

Šta je statistika?

Statistika je nauka koja se bavi prikupljanjem, analizom, tumačenjem i prikazivanjem podataka. Koristi se u mnogim oblastima kako bi se donijele informisane odluke na osnovu podataka.

Ciljevi učenja:

1. Opisati opseg primjene statistike.
2. Identifikovati situacije u kojima statistika može biti pogrešno tumačena.
3. Definisati šta je statistika.

Primjeri statistike u svakodnevnom životu:

- **Najveći zabilježeni zemljotres** imao je jačinu od 9.2 po Rihterovoj skali.
- **Muškarci** su najmanje 10 puta skloniji da počine ubistvo u poređenju sa ženama.
- **Svaka osma osoba** u Južnoafričkoj Republici je HIV pozitivna.
- **Do 2020. godine**, biće 15 osoba starijih od 65 godina na svakog novorođenog.

Pogrešna tumačenja statistike:

1. **Reklama za sladoled:** Nova reklama za Ben i Jerry's sladoled rezultirala je povećanjem prodaje od 30% u naredna tri mjeseca. Zaključak je da je reklama bila efikasna. Greška: Potrošnja sladoleda generalno raste u ljetnim mjesecima bez obzira na reklame.
2. **Crkve i kriminal:** Više crkava u gradu, više kriminala. Zaključak je da crkve dovode do kriminala. Greška: Povećanje broja crkava i stope kriminala mogu se objasniti većom populacijom u većim gradovima.
3. **Međurasni brakovi:** Ove godine se događa 75% više međurasnih brakova nego prije 25 godina. Zaključak je da društvo prihvata međurasne brakove. Greška: Nemamo dovoljno informacija o stopi sklopljenih brakova.

Značaj statistike:

Statistika nije samo zbir činjenica i brojki; ona uključuje različite tehnike i postupke za analiziranje, tumačenje, prikazivanje i donošenje odluka na osnovu podataka. Statistika se svakodnevno koristi u raznim oblastima i daje kredibilitet argumentima. Da biste pravilno procijenili podatke i tvrdnje koje vas svakodnevno bombarduju, potrebno je da budete inteligentan korisnik statistike.

Terminologija:

- **Statistika:** Razvoj i primjena metoda za prikupljanje, analizu i tumačenje opaženih informacija (podataka) iz planiranih istraživanja.
- **Pojedinci:** Objekti mjerena u statističkom problemu.

- **Varijabla:** Karakteristika koju želimo mjeriti na pojedincima.
- **Podaci:** Stvarne izmjerene vrijednosti na pojedincima.

Opseg primjene statistike:

Statističko razmišljanje se može koristiti u svim disciplinama. Na primjer:

- **Studija pouzdanosti:** Tribolozi žele utvrditi glavni uzrok kvara u sklopu mlaznog motora.
- **Marketinški projekat:** Upravitelji trgovina žele znati koja marka kafe je najomiljenija među populacijom od 18-24 godine.
- **Kliničko ispitivanje:** Ljekari žele utvrditi koji od dva lijeka je efikasniji u liječenju HIV-a u ranim fazama bolesti.
- **Proizvodnja automobilskih boja:** Inženjeri žele razumjeti kako temperatura i brzina miješanja utiču na promjenu boje.

Opservacione studije:

- **Opservaciona studija:** Istraživanje koje se provodi za opažanje pojedinaca, bez pokušaja utjecaja na njihove odgovore. Svrha opservacione studije je opisati grupu ili situaciju.

Populacije i uzorci:

- **Populacija:** Cijela grupa pojedinaca o kojoj želimo donijeti neke zaključke.
- **Uzorak:** Dio populacije koji zapravo posmatramo.

Deskriptivna statistika:

- **Definicija:** Brojke koje se koriste za sumiranje i opisivanje podataka.
- **Primjeri:** Prosječne plate za različita zanimanja, broj neudatih muškaraca na 100 neudatih žena u američkim metropolitanskim područjima.

Deskriptivna statistika je centralna za svijet sporta. Svaki sportski događaj proizvodi brojne statistike, poput postotka uspješnosti šutiranja košarkaških igrača. Za Olimpijski maraton (utrka na 42.2 kilometra), posjedujemo podatke koji pokrivaju više od vijeka natjecanja.

Zaključak:

Statistika je moćan alat koji nam pomaže da razumijemo svijet oko nas kroz analizu podataka. Kroz pravilno korištenje statistike, možemo donositi informisane odluke i izbjegavati pogrešna tumačenja podataka.

Dodatak 1: Kako se statistika koristi u zdravstvu? COVID-19 primjeri

Statistika igra ključnu ulogu u zdravstvu, posebno tokom pandemije COVID-19. Evo nekoliko interesantnih primjera kako se statistika koristi u ovoj oblasti:

1. Praćenje širenja virusa

- **Broj slučajeva:** Statistika se koristi za praćenje broja zaraženih, oporavljenih i preminulih osoba. Ovi podaci pomažu zdravstvenim vlastima da prate širenje virusa i donose odluke o javnom zdravlju.
- **Reprodukcijski broj (R0):** Ovaj broj pokazuje koliko ljudi u prosjeku zarazi jedna zaražena osoba. Ako je R0 veći od 1, virus se širi eksponencijalno.

2. Efikasnost mjera

- **Lockdown mjere:** Statistički podaci se koriste za procjenu efikasnosti mjera poput karantina, socijalnog distanciranja i nošenja maski. Na primjer, smanjenje broja novih slučajeva nakon uvođenja mjera može ukazivati na njihovu efikasnost.
- **Vakcinacija:** Statistika prati stope vakcinacije i analizira koliko su vakcine efikasne u sprečavanju infekcija i teških oblika bolesti.

3. Klinička ispitivanja

- **Testiranje lijekova i vakcina:** Statistika je ključna u kliničkim ispitivanjima za procjenu sigurnosti i efikasnosti novih lijekova i vakcina. Na primjer, randomizirana kontrolisana ispitivanja koriste statističke metode za analizu podataka i donošenje zaključaka.
- **Nuspojave:** Praćenje i analiza nuspojava vakcina i lijekova pomaže u identifikaciji potencijalnih rizika i poboljšanju sigurnosti.

4. Modeliranje pandemije

- **Predikcijski modeli:** Statistički modeli se koriste za predviđanje budućeg toka pandemije. Ovi modeli uzimaju u obzir različite faktore kao što su stope zaraze, stope vakcinacije i efikasnost mjera.
- **Scenariji:** Različiti scenariji se modeliraju kako bi se procijenili potencijalni ishodi i pomoglo u planiranju odgovora na pandemiju.

5. Javno zdravstvo

- **Praćenje simptoma:** Statistika se koristi za analizu podataka o simptomima prijavljenim od strane pacijenata kako bi se identifikovali najčešći simptomi i njihova ozbiljnost.
- **Zdravstveni resursi:** Analiza podataka o kapacitetima bolnica, broju kreveta, respiratora i medicinskog osoblja pomaže u planiranju i alokaciji resursa tokom pandemije.

Primjeri iz COVID-19 pandemije:

- **Johns Hopkins COVID-19 Dashboard:** Ova platforma koristi statističke podatke za praćenje globalnog širenja virusa, prikazujući broj slučajeva, smrtnih slučajeva i oporavaka u realnom vremenu (Johns Hopkins University, n.d.).
- **Studije efikasnosti vakcina:** Statističke analize su pokazale da su vakcine protiv COVID-19 značajno smanjile rizik od teških oblika bolesti i hospitalizacije (Centers for Disease Control and Prevention, 2021).
- **Analiza smrtnosti:** Statistika je korištena za analizu stope smrtnosti među različitim starosnim grupama i identifikaciju faktora rizika koji povećavaju vjerovatnoću smrtnog ishoda (World Health Organization, 2021).

Ovi primjeri pokazuju koliko je statistika važna za razumijevanje i borbu protiv pandemije. Ako imaš još pitanja ili trebaš dodatne informacije, slobodno mi se obrati!

Reference:

1. Johns Hopkins University. (n.d.). *COVID-19 Dashboard*. Retrieved February 23, 2025, from <https://coronavirus.jhu.edu/map.html>
2. Centers for Disease Control and Prevention. (2021). *COVID-19 vaccine effectiveness*. Retrieved February 23, 2025, from <https://www.cdc.gov/coronavirus/2019-ncov/vaccines/effectiveness.html>
3. World Health Organization. (2021). *COVID-19 mortality analysis*. Retrieved February 23, 2025, from <https://www.who.int/data/sets/covid-19-mortality-data>

Dodatak: Zablude u statistici

1. Korelacija podrazumijeva uzročnost

Zabluda: Ako su dvije varijable povezane, jedna mora uzrokovati drugu.

Stvarnost: Korelacija ne podrazumijeva uzročnost. Na primjer, prodaja sladoleda i broj utapanja su povezani jer oba rastu tokom ljeta, ali jedno ne uzrokuje drugo.

Primjer: U jednom istraživanju je pronađena korelacija između broja požara i broja vatrogasaca na terenu. Međutim, to ne znači da vatrogasci uzrokuju požare. Umjesto toga, veći broj požara zahtijeva više vatrogasaca.

Anegdota: Tokom 1970-ih, istraživači su otkrili da je broj roda u određenim regijama bio povezan sa brojem novorođenih beba. Naravno, rode ne donose bebe, već su oba fenomena povezana sa ruralnim područjima gdje su rode češće i gdje je natalitet viši.

2. P-vrijednost

Zabluda: P-vrijednost pokazuje vjerovatnoću da je nulta hipoteza tačna.

Stvarnost: P-vrijednost pokazuje vjerovatnoću posmatranja vaših podataka, ili nečeg ekstremnijeg, pod pretpostavkom da je nulta hipoteza tačna. Ne mjeri vjerovatnoću da je nulta hipoteza tačna.

Primjer: Ako dobijete p-vrijednost od 0.03, to znači da postoji 3% šanse da biste dobili ovakve rezultate (ili ekstremnije) ako je nulta hipoteza tačna. To ne znači da postoji 97% šanse da je alternativna hipoteza tačna.

Anegdota: U medicinskim istraživanjima, p-vrijednosti se često koriste za procjenu efikasnosti tretmana. Međutim, niska p-vrijednost ne garantuje da je tretman efikasan, već samo da su rezultati statistički značajni.

3. Veći uzorak uvijek znači bolje rezultate

Zabluda: Veći uzorci uvijek vode ka tačnijim rezultatima.

Stvarnost: Iako veći uzorci mogu povećati statističku moć, mogu također uvesti pristrasnost i šum ako podaci nisu reprezentativni ili pravilno prikupljeni.

Primjer: Ako istražujete mišljenja o političkom kandidatu i uzmete veliki uzorak samo iz jednog grada, rezultati neće biti reprezentativni za cijelu zemlju.

Anegdota: Tokom predsjedničkih izbora u SAD 1936. godine, časopis "Literary Digest" je sproveo anketu sa velikim uzorkom, ali su koristili telefonske imenike i registarske liste automobila, što je dovelo do pristrasnosti prema bogatijim biračima. Rezultati su bili potpuno pogrešni.

4. Neznačajni rezultati znače da nema efekta

Zabluda: Ako rezultat nije statistički značajan, nema efekta.

Stvarnost: Neznačajni rezultati mogu se pojaviti zbog nedovoljne veličine uzorka ili dizajna studije. To ne znači nužno da nema efekta.

Primjer: Mala studija o efektima novog lijeka može pokazati neznačajne rezultate zbog malog uzorka, iako lijek može biti efikasan.

Anegdota: U istraživanju o povezanosti između pušenja i raka pluća, rane studije sa malim uzorcima nisu pokazale značajnu povezanost. Kasnije, veće studije su jasno pokazale snažnu povezanost.

5. Ignorisanje pretpostavki statističkih testova

Zabluda: Statistički testovi se mogu primijeniti bez razmatranja njihovih osnovnih pretpostavki.

Stvarnost: Svaki statistički test ima specifične pretpostavke (npr. normalnost, nezavisnost) koje moraju biti ispunjene da bi rezultati bili validni.

Primjer: t-test pretpostavlja normalnu distribuciju podataka. Ako podaci nisu normalno distribuirani, rezultati testa mogu biti netačni.

Anegdota: U jednoj studiji, istraživači su koristili t-test na podacima koji nisu bili normalno distribuirani, što je dovelo do pogrešnih zaključaka. Kada su koristili odgovarajući neparametrijski test, rezultati su se značajno promijenili.

6. Jedan test je dovoljan da dokaže nalaz

Zabluda: Jedan statistički test može definitivno dokazati hipotezu.

Stvarnost: Replikacija i višestruki testovi su neophodni da bi se potvrdili nalazi i osigurala robusnost.

Primjer: Ako jedan eksperiment pokaže da određena dijeta smanjuje težinu, potrebno je više studija da bi se potvrdili rezultati i osigurala njihova validnost.

Anegdota: U psihologiji, mnogi eksperimenti su se pokazali teško replikabilnim, što je dovelo do "krize replikacije" u ovoj oblasti. To je naglasilo važnost ponovljenih testova i replikacija.

7. Preterano generalizovanje rezultata

Zabluda: Rezultati iz uzorka mogu se direktno primijeniti na cijelu populaciju.

Stvarnost: Preterano generalizovanje može dovesti do netačnih zaključaka. Važno je razmotriti reprezentativnost uzorka i kontekst studije.

Primjer: Ako istražujete prehrambene navike studenata na jednom univerzitetu, ne možete generalizovati rezultate na sve studente u zemlji.

Anegdota: U jednoj studiji o efektima vježbanja na zdravlje, uzorak je bio sastavljen samo od mladih, zdravih muškaraca. Generalizovanje rezultata na cijelu populaciju, uključujući starije osobe i žene, bilo bi netačno.

8. Nerazumijevanje deskriptivne i inferencijalne statistike

Zabluda: Deskriptivna statistika se može koristiti za izvođenje zaključaka o populaciji.

Stvarnost: Deskriptivna statistika sumira podatke, dok se inferencijalna statistika koristi za predviđanja ili izvođenje zaključaka o populaciji na osnovu uzorka.

Primjer: Prosječna ocjena studenata na jednom kursu je deskriptivna statistika. Korišćenje te prosječne ocjene za predviđanje uspjeha svih studenata na univerzitetu zahtijeva inferencijalnu statistiku.

Anegdota: U jednoj anketi, prosječna plata zaposlenih u jednoj kompaniji je bila visoka zbog nekoliko veoma visokih plata menadžera. Korišćenje te prosječne plate za zaključivanje o platama svih zaposlenih bilo bi varljivo.