

MJESTO I ULOGA STATISTIČKIH METODA U NAUČNIM ISTRAŽIVANJIMA

PLACE AND ROLE OF STATISTICAL METHODS IN SCIENTIFIC RESEARCH

Dr Stevan R. Stević, redovni profesor

Ekonomski fakultet Brčko Univerziteta u Istočnom Sarajevu

stevan.stevic.efb@gmail.com

APSTRAKT

Svako naučno istraživanje, kako teorijsko tako i empirijsko, u osnovi počiva na primjeni najvažnijih metoda naučnoistraživačkog rada. Zbog toga su u radu obrađeni pojam i etape naučnog istraživanja, sa naglaskom na metodologiji istraživanja, odnosno na mjestu, ulozi i značaju statistike i statističkih metoda.

U empirijskim istraživanjima nezaobilaznu ulogu imaju metodi deskriptivne statistike koji se tiču definisanja osnovnog skupa, vrste i veličine uzorka, te prikupljanja, grupisanja i sređivanja, prikazivanja i obrade podataka o posmatranoj pojavi. Zavisno od odabranog predmeta istraživanja, postavljenih ciljeva i istraživačkih hipoteza, koriste se odgovarajući statistički metodi pomoću kojih će se provjeravati hipoteze i odgovarati na najvažnija pitanja kojima se ostvaruje naučni doprinos istraživanja.

U radu su, u najkraćim crtama, obrađeni najvažniji statistički metodi koji se koriste u empirijskim istraživanjima.

Ključne riječi: naučno istraživanje, etape istraživanja, statistika, statistički metodi

ABSTRACT

Any scientific research, either theoretical or empirical, is essentially based on the application of fundamental research methods. Therefore, this paper analyzes the concept and the various stages of scientific research, with an emphasis on research methodology, i.e., the place, role and importance of statistics and statistical methods.

Descriptive statistics methods relating to the definition of the basic compilation, type, and size of the sample, as well as collection, grouping, sorting, displaying, and processing of data on the observed phenomenon, play an indispensable role in empirical research. Appropriate statistical methods should be chosen depending on the research subject, preset objectives and research hypothesis, in order to reliably verify the hypotheses and answer the most important questions that lead to a meaningful scientific contribution.

Succinctly, this paper analyzes the most important statistical methods used in empirical research.

Key words: scientific research, stages of research, statistics, statistical methods.

UVOD

Naučno istraživanje predstavlja sistematski, kritički i kontrolisani ponovljivi proces sticanja novih znanja, koja su neophodna za identifikovanje, određivanje i rješavanje određenih teorijskih i empirijskih naučnih problema.

Istraživanje treba da ispuni određene zahtjeve. Najčešće se ističu sljedeći zahtjevi: sistematičnost, kontrolisanost i kritičnost.

Sistematičnost istraživanja podrazumijeva njegovo dobro planiranje, organizovanje i uređenje, kako bi se obezbijedio prirodan i logičan slijed procesa istraživanja. Kontrolisanost istraživanja označava njegovo provođenje u, što je moguće bolje, praćenim i kontrolisanim uslovima kako bi se isključile alternativne mogućnosti. Kritičnost istraživanja podrazumijeva zahtjev da se najvažnije hipoteze o pojavama, procesima, odnosima i svojstvima u oblasti koja se proučava moraju ozbiljno, iskreno i strogo procjenjivati, provjeravati i obrazlagati odgovarajućim dokazima.

Posebno mjesto u svakom naučnom istraživanju zauzima metodologija rada, u okviru koje se koriste odgovarajući naučni metodi, kako opšti tako i metodi primjereni specifičnostima svakog konkretnog istraživanja.

Cilj rada je da se ukaže na mjesto, ulogu i značaj statističkih metoda u naučnim istraživanjima, prije svega u empirijskim istraživanjima. U tom smislu, pored objašnjenja uloge i značaja najvažnijih metoda deskriptivne statistike, posebna pažnja posvećena je najčešće korišćenim metodima statističke analize, bez kojih nije moguća provjera postavljenih hipoteza i davanje odgovora na najvažnija istraživačka pitanja.

POJAM I ETAPE NAUČNOG ISTRAŽIVANJA

Naučno opravdano istraživanje je ono koje doprinosi produblivanju, proširivanju, pouzdanosti i primjenjivosti naučnog saznanja o predmetima i metodima nauke i istraživanja. Društvena opravdanost naučnog istraživanja ocjenjuje se na osnovu aktuelnog i potencijalnog doprinosa provođenjem i rezultatima istraživanja u rješavanju određenog društvenog problema.

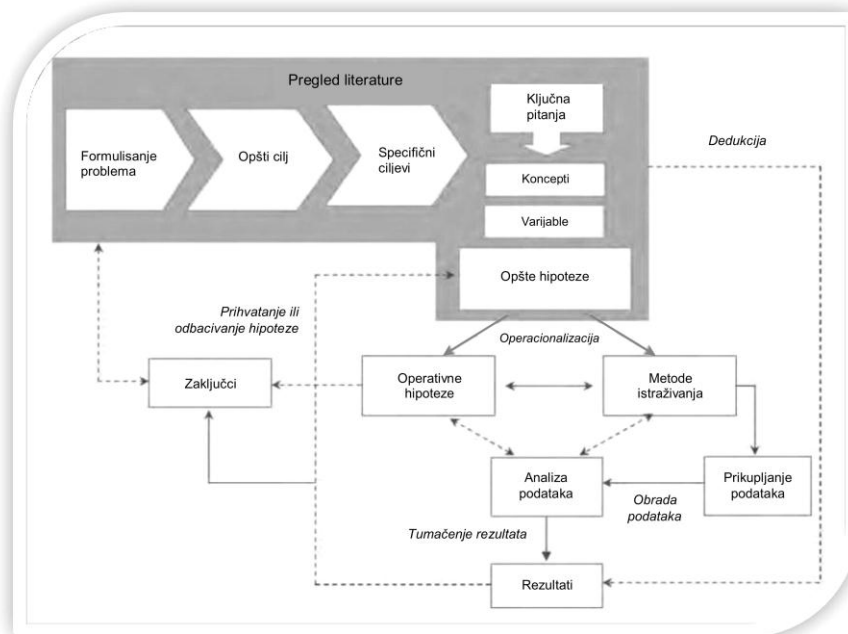
Između naučnog i društvenog doprinosa, odnosno naučnog i društvenog značaja i naučnih i društvenih ciljeva postoji veza, ali se oni međusobno ne mogu izjednačiti. Naučni i društveni doprinos ne ostvaruje se samo rezultatima istraživanja ili razrješavanjem određenog društvenog problema, već i samim provođenjem istraživanja. Zbog toga se može govoriti o posrednom i neposrednom doprinosu, kao i o kratkoročnom i dugoročnom doprinosu istraživanja.

Svako naučno istraživanje započinje izučavanjem relevantne literature, nakon čega slijedi izbor područja istraživanja i definisanje hipoteza. U ovoj etapi rada postavlja se istraživački problem, definišu se opšti i posebni ciljevi i ključna pitanja na koja se istraživanjem želi dati odgovor.

U narednoj fazi, zavisno od načina na koji su postavljene hipoteze, slijedi projektovanje istraživanja i izbor odgovarajuće metodologije koja će omogućiti da se obezbijede relevantni rezultati i odgovor na istraživačka pitanja. Metodologija istraživanja obuhvata izbor načina prikupljanja podataka, planiranje i realizaciju uzorka, kao i izbor statističkih metoda za obradu i analizu podataka.

Interpretacija rezultata podrazumijeva sagledavanje ostvarenja ciljeva istraživanja, u smislu davanja odgovora na postavljena ključna pitanja u radu, kao i potvrdu ili odbacivanje istraživačkih hipoteza. Nakon toga daje se rezime istraživanja sa opštim zaključcima o ostvarenim rezultatima i značaju istraživanja, kao i budućim pravcima istraživanja u određenoj oblasti.

Proces naučnog istraživanja može se, šematski, prikazati na sljedeći način:



Slika 1. Etape naučnog istraživanja
Prilagođeno prema: Breda, Z. (2010), strana 144.

Obaveza istraživača je da naučnu javnost, kao i sve potencijalne korisnike rezultata istraživanja, obavijesti o sljedećim pojedinostima vezanim za istraživanje: o problemu istraživanja, postavljenim ciljevima istraživanja, hipotezama, načinu istraživanja, rezultatima istraživanja i zaključcima izvedenim iz njih.

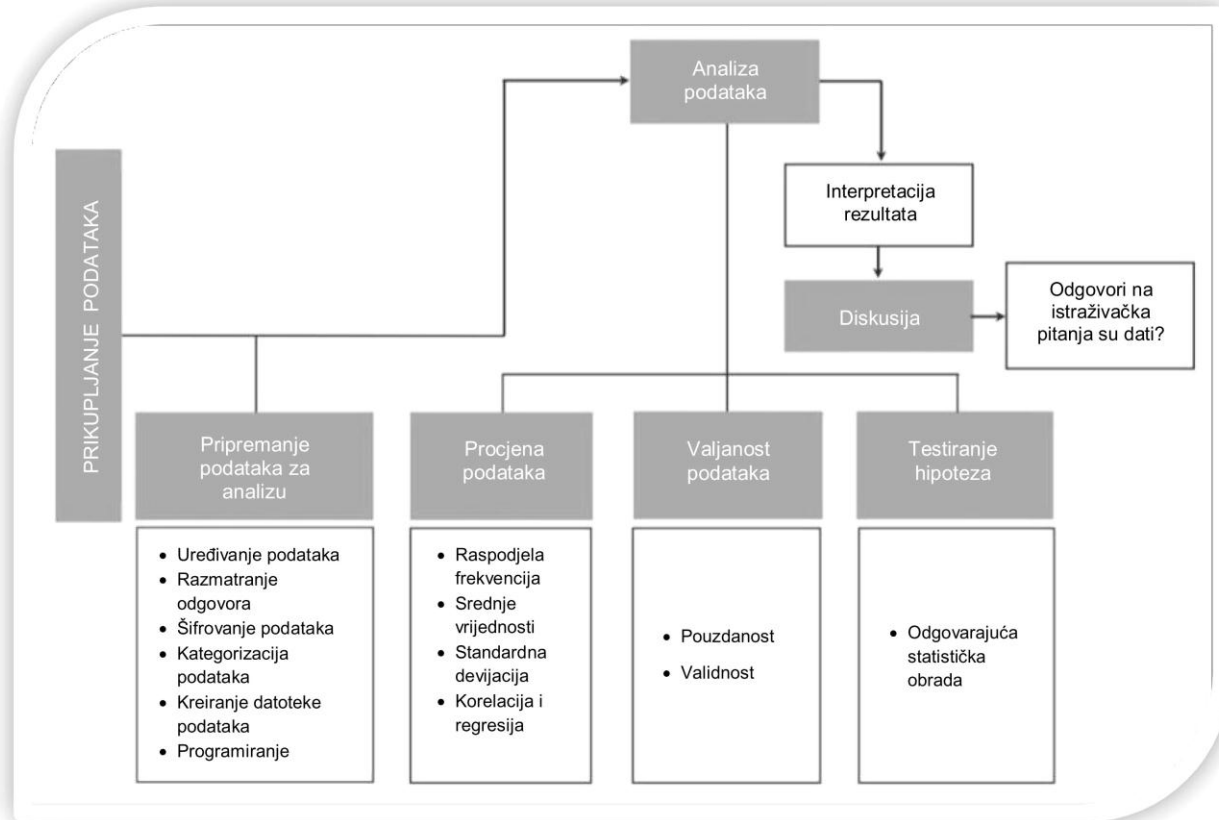
U svakom istraživanju, bilo empirijskom ili teorijskom, primjenjuje se bar jedan opštenaučni metod. Za teorijska istraživanja primjenjuje se analitičko-deduktivni i aksiomatski opštenaučni metod. Najviše se koriste metodi ispitivanja i analiza postojećih dokumenata. Za empirijske nauke karakteristična je primjena različitih statističkih metoda, hipotetičko-deduktivnog metoda i metoda modeliranja.

MJESTO STATISTIKE U NAUČNIM ISTRAŽIVANJIMA

Statistika se najčešće shvata kao kvalitativno-kvantitativni naučni metod analize varijabilnih pojava (pojava na koje djeluje veći broj faktora, zbog čega one pokazuju različite vrijednosti od jednog do drugog slučaja svog ispoljavanja i čije vrijednosti nije moguće sa sigurnošću predvidjeti), zasnovan na teoriji vjerovatnoće. Međutim, nešto sadržajnije definicija statistike ima trostruko značenje: Deskriptivna statistika (kao skup metoda kojima se podaci o odgovarajućoj pojavi prikupljaju, sređuju, grupišu, obrađuju i na adekvatan način prikazuju), statistička analiza (koja podrazumijeva primjenu različitih metoda analize podataka radi ustanovljavanja odgovarajućih pravilnosti i zakonitosti) i statistička teorija (čiji je zadatak da unapređuje postojeće i razvija nove metode statističke analize).

Ni jedno ozbiljnije empirijsko istraživanje nije moguće realizovati bez primjene nekog od najvažnijih statističkih metoda. U prvom redu koriste se metodi deskriptivne statistike (prikupljanje, sređivanje i grupisanje, prikazivanje i obrada podataka). Zavisno od predmeta i ciljeva istraživanja najčešće se koriste sljedeći metodi statističke analize: hi-kvadar test, analiza varijanse, faktorska analiza, regresiona i korelaciona analiza, kao i metodi dinamičke analize i analize vremenskih serija.

Postupak prikupljanja, vrednovanja, upotrebe i analize podataka može se prikazati na sljedeći način:



Slika 2. Postupak obrade i analize podataka

Prilagođeno prema: Breda, Z. (2010), strana 178.

Prikupljanje podataka

Planom naučnog istraživanja određuje se varijabilna pojava čijom analizom se dolazi do zaključaka koji su relevantni za provjeru postavljenih hipoteza i odgovor na istraživačka pitanja. Cilj je da se obezbijede kvalitetni podaci o posmatranoj varijabilnoj pojavi. Posmatranje i prikupljanje podataka vrši se na osnovu prethodno postavljenog plana prikupljanja podataka.

Plan posmatranja i prikupljanja podataka sadrži sljedeće elemente: definisanje cilja posmatranja, određivanje statističkog skupa i jedinica posmatranja, izbor obilježja i definisanje modaliteta obilježja, sastavljanje upitnika za prikupljanje podataka, određivanje načina posmatranja i prikupljanja podataka.

Prema izvoru podataka koji se koriste u statističkom istraživanju, može se govoriti o primarnim i sekundarnim statističkim podacima. Primarni statistički podaci prikupljaju se u toku samog istraživanja, odgovarajućim postupcima statističkog posmatranja i prikupljanja podataka. Sekundarni podaci obezbjeđuju se iz sekundarnih izvora (publikacije zavoda za statistiku, ili ovlaštenih institucija kao što su: centralna banka, carinska služba, matične službe opština, preduzeća i sl.).

Statističko istraživanje može se zasnivati na potpunom obuhvatu svih jedinica skupa (potpuno posmatranje na bazi popisa ili statističkog izvještaja), ili samo na jednom dijelu njegovih jedinica (djelimično posmatranje na bazi uzorka).

Prikupljeni podaci o varijabilnoj pojavi sređuju se prema odabranim obilježjima (odnosno varijablama) i grupišu u odgovarajuće serije (najčešće serije strukture, odnosno rasporede frekvencija). Statističke serije su rezultat grupisanja podataka, a istovremeno i jedan od vidova prikazivanja podataka prije njihove statističke obrade. Pored statističkih serija, podaci se prikazuju u obliku odgovarajućih statističkih tabela, koje mogu da budu: proste, složene i kombinovane, zavisno od sadržaja i vrste podataka, odnosno: iskazne i obradne, zavisno od njihove namjene.

Statistički podaci o varijabilnoj pojavi mogu da se, na slikovit način, prikažu i grafički, najčešće pomoću dijagrama kao odgovarajućih geometrijskih oblika. Zavisno od prirode podataka i njihove upotrebe, najčešće se koriste: tačkasti, linijski, površinski i prostorni dijagrami.

Pojam i planiranje uzorka

U velikom broju statističkih istraživanja često se koristi djelimično (nepotpuno) posmatranje, zasnovano na statističkom uzorku. Statističko uzorkovanje predstavlja metod po kome se na osnovu posmatranja jednog dijela jedinica skupa zaključuje o karakteristikama i ponašanju cijelog skupa.

Da bi zaključci na osnovu uzorka bili relevantni za cijeli skup, potrebno je da uzorak bude reprezentativan. Za uzorak kažemo da je reprezentativan ako vjerno odslikava strukturu osnovnog skupa iz kojeg je odabran. Zbog toga se prilikom korišćenja uzorka u istraživanju polazi od definisanja osnovnog skupa i utvrđivanja okvira uzorka (liste, fajla, direktorijuma), odnosno spiska svih jedinica skupa koji se želi posmatrati. Nakon toga se određuje vrsta uzorka i njegova veličina, kao i sam postupak provođenja uzorkovanja.

Smisao primjene uzorka je u nastojanju da se dobiju dovoljno precizne ocjene najvažnijih karakteristika svih jedinica skupa, a da se pri tome koristi samo jedan njegov dio. Što je osnovni skup homogeniji, moguće je i na manjem broju jedinica ocijeniti njegove najvažnije karakteristike.

U slučaju nepoznavanja homogenosti jedinica skupa, veličina uzorka utvrđuje se u zavisnosti od varijabiliteta jedinica skupa, stepena pouzdanosti ocjene odgovarajućih parametar skupa i prihvatljivog odstupanja koje se toleriše prilikom istraživanja. Ukoliko se povećava stepen pouzdanosti ocjene, utoliko se povećava i veličina uzorka i obratno. S druge strane, veći uzorak će obezbijediti i precizniju ocjenu karakteristika skupa. Kada se govori o dozvoljenoj toleranciji u smislu granica odstupanja ocjene, važi sljedeće pravilo: ukoliko je tolerancija manja potrebno je raditi sa većim uzorcima, dok je za slučaj u kome je prihvatljivo veće odstupanje moguće raditi i sa manjim uzorcima.

Kada je u pitanju veličina uzorka, prihvaćeno je stanovište da se u male uzorke ubrajaju oni koji sadrže do 30 jedinica skupa odnosno najviše 5% ukupne populacije, a da su u suprotnom veliki uzorci.

Primjena metoda uzorka neminovno dovodi do mogućnosti greške u statističkom zaključivanju. Razlog je u tome da uzorak skoro nikada ne može biti savršeno reprezentativan (sem u sljedeća dva slučaja: 1) kada je po veličini jednak skupu, i 2) ako bi skup bio apsolutno homogen, odnosno sastavljen od identičnih jedinica posmatranja). Greške mogu da budu slučajne i sistematske. Cilj je smanjenje sistematske greške koja može da utiče na rezultat, dok se slučajna greška smanjuje ili potpuno gubi u velikom broju podataka.

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, počev od toga da se radi o jedinicama koje su na pravom mjestu u pravo vrijeme, ili su poznate rukovodiocima istraživanja, ili je u pitanju nešto drugo.

Statistički metodi, najvećim dijelom, bazirani su na pretpostavci o slučajnom uzorkovanju. U najvećem broju statističkih istraživanja nije problem da se ispuni ova pretpostavka. S druge strane, neslučajan, odnosno namjeren izbor elemenata u uzorak uvijek može da dovede do manje ili veće pristrasnosti prilikom formiranja uzorka, a na taj način i do pristrasnosti u dobijenim vrijednostima parametara, kao i u zaključivanju.

Postoje četiri osnovne tehnike slučajnog uzorkovanja: jednostavno slučajno uzorkovanje ili izbor prostog slučajnog uzorka, stratifikovano slučajno uzorkovanje, sistematsko slučajno uzorkovanje i klaster uzorkovanje.

Izbor vrste uzorka, odnosno tehnike uzorkovanja koja će se koristiti, zavisi od predmeta i ciljeva istraživanja, kao i od prirode podataka koji se žele obezbijediti metodom uzorka. Primjena metoda uzorka podrazumijeva anketiranje, kao jednoobrazno prikupljanje podataka, koje se obezbjeđuje adekvatnim upitnicima i pripremom anketara i lica koja će ih popunjavati.

Rezultati istraživanja dobijeni primjenom reprezentativnog uzorka smatraju se više ili manje pouzdanom ocjenom najvažnijih karakteristika skupa iz koga su izabrani. Zbog toga je često, zavisno od prirode istraživanja, nužno provjeravati da li se navedene ocjene mogu smatrati relevantnim za cijeli skup. U tu svrhu, u statistici su razvijeni odgovarajući metodi ocjene i testiranja parametara skupa na osnovu uzorka.

Ocjenjivanje na osnovu uzorka, kao metod, najčešće se koristi kada istraživač nema nikakvu pretpostavku o veličini odgovarajućeg parametra skupa (prosječna vrijednost, proporcija i sl.) i podrazumijeva intervalno ocjenjivanje parametara skupa, gdje se odgovarajuća uzoračka karakteristika (statistika uzorka), koriguje zadatim koeficijentom pouzdanosti i utvrđenom standardnom greškom ocjene. Dobijeni interval povjerenja predstavlja raspon ili okvir vrijednosti za koji se vjeruje (sa određenim stepenom pouzdanosti) da sadrži i traženi parametar skupa.

S druge strane, testiranje statističkih hipoteza koristi se kada postoje određene pretpostavke o veličini odgovarajućih parametara skupa, koje je neophodno provjeriti. Taj postupak provjere naziva se testiranje statističkih hipoteza, a sama procedura provjere naziva se statističkim testom. U zavisnosti od prirode podataka koji se analiziraju i ciljeva istraživanja, mogu se koristiti parametarski ili neparametarski testovi, zasnovani na jednom uzorku, na dva ili više uzoraka. Smisao testiranja hipoteza je da se provjeri statistička značajnost razlika između uzoračke i hipotetičke vrijednosti odgovarajućeg parametra skupa na kome se varijabilna pojava ispoljava.

Mjerne skale i izbor statističkih metoda

Pošto se odabrane karakteristike jedinica statističkog skupa mjere, svako prikupljanje podataka podrazumijeva mjerenje. Mjerenje predstavlja pridruživanje brojeva ili određenih oznaka jedinicama skupa, prema određenom pravilu. Pravila pridruživanja data su mjernim skalama, tako da svako mjerenje

podrazumijeva korišćenje određene mjerne skale. Rezultati statističkog istraživanja mjere se korišćenjem sljedećih mjernih skala: nominalne, ordinalne, intervalne i skale odnosa.

Priroda podataka koji se prikupljaju i mjerne skale na kojima se obuhvataju opredjeljuje i izbor statističkih metoda za njihovu obradu i analizu. Izbor odgovarajuće mjerne skale je dio procjene prirode i validnosti statističkih podataka i direktno utiče na izbor odgovarajućih deskriptivnih statističkih mjera: mjera centralne tendencije, mjera disperzije i mjera oblika rasporeda.

Nominalna skala data je u vidu liste naziva, kategorija ili određenih atributa po kojima se jedinice statističkog skupa razlikuju. Ukoliko je riječ o atributivnim obilježjima koja imaju veliki broj modaliteta, klasifikacija pojedinih modaliteta vrši se u srodne grupe, u okviru posmatranog obilježja. Na taj način formiraju se jednoobrazne grupe i podgrupe, koje se najčešće nazivaju nomenklaturama (npr. nomenklatura zanimanja). Iz tog razloga nominalna skala je i najmanje precizna.

Podaci mjereni na nominalnoj mjernoj skali mogu se koristiti za izračunavanje relativnih brojeva (procenata) i modusa kao mjere centralne tendencije, i grafički prikazivati izvan koordinatnih sistema. Od statističkih metoda, moguće je koristiti Hi-kvadrat test (za testiranje međuzavisnosti različitih obilježja na bazi analize tabela kontingencije) i Koeficijent kontingencije, kao i Binomni test kada su u pitanju dihotomna obilježja.

Ordinalna skala se koristi ukoliko je moguće modalitete obilježja rangirati prema značaju u odnosu na usvojene kriterijume. Ova skala jedinicama skupa pridružuje brojeve, slovne oznake ili određene simbole, prema stepenu određenog svojstva ili odlike. Mjesto modaliteta na mjernoj skali predstavlja njegov rang, a ne određenu mjernu veličinu. Mjerna skala predstavlja redosljednu klasifikaciju, odnosno rang listu modaliteta obilježja, pri čemu njihov relativan značaj zavisi od broja modaliteta.

Od deskriptivnih statističkih mjera moguće je izračunavanje kvartila i medijane kao mjere centralne tendencije. Za podatke mjerene na navedenoj mjernoj skali izračunavaju se pokazatelji korelacije ranga, zavisno od toga da li je riječ o istraživanju međuzavisnosti dvije pojave (Spearman-ov koeficijent korelacije ranga) ili međuzavisnosti više pojava (Kendalov koeficijent saglasnosti).

Intervalna skala svakom modalitetu obilježja pripisuje određenu jedinicu mjere. Ovom skalom jedinicama skupa pridružuju se brojevi, pri čemu jednake razlike brojeva predstavljaju jednake razlike mjerene karakteristike. Intervalna skala omogućava utvrđivanje redosljeda modaliteta u skupu, kao i mjeru njihovog razlikovanja.

Za podatke koji su mjereni na intervalnoj mjernoj skali moguće je izračunavati aritmetičku sredinu kao mjeru centralne tendencije i standardnu devijaciju kao pokazatelj disperzije vrijednosti obilježja. Od statističkih metoda moguće je koristiti korelacionu analizu (koeficijente proste, parcijalne i višestruke korelacije, zavisno od broja pojava koje se posmatraju), regresionu analizu, t-test, analizu varijanse i faktorsku analizu.

Skala odnosa pokazuje i redosljed modaliteta i mjeru njihovog razlikovanja. Skala odnosa sadrži brojeve za koje važi pravilo da njihove jednake razlike predstavljaju i jednake razlike mjerene karakteristike jedinica skupa. Ova mjerna skala obezbjeđuje najviši nivo mjerenja. Skalu odnosa ne karakteriše samo upotreba jedinice mjerenja, nego i prava nulta tačka koja ukazuje na nepostojanje određene karakteristike. Zbog toga ova skala omogućava iskazivanje proporcionalnih odnosa modaliteta obilježja koja se mjere. Skala odnosa je najpreciznija mjerna skala.

Podaci mjereni na navedenoj mjernoj skali obrađuju se koristeći sve mjere deskriptivne statistike: pokazatelje centralne tendencije, mjere disperzije i mjere oblika rasporeda. Za analizu podataka mjenjenih na skali odnosa moguće je koristiti sve raspoložive metode statističke analize.

NAJVAŽNIJI METODI STATISTIČKE ANALIZE

Metodologija naučnog istraživanja, pored opštih naučnih metoda, svakako podrazumijeva i primjenu određenih statističkih metoda. Zavisno od vrste istraživanja, odabranog problema, predmeta i ciljeva istraživanja, najčešće se koriste sljedeći metodi: hi-kvadtar test, analiza varijanse, faktorska analiza, regresiona i korelaciona analiza, kao i metodi dinamičke analize i analize vremenskih serija.

Hi-kvadrat test (χ^2 test)

Hi-kvadrat test je zasnovan na hi-kvadrat distribuciji i koristi tipski za rješavanje nekoliko problema: Testiranje značajnosti razlike između opaženih i teorijskih frekvencija, testiranje međusobne povezanostirazličitih obilježja posmatrane pojave i testiranje jednakosti (ili razlike) proporcija tri ili više skupova. U sva tri navedena slučaja ovaj test ima odlike neparametarskog testa. Samo u slučaju kada se hi-kvadrat test koristi za testiranje varijanse, radi se o parametarskom testu.

1. Testiranje značajnosti razlike između opaženih i teorijskih frekvencija različitih rasporeda vjerovatnoće (test prilagođenosti, test saglasnosti ili test oblika rasporeda). Test prilagođenosti je statistički test koji treba da odgovori na pitanje, u kojoj mjeri su empirijski podaci, odnosno opservacije o broju pojavljivanja određenih vrijednosti ili modaliteta neke pojave, prilagođeni ili odgovaraju, ili prate, ili se ponašaju prema nekom teorijskom modelu rasporeda vjerovatnoće. Pri tome se polazi od pretpostavke da se posmatrana populacija prilagođava nekom teorijskom rasporedu vjerovatnoće. χ^2 test, prilikom testiranja prilagođenosti, može da se primijeni za bilo koju distribuciju populacije ili slučajne promjenljive, ali je to od posebne važnosti i najčešće primjenjivano u slučajevima uniformnog i normalnog rasporeda. Testiranje se svodi na provjeravanje značajnosti razlike između opservisanih i očekivanih frekvencija.

2. Analiza tabela kontingencije i testiranje međusobne povezanosti različitih obilježja posmatrane pojave. Kada su posmatrana obilježja mjerena na nominalnoj mjernoj skali i u slučajevima kada je moguće formirati tabelu kontingencije, sa dva ulaza (za jedno i za drugo obilježje), uz pomoć χ^2 testa može se dobiti odgovor na pitanje: Da li između dva obilježja elemenata jednog skupa postoji veza i da li je ta veza statistički značajna? Određivanje stepena međusobne zavisnosti dva obilježja jednog skupa, mjerena na nominalnoj mjernoj skali, kada su njihove vrijednosti samo riječima klasifikovane u različite grupe, može da se ustanovi na osnovu Pearson-ovog koeficijenta kontingencije.

3. Testiranje jednakosti (ili razlike) proporcija tri i više skupova, odnosno testiranje homogenosti posmatrane pojave. 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. Primjenom χ^2 testa mogu se dobiti odgovori na prethodno postavljena pitanja. Test jednakosti (razlike) proporcija za više populacija naziva se i test homogenosti.

