

DEFINISANJE PROBLEMA (PREDMETA) ISTRAŽIVANJA

Istraživanje u nauci obično započinje tako što postoji neki problem, volja za rešavanjem tog problema ili interesovanje za taj problem. Problem nekog istraživanja generalno predstavlja neku poteškoću sa kojom se istraživač sreće u teoriji ili praksi i koju on želi da reši, a može biti produkt svakodnevnog iskustva, potreba za testiranjem teorija ili može nastati kao nerešeno pitanje nekog prethodnog istraživanja. Naučni problem predstavlja svest o nepoznatom, tj. pitanje na koje trenutna naučna saznanja nemaju odgovor. Pored ovoga, naučni problem može biti i neka već istraživana pojava o kojoj se trenutni rezultati iz nekoliko istraživanja ne poklapaju; pojava čija se neka činjenica/činjenice ne mogu do kraja razjasniti postojećim naučnim znanjem; pojava čiji se rezultati ne uklapaju sa teorijom koja postoji na tu temu, itd...

Postoje dva osnovna tipa naučnog problema:

1. problemi koji se odnose na neka stanja u prirodi
2. problemi koji predstavljaju uzajamne odnose između dve ili više varijabli

Problem istraživanja mora biti veoma pažljivo odabran, pri čemu je to zaista izazov iako tako ne deluje na prvi pogled. Na samom početku planiranja istraživanja/master teze/doktorske disertacije, istraživač mora da zna koja oblast je njemu interesantna i shodno svojim afinitetima pronaći do sada nerešen problem u okviru te oblasti. Ukoliko je u pitanju master teza ili doktorska disertacija, kandidat sam osmišlja i predlaže temu mentoru, koju zatim zajedno razrađuju i usavršavaju shodno postojećim uslovima (trenutna znanja raspoloživa za rešavanje postojećeg problema, raspoloživa oprema, materijal, ljudi, prostor i finansije za rad). Opšta preporuka je da se najpre odredi šira tema, koja se zatim definiše u specifični problem istraživanja tj. naučni problem.

U procesu formulisanja naučnog problema postoje sledeći koraci:

1. definisanje skupa objekata koji sačinjavaju neku pojavu (naučni problem) i definisanje merljivih parametara koji opisuju ove objekte i njihove interakcije. Pri tome istraživač postavlja pitanje 'Koji će podaci biti sakupljeni?'. Preduslov za ovaj korak je izvesno prethodno razumevanje planiranog predmeta istraživanja (pojave, naučnog problema) na osnovu predznanja ili iskustva istraživača koji će se baviti ovom tematikom. Na primer, ako želimo da ispitujemo uticaj svetlosti određene talasne dužine na morfološke karakteristike lista neke biljne vrste, moramo najpre da odredimo šta ćemo meriti. Da li je to samo dužina ili je to znatno više parametara, npr. dužina, širina, izgled ivice, debljina nerava, broj žlezdanih dlaka itd.

- ☒ dodatno detaljno upoznavanje sa problemom iščitavanjem literaturnih podataka (teorijskih i eksperimentalnih u vezi sa naučnim problemom) i uže definisanje nepoznatih činjenica. Pri tome u stvari postavljamo pitanje “šta su drugi istraživači radili na tu temu?” i čitamo rezultate njihovih istraživanja, tj. šta su istraživali i šta su otkrili, tj. do kojih zaključaka su došli u vezi sa našim predmetom istraživanja. Na primer, nećemo ispitivati koja organela je odgovorna za fotosintezu, jer je poznato da su to hloroplasti pa je istraživanje na tu temu apsolutno nepotrebno. Sa druge strane, pitanje kakav je uticaj svetlosti na gene koji kodiraju enzime odgovorne za ovaj metabolički proces predstavlja izuzetno smisleno pitanje i aktuelnu temu za istraživanje. Traženjem literature na tu temu, utvrdićemo stepen istraženosti ovog problema i naći naš potencijalni udeo u doprinosu o ukupnom znanju na ovu temu. Naravno, u celom planu treba imati u vidu mogućnosti laboratorije u kojoj ćemo raditi istraživanje. Kako pretraživati literaturu u ovoj fazi istraživanja? Pri definisanju problema, neophodno je pretraživati dve vrste literature, tzv. konceptualnu (knjige, enciklopedije, udžbenici) i istraživačku (pregledati šta su drugi dobijali istražujući istu ili sličnu temu).

- ☒ definisanje tačnih izraza tj. termina (delova sistema, parametara) koji predstavljaju problem i načina njihovog ispitivanja (metoda i tipa istraživanja koji će biti upotrebljeni). U ovom koraku određuju se koje su osobine značajne za istraživanje, koji će odnosi između objekata biti istraživani, koje će tehnike i metode biti upotrebljavane, kao i koja će forma biti korišćena za prezentaciju dobijenih rezultata (nije obavezno u ovoj fazi).

- ☒ ocenjivanje da li je željeni problem podložan metodološkom ispitivanju u trenutnim, raspoloživim uslovima. Ukoliko se ispostavi da objektivno i više puta provereno (rađeno naučnim metodom) istraživanje ne možemo izvršiti, bolje je uopšte ne raditi takvo istraživanje jer ne doprinosi nauci zbog nedovoljne relevantnosti dobijenih podataka. