka. Govori koliko su vrednosti udaljene od srednje vrednosti (proseka) u skupu. Standardna devijacija jedna je od najčešće korišćenih mera disperzije u statistici i utvrđuje se izračunavanjem kvadratnog korena varijance.

Izračunavanje standardne devijacije:

- Prvo izračunajte varijansu. Varijansa se izračunava uzimanjem proseka svih vrednosti u skupu za svaku vrednost u skupu, zatim kvadriranjem i sabiranjem tih razlika.
- Kada dobijete vrednost varijanse (σ^2), izračunajte standardnu devijaciju izračunavanjem kvadratnog korena varijanse. To se radi vađenjem kvadratnog korena iz σ^2 :

Standardna devijacija $\sigma = \sqrt{\sigma^2}$





Standardna devijacija meri koliko su vrednosti disperzovane oko srednje vrednosti u skupu podataka. Veća vrednost standardne devijacije znači da su vrednosti raširenije i više se razlikuju od srednje vrednosti, dok niža vrednost standardne devijacije označava manju raspršenost.

Primer: zamislite skup podataka koji predstavlja ocene učenika na ispitu iz matematike: 80, 85, 90, 75, 95. Formula koja će biti prikazana u nastavku važi samo ako pet vrednosti s kojima smo započeli čine celokupnu populaciju. Prvo izračunavate prosek (srednju vrednost), koji iznosi 85. Zatim izračunavate varijansu, koja iznosi 50.

Prvo izračunajte odstupanja svakog podatka od srednje vrednosti i kvadrirajte rezultat svake:

$$(80 - 85)^2 = (-5)^2 = 25, \quad (85 - 85)^2 = (0)^2 = 0, \quad (90 - 85)^2 = (5)^2 = 25, \quad (75 - 85)^2 = (-10)^2 = 100, (95 - 85)^2 = (10)^2 = 100$$

Varijansa je srednja vrednost ovih vrednosti:

$$\sigma^2 = \frac{25 + 0 + 25 + 100 + 100}{5} = \frac{250}{5} = 50$$

Na kraju, izračunavate standardnu devijaciju uzimajući kvadratni korijen varijanse:

$$\text{Standardna devijacija} = \sqrt{50} \approx 7.07$$

Dakle, standardna devijacija je u ovom slučaju oko 7,07. To znači da su u proseku rezultati učenika udaljeni oko 7,07 jedinica od proseka. Standardna devijacija se često koristi u analizi distribucije podataka i u proceni varijabilnosti vrednosti u skupu.

Kvartili

Kvartili su vrednosti koje dele uređene podatke u određene delove. Na primer, kvartili dele podatke na četiri jednaka dela. Prvi kvartil (Q1) deli donjih 25% podataka, drugi kvartil (Q2) jednak je medijani, a treći kvartil (Q3) deli gornjih 25% podataka.



Primer: u skupu podataka 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, prvi kvartil (Q1) jednak je 6, drugi kvartil (Q2) jednak je 11, a treći kvartil (Q3) je jednak 16.

1.4 Prikaz statistike

Prikaz statistike uključuje korišćenje različitih metoda i alata, s ciljem da se podaci prezentuju na jasan, transparentan i informativan način.



Evo nekoliko uobičajenih načina za prikaz statistike:

Tabele

Tabele su osnovna metoda za prikaz podataka. Primeri uključuju tabele frekvencija, koje pokazuju broj pojavljivanja za različite vrednosti, i tabele podataka koje prikazuju više informacija o podacima.

Studenti – ostvarene ocjene	Zbrojne oznake	Frekvencija
41 - 49		3
50 - 58	/	6
59 - 67	/	5
68 - 76	/	6
77 - 85		2
		Total =22

Slika 1.1 Primer tabele.

Grafički prikazi

Grafički prikazi su koristan alat za vizualizaciju podataka. Uključuju različite vrste grafikona kao što su stubičasti grafikoni, linijski grafikoni, pita grafikoni, histogrami, kutijasti dijagram itd.



Slika 1.2 Primeri grafičkih prikaza podataka.

Linijski grafikoni koriste se za vizualizaciju trendova i promena tokom vremena, što ih čini idealnim za praćenje podataka koji se kontinuirano razvijaju. Posebno su učinkoviti za prikazivanje odnosa između varijabli i isticanje uzoraka, kao što su povećanja, smanjenja ili fluktuacije. Linijski grafikoni obično se koriste u područjima kao što su finansije, nauka i poslovanje za analizu vremenskih serija podataka, poređenje trendova u kategorijama ili predviđanje budućeg razvoja na osnovu istorijskih podataka.



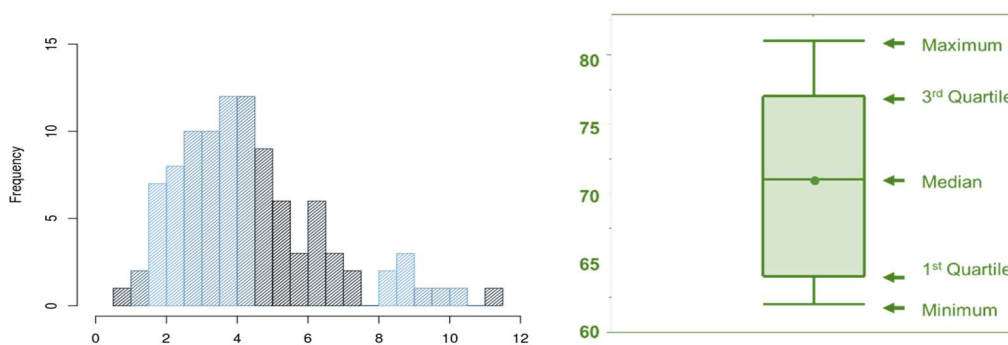
Stubičasti grafikoni koriste se za poređenje količina u različitim kategorijama, što ih čini idealnim za predstavljanje diskretnih podataka. Posebno su delotvorni za isticanje razlika, sličnosti i trendova među grupama. Stubičasti grafikoni obično se koriste kada treba prikazati frekvencije, procenite ili druge numeričke mere na jasan i vizualno jednostavan način. Široko se primenjuju u poslovanju, obrazovanju i istraživanju za analizu i iznošenje kategoričkih podataka.

Polarni grafikoni, takođe poznati kao radarski ili paukovi grafikoni, koriste se za prikaz višestrukih podataka u više dimenzija u kružnom formatu. Idealni su za poređenje nekoliko varijabli ili entiteta prema istim kriterijima, ističući prednosti i slabosti na jasan, vizualan način. Često se koriste u analizi učinka, donošenju odluka i konkurentskim poređenjima, kao što je procena karakteristika proizvoda, timskih veština ili rezultata ankete u različitim kategorijama.

Pita grafikoni koriste se za predstavljanje proporcija ili procenata celine, što ih čini idealnim za vizualizaciju relativnih veličina različitih kategorija. Posebno su učinkoviti kada treba pokazati kako delovi doprinose ukupnom iznosu ili na prvi pogled uporediti proporcije. Tortni grafikoni obično se koriste u izveštajima, prezentacijama i istraživanjima za prikaz podataka poput tržišnog učešća, raspodele proračuna ili demografske distribucije.

Histogrami

Histogrami su grafički prikazi distribucije podataka. Koriste se za prikaz frekvencija vrednosti varijable u različitim intervalima.



Slika 1.3 Histogram i kutijasti dijagram (Box-Plot).

Kutijasti dijagram (Box plot)



Kutijasti dijagram, ili okvir s brkovima, vrsta je grafikona koji se koristi u deskriptivnoj statistici kao prikladan način grafičkog predstavljanja grupa numeričkih podataka njihovim sažimanjem s pet brojeva: minimum, prvi kvartil, medijana, treći kvartil i maksimum.

Izbor metode za prikaz statistike zavisi od prirode podataka, ciljevima analize i ciljanoj publici. Važno je odabrati metodu koja najbolje odgovara vašoj poruci i čini podatke razumijivijim.

1.5 Distribucija frekvencija

Distribucija frekvencija, takođe poznata kao tabela frekvencija ili histogram, način je prikazivanja broja pojavljivanja različitih vrednosti varijable u skupu podataka. Pomoću distribucije frekvencija možete identifikovati obrasce, distribucije i učestalosti vrednosti u podacima. Obično se koristi za analizu kvalitativnih (kategoričkih) varijabli, ali se takođe može koristiti za prikaz diskretnih vrednosti kvantitativnih (numeričkih) varijabli.