Analiza varijanse

Da bi se istovremeno, jednim postupkom, ispitala jednakost aritmetičkih sredina više skupova, u statistici se koristi metod analiza varijanse (u literaturi se često navodi skraćenica ANOVA). Na prvi pogled može izgledati da naziv analiza varijanse nije odgovarajući, jer se ne testiraju varijanse već aritmetičke sredine. Ispitivanje jednakosti aritmetičkih sredina više skupova upravo se i provodi upoređivanjem odgovarajućih varijansi.

Analiza varijanse je u početku najveću primjenu našla u eksperimentalnim studijama u kojima je cilj istraživanje uticaja jednog ili više faktora na varijabilitet određene pojave. Pri tome, jedinice na kojima se primjenjuje neki eksperiment nazivaju se ekperimentalne jedinice. Ukoliko su to ljudska bića, nazivamo ih ekperimentalni subjekti. Međutim, danas se analiza varijanse široko koristi i u studijama posmatranja.

Navedena razlika između dva vida studija je veoma važna, jer u statistici samo putem eksperimenata možemo dokazati da postoji uzročna veza između pojava. Eksperiment se mora pažljivo i precizno dizajnirati, kako bi rezultati mogli da se generalizuju.

U studijama posmatranja samo se mjeri varijabilitet pojava, odnosno jedinica posmatranja i ni na koji način se ne utiče na njihove varijacije. Međutim, u eksperimentalnim studijama istraživač planira i kontroliše izbor ekperimentalnih jedinica i nameće im posebne uslove, s ciljem da se ispituju njihove reakcije, odnosno varijacije. Specifični ekperimentalni uslovi kojima se podvrgavaju ekperimentalne jedinice nazivaju se tretmani.

Najjednostavnija vrsta eksperimentalnog plana zasniva se na ispitivanju uticaja jednog faktora na varijabilitet posmatrane pojave, pri čemu se tretmani primjenjuju na slučajno odabrane eksperimentalne jedinice. Takav plan naziva se potpuno slučajnim planom, a sam postupak kojim se dobijeni podaci ispituju naziva se analizom varijanse sa jednim faktorom. Kada postoje indicije da na posmatranu pojavu značajno utiče više faktora, u statistici se primjenjuje analiza varijanse sa dva ili više faktora. U ovom slučaju se ne istražuje samo uticaj pojedinačnih faktora, nego i interakcija između posmatranih faktora, što je zapravo i najvažniji smisao analize varijanse.

Analiza varijanse se, prije svega, koristi kod atributivnih faktora. Ukoliko se faktori mogu iskazati numerički, tada je bolje koristiti metode regresione analize. S druge strane, zavisna promjenljiva mora biti mjerljiva bar na nivou intervalne mjerne skale.

Regresija i korelacija

Izučavanje međuzavisnosti dvije ili više kvantitativnih promjenljivih vrši se pomoću dva, vjerovatno najvažnija statistička parametarska metoda, korelacije i regresije.

Prilikom istraživanja međuzavisnosti varijacija dvije ili više numeričkih varijabli, u statistici se primjenjuju metodi regresione i korelacione analize. Kod istraživanja međusobnih veza dvije promjenljive primjenjuju se metodi proste regresione i korelacione analize, a u slučaju posmatranja više promjenljivih, metodi višestruke (multiple ili složene) regresije i korelacije.

Iako su ovi statistički metodi u bliskoj vezi i međusobno se dopunjuju, između njih postoje i značajne razlike.

Kod korelacione analize dvije pojave svedeno je koja pojava se označava kao nezavisna, 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. Ukoliko se ispituje zavisnost između pojava koje su mjerene na ordinalnoj mjernoj skali, odnosno ako su podaci dati u vidu rangova, za istraživanje njihove međuzavisnosti koriste se neparametarski koeficijenti korelacije. Za istraživanje zavisnosti dvije pojave najčešće se koristi Spearman-ov koeficijent korelacije, dok se u analizi međuzavisnosti tri ili više promjenljivih primjenjuje Kendall-ov koeficijent saglasnosti.

Kod regresione analize nužno je unaprijed identifikovati koja pojava će imati ulogu zavisne promjenljive, a koja nezavisne promjenljive. U statistici se kod regresije umjesto termina nezavisna promjenljiva koristi izraz objašnjavajuća promjenljiva ili regresor.

Svrha regresije je 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. Elementi jednačine regresije matematički se utvrđuju primjenom metoda najmanjih kvadrata.

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.

Faktorska analiza

Faktorska analiza je metod multivarijacione analize koji se koristi za opisivanje međusobne zavisnosti velikog broja promjenljivih, putem definisanja seta zajedničkih osnovnih dimenzija poznatih kao faktori. Zadatak faktorske analize je sažimanje većeg broja međusobno povezanih izvornih promjenljivih u manji broj zajedničkih faktora koji će ih opisivati i objasniti njihovu međusobnu povezanost.

Faktorska analiza se primjenjuje za identifikovanje dimenzija koje nisu odmah uočljive, za testiranje postavljenih hipoteza i za istraživanje novog područja. Pri tome se polazi od pretpostavke da među varijablama postoji linearna korelacija, a da svaki izdvojeni faktor nije u korelaciji sa drugim faktorima. Umjesto nad velikim brojem koreliranih varijabli analiza se provodi nad nekoreliranim faktorima i na taj način se otklanja problem kolinearnosti promjenljivih.

Može se govoriti o faktorskoj analizi u užem smislu čiji je zadatak da identifikuje osnovne faktore koji odražavaju zajedničke karakteristike varijabli, ili o komponentnoj analizi koja sumira većinu originalnih informacija u minimalan broj faktora za svrhe predviđanja.

Koji metod ekstrakcije će se izabrati, zavisi od ciljeva faktorske analize i od obima prethodnog znanja o varijansi promjenljivih.

Pored navedenog, za faktorsku analizu se može reći da predstavlja veoma koristan multivarijacioni statistički metod koji služi za efektivno ekstrakovanje informacija iz velikih baza podataka. Ona ukazuje na odnose koji nisu vidljivi iz ispitivanja sirovih podataka ili korelacione matrice i ima sposobnost da identifikuje setove srodnih varijabli čime istraživači imaju mogućnost ostvarenja boljeg razumijevanja strukture posmatranih podataka.

Analiza vremenskih serija

Postoje raznovrsni metodi statističke analize vremenskih serija. S obzirom na pristup analizi, ovi metodi se mogu razvrstati u dvije grupe: metodi koji analizu vremenskih serija pristupaju sa stanovišta vremena i metodi koji joj prilaze sa stanovišta frekvencija.

Prva grupa metoda koristi se u istraživanju modela razvojnih tendencija vremenskih serija u funkciji vremena, dok druga grupa metoda spada u područje harmonijske i spektralne analize.

Analiza vremenskih serija u funkciji vremena polazi od pretpostavke da na promjene posmatranih pojava tokom vremena utiču četiri komponente: razvojna tendencija pojave u dužem posmatranom periodu - trend (odnosno sekularna tendencija), ciklična komponenta (kolebanja koja se ponavljaju u određenim, često nejednakim, periodima od više godina), sezonske varijacije (koje se ispoljavaju u razmacima manjim od jedne godine i ponavljaju na isti način u dužem nizu godina) i neregularni uticaji (rezidualna komponenta), koji se pojavljuju kao slučajne varijacije pojave.

Analiza vremenskih serija pokušava da identifikuje obrazac ponašanja navedenih komponenti u prošlosti. Zatim, pretpostavljajući da će identičan obrazac da se nastavi u budućnosti, modeli vremenskih serija ekstrapoliraju ove obrasce da bi se prognozirale buduće vrijednosti pojave. Zbog toga se može reći da je jedan od glavnih ciljeva analize vremenskih serija prognoziranje budućih vrijednosti pojave.

Osnovna pretpostavka pri prognoziranju u analizi vremenskih serija je da će faktori koji su uticali na nivo pojave u prošlosti i sadašnjosti djelovati na isti način u budućnosti i da neće biti uključivanja novih faktora.

ZAKLJUČAK

Metodologija istraživanja smatra se najvažnijim dijelom projekta naučnog istraživanja. Ona sadrži osnovne i bitne odluke o tome kako će se istraživati prethodno određeni predmet posmatranja, kako će se istraživanjem ostvariti opšti i posebni ciljevi i kako će se provjeriti postavljene hipoteze.

U svakom istraživanju, bilo empirijskom ili teorijskom, primjenjuje se bar jedna opštenaučna metoda. U teorijskim istraživanjima se najčešće primjenjuje analitičko-deduktivni i aksiomatski opštenaučni metod. Najviše se koriste metod ispitivanja i analiza postojećih dokumenata. Za empirijska naučna istraživanja, pored hipotetičko-deduktivnog i metoda modeliranja, najviše se primjenjuju različiti statistički metodi.

Ni jedno empirijsko istraživanje nije moguće realizovati bez primjene određenih statističkih metoda. U prvom redu, to su metodi deskriptivne statistike (prikupljanje, sređivanje i grupisanje, obrada i prikazivanje

podataka). Zavisno od predmeta i ciljeva istraživanja, najčešće se koriste: hi-kvadrat test, analiza varijanse, metodi regresione i korelacione analize, faktorska analiza, kao i određeni metodi dinamičke analize i analize vremenskih serija.

LITERATURA

- [1] Breda, Z. (2010), *Network relationships and the internationalisation of the tourism economy*. PhD dissertation, University of Aveiro, Portugal
- [2] Kukić, S., Markić. B.(2006), *Metodologija društvenih znanosti*, Ekonomski fakultet Sveučilišta u Mostaru
- [3] Lovric, M. (Editor) (2011), *International Encyclopedia of Statistical Science*, Springer-verlag, Berlin Heidelberg
- [4] Lovrić, M., Komić, J., Stević, S. (2006), *Statistička analiza- osnovi i primjena*, Ekonomski fakultet, Banja Luka
- [5] Miljević, M. (2007), *Metodologija naučnog rada*, Filozofski fakultet, Istočno Sarajevo (Pale)
- [6] Velispahić, E., Duraković, S. (2011), *Osnovne osobine naučnoistraživačkog rada*, Prirodno-matematički fakultet, Sarajevo
- [7] Zelenika, R.(2000), *Metodologija i tehnologija znanstvenog i stručnog djela*, četvrto izdanje, Ekonomski fakultet, Rijeka