Proces stvaranja distribucije frekvencija uključuje sledeće korake:

- Prikupljanje podataka: prvo prikupite podatke za koje želite uraditi distribuciju frekvencija.
- Identifikujte različite vrednosti: identifikujte različite vrednosti koje se pojavljuju u vašim podacima. Ovo su kategorije ili diskretne vrednosti koje želite analizirati.
- Brojanje pojavljivanja: izbrojite koliko puta se svaka vrednost pojavljuje u skupu podataka.
- Napravite tabelu frekvencija: izradite tabelu koja prikazuje sve različite vrednosti varijable i broj pojavljivanja za svaku vrednost.
- Crtanje histograma: ako imate veliki broj različitih vrednosti, možete izraditi histogram koji prikazuje distribuciju frekvencija. Ovo je grafički prikaz koji prikazuje broj pojavljivanja svake vrednosti u obliku stupaca.

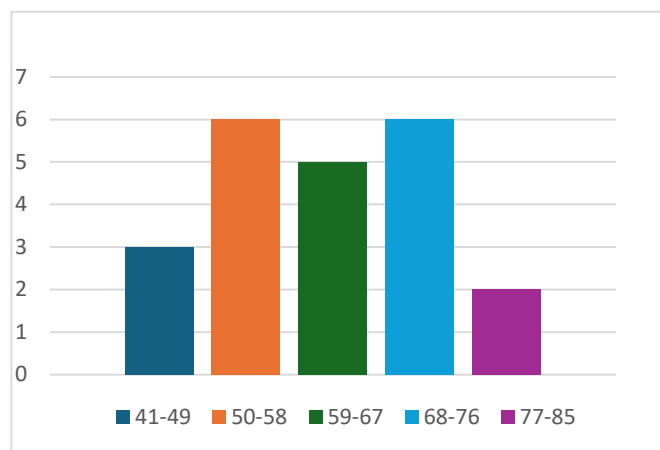
Primer distribucije frekvencija: Zamislite da analiziramo distribuciju učestalosti ocena koje su postigli učenici. Prikupili smo podatke od 22 učenika i želimo videti koliko je učenika osvojilo određeni broj bodova.



Studenti – ostvarene ocjene	Zbrojne oznake	Frekvencija
41 - 49		3
50 - 58		6
59 - 67		5
68 - 76		6
77 - 85		2
		Total =22

Slika 1.22 Tabela distribucije frekvencija.

Grafikon distribucije frekvencija (histogram) bi prikazao stupce za svaki raspon ocena s visinom koja predstavlja broj učenika (frekvenciju) u svakoj klasi. Na taj način možemo jasno videti koja je klasa frekvencija najčešća i kako su ostale oznake u skupu podataka raspoređene. Distribucije frekvencija su koristan alat za vizualizaciju i analizu kvalitativnih podataka i za brzo prepoznavanje obrazaca.



Slika 1.23 Grafikon distribucije frekvencija.

1.6 Deskriptivna i inferencijalna statistika

Deskriptivna statistika: deskriptivna statistika bavi se opisom i sažimanjem podataka iz uzorka ili populacije koja se proučava. Koristi se za analizu i razumevanje podataka, ali ne i za donošenje zaključaka o populaciji kao celini. Glavni cilj deskriptivne statistike je opisati karakteristike podataka, na primer izračunati srednju vrednost, medijanu, raspon, standardnu devijaciju i dati grafičke prikaze kao što su histogrami ili grafikoni. Koristi se za izradu sažetaka i grafikona koji pomažu u vizualizaciji podataka.





Inferencijalna statistika: inferencijalna statistika bavi se donošenjem zaključaka o populaciji iz uzorka. To znači da inferencijalna statistika omogućuje izvođenje zaključaka o populaciji kao celini iz analize uzorka. Koristi različite statističke metode kao što su testiranje hipoteza, intervali pouzdanosti i regresiona analiza kako bi se razumelo da li se promatrani rezultati uzorka mogu generalizirati na populaciju. Na primer, ako želimo otkriti je li srednja starost u uzorku reprezentativna za populaciju u celini, koristićemo se inferencijalnom statistikom.

Inferencijalna statistika

Inferencijalna statistika grana je statistike koja se fokusira na zaključke koje možemo izvući iz podataka koje prikupljamo. Glavni zadatak je izvući opšte zaključke o populaciji ili uzorku iz analize uzorka podataka.

Glavni ciljevi inferencijalne statistike su:

Procena parametara populacije: inferencijalna statistika omogućuje nam procenu parametara populacije kao što su srednja vrednost, varijansa, proporcije i druge karakteristike iz uzorka.

Testiranje hipoteza: inferencijalna statistika može se koristiti za testiranje hipoteza o populaciji na osnovu uzorkovanih podataka. To uključuje statističko testiranje, gde upoređujemo uzorak s pretpostavkama o populaciji.

Stvaranje intervala pouzdanosti: inferencijalna statistika omogućuje izračunavanje intervala koji sadrže procenjene vrednosti parametara populacije s određenom nivom pouzdanosti.

Primer inferencijalne statistike: pretpostavimo da želimo proceniti prosečnu visinu svih studenata na univerzitetu. Budući da je nemoguće proveriti sve učenike, uzimamo uzorak od 100 učenika i merimo njihovu visinu.

Zatim koristimo inferencijalnu statistiku za izračunavanje intervala pouzdanosti za prosečnu visinu svih učenika. Naš uzorak ima srednju visinu od 170 cm i standardnu devijaciju od 5 cm.

Uz pretpostavku da su visine učenika u populaciji **približno normalno distribuirane**, možemo koristiti standardnu grešku srednje vrednosti za izračunavanje intervala



pouzdanosti. Na primer, ako želimo interval pouzdanosti od 95%, koristimo standardnu grešku i kvantile normalne distribucije.

Približan interval pouzdanosti od 95% za prosečnu visinu svih studenata na univerzitetu bio bi:

$$170 \text{ cm} \pm 1.96 \times \left(\frac{5 \text{ cm}}{\sqrt{100}} \right) = 170 \text{ cm} \pm 0.98 \text{ cm}$$

To znači da s 95%-tnom sigurnošću možemo reći da je prosečna visina svih učenika između približno 169,02 cm i 170,98 cm. Ovaj interval pouzdanosti omogućuje nam da zaključimo o prosječnoj visini svih studenata na univerzitetu iz ukupnog uzorka.

Zajedno ove statističke metode omogućuju logističkim kompanijama bolje razumevanje njihovih procesa, predviđanje budućih događaja i donošenje informisanih odluka za poboljšanje učinkovitosti i konkurentnosti.



1.7 Korelacija i regresija

To su statističke metode koje se koriste za proučavanje odnosa između varijabli i predviđanje vrednosti. Obe metode pomažu da se razume kako jedna varijabla utiče na drugu i koliko se dobro jedna varijabla može koristiti za predviđanje druge. Evo objašnjenja svake od ove dve metode:

Korelacija

Korelacija se koristi za merenje stepena povezanosti između dve kvantitativne (numeričke) varijable. Ona ukazuje da li postoji linearna veza između dve varijable i koliko je jaka ta veza. Korelacija se meri koeficijentom korelacije, koji uzima **vrednosti između -1 i 1**.

Koeficijent korelacije 1 znači savršenu pozitivnu korelaciju, što znači da su varijable savršeno korelirane i da se kreću u istom smeru.



Koeficijent korelacije -1 znači savršenu negativnu korelaciju, što znači da su dve varijable potpuno obrnuto korelirane i da se kreću u suprotnim smerovima.

Koeficijent korelacije 0 znači da ne postoji linearna povezanost između varijabli.

Primer: korelacija između broja sati učenja i ocena koje studenti postižu biće pozitivna ako povećanje broja sati učenja obično odgovara višim ocenama.

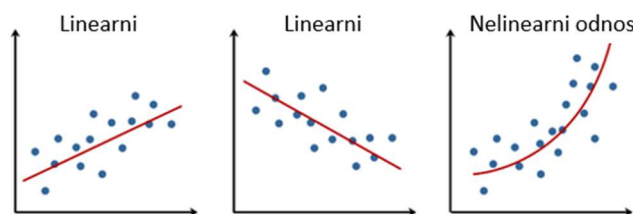
Regresija

Regresija se koristi za modeliranje i predviđanje vrednosti jedne kvantitativne varijable (zavisne varijable) iz vrednosti druge kvantitativne varijable (nezavisne varijable). Postoje različite vrste regresije, uključujući **jednostavnu linearnu regresiju, višestruku linearnu regresiju**, logističku regresiju itd.



Jednostavna linearna regresija: koristi se za modeliranje odnosa između jedne nezavisne varijable i jedne zavisne varijable. Model je linearan i obično se prikazuje jednačinom prave linije ($y = a + bx$), gde je a secište s y -osom, a b nagib krive.

Višestruka linearna regresija: koristi se kada želite modelirati odnos između nekoliko nezavisnih varijabli i jedne zavisne varijable.



Slika 1.24 Grafikon jednostavne linearne regresije.

Primer: jednostavna linearna regresija može se koristiti za modeliranje odnosa između broja obavljenih zadataka učenja (nezavisna varijabla) i završne ocene ispita (zavisna varijabla).

1.8 Distribucija verovatnoća

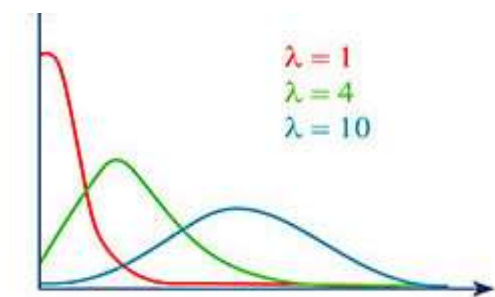
U statistici, distribucija verovatnoća opisuje verovatnost različitih vrednosti koje varijabla može poprimiti. To je matematički model koji pomaže da se





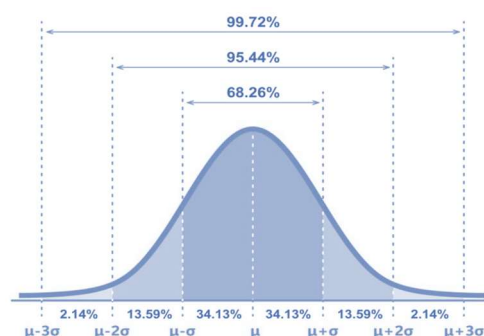
razumeju i analiziraju slučajne pojave i da se predvidi kako će se vrednosti raspodeliti pod određenim okolnostima. Postoji više različitih distribucija verovatnoća, svaka sa svojim karakteristikama i primenama u različitim situacijama. Evo nekih od najpoznatijih distribucija verovatnoća u statistici:

Normalna (Gaussova) distribucija: normalna distribucija jedna je od najvažnijih i najčešće korišćenih distribucija. Opisuje simetričnu i zvonastu distribuciju s poznatim parametrima: srednjom vrednošću (μ) i standardnom devijacijom (σ). Mnogi prirodni fenomeni približni su normalnoj distribuciji.



Slika 1.26 Grafikon Poissonove distribucije.

Binomna



Slika 1.25 Grafikon normalne distribucije.

distribucija: binomna distribucija koristi se za modeliranje broja uspeha (npr. broja "glava") u određenom broju neovisnih Bernoullijevih eksperimenata. Ima dva parametra: broj pokušaja (n) i verovatnost uspeha (p).

Poissonova distribucija: Poissonova distribucija koristi se za modeliranje broja događaja koji se događaju u određenom vremenskom ili prostornom razdoblju. Obično se koristi za modeliranje retkih događaja kao što su nesreće, pozivi hitnim službama itd. Parametar distribucije je prosečna stopa događaja (λ).

Eksponencijalna distribucija: eksponencijalna distribucija poseban je slučaj gama distribucije i koristi se za modeliranje vremena do prvog događaja u Poissonovom procesu. Parametar distribucije je prosečna stopa događaja (λ).

Studentova t-distribucija: Studentova t-distribucija koristi se za procenu intervala pouzdanosti i testiranje hipoteza kada imate mali uzorak i ne znate standardnu devijaciju populacije. To je važno pri analizi uzoraka gde pretpostavka o normalnoj distribuciji može biti slaba.



Hi-kvadrat distribucija: Hi-kvadrat distribucija koristi se za analizu distribucije frekvencija u tabelama, za testiranje nezavisnosti i za testiranje hipoteza. Često se koristi u statističkim testovima kao što je hi-kvadrat test.

F-distribucija: F-distribucija se koristi kada se upoređuje varijabilnost između dva uzorka. Koristi se u analizi varijanse (ANOVA) i drugim statističkim testovima.

Ove distribucije verovatnoća osnovni su gradivni blokovi u statistici i koriste se za modeliranje i analizu različitih vrsta podataka u različitim kontekstima. Izbor ispravne distribucije verovatnoće ključan je pri sprovođenju statističkih analiza i predviđanju rezultata.

Literatura 1. poglavlja

- *Introductory Statistics*. Bentham Science Publishers, Kahl, A. (Publish 2023). DOI:10.2174/97898151231351230101
- *Introductory Statistics 2e*, Openstax, Rice University, Houston, Texas 77005, Jun 23, senior contributing authors: Barbara Illowsky and Susan dean, De anza college, Publish Date: Dec 13, 2023, (<https://openstax.org/details/books/introductory-statistics-2e>);
- *Introductory Statistics 4th Edition*, Susan Dean and Barbara Illowsky, Adapted by Riyanti Boyd & Natalia Casper (Published 2013 by OpenStax College) July 2021, (<http://dept.clcillinois.edu/mth/oer/IntroductoryStatistics.pdf>);
- *Introductory Statistics 7th Edition*, Prem S. Mann, eastern Connecticut state university with the help of Christopher Jay Lacke, Rowan university, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030-5774, 2011
- *Journal of the Royal Statistical Society 2024*, A reputable journal publishing cutting-edge research and articles on various aspects of statistics, including theoretical advancements and practical applications. Recent issues have featured studies on sampling and hypothesis testing.
- *Introduction to statistics, made easy second edition*, Prof. Dr. Hamid Al-Oklah Dr. Said Titi Mr. Tareq Alodat, March 2014



- Statistics for Business and Economics, Thirteenth Edition, David R. Anderson, Dennis J. Sweeney, Thomas A. Williams, Jeffrey D. Camm, James J. Cochran, 2017, 2015 Cengage Learning®
- Statistics for Business, First edition, Derek L Waller, 2008 Copyright © 2008, Derek L Waller, Published by Elsevier Inc. All rights reserved

Dodatne poveznice na literaturu i Youtube videozapise 1. poglavlja

- <https://open.umn.edu/opentextbooks/textbooks/196>
- <https://www.scribbr.com/category/statistics/>
- https://stats.libretexts.org/Bookshelves/Introductory_Statistics
- https://assets.openstax.org/oscms-prodcms/media/documents/IntroductoryStatistics-OP_i6tAI7e.pdf
- https://saylordotorg.github.io/text_introductory-statistics/
- [https://drive.uqu.edu.sa/_/mskhayat/files/MySubjects/20178FS%20Elementary%20Statistics/Introductory%20Statistics%20\(7th%20Ed\).pdf](https://drive.uqu.edu.sa/_/mskhayat/files/MySubjects/20178FS%20Elementary%20Statistics/Introductory%20Statistics%20(7th%20Ed).pdf)
- <https://dept.clcillinois.edu/mth/oer/IntroductoryStatistics.pdf>
- <https://www.geeksforgeeks.org/introduction-of-statistics-and-its-types/>
- https://onlinestatbook.com/Online_Statistics_Education.pdf
- https://www.researchgate.net/profile/Tareq-Alodat-2/publication/340511098_INTRODUCTION_TO_STATISTICS_MADE_EASY/links/5e8de3dc4585150839c7b58a/INTRODUCTION-TO-STATISTICS-MADE-EASY.pdf
- <https://byjus.com/maths/statistics/>
- <https://www.khanacademy.org/math/statistics-probability>
- <https://www.youtube.com/watch?v=XZo4xyJXCak>
- <https://www.youtube.com/watch?v=LMSyiAJm99g>
- https://www.youtube.com/watch?v=VPZD_ajj8H0
- <https://www.youtube.com/watch?v=TLwp5DwcqD4>
- <https://www.youtube.com/watch?v=fpFj1Re1l84>



BAS4SC - Business Analytics Skills for the Future-proof Supply Chains

- https://youtube.com/playlist?list=PLqzoL9-eJTNAB5st3mtP_bmXafGSH1Dtz&si=z-IXQ1iKbw2-ieJW
- <https://www.youtube.com/watch?v=44MJyNTxaP8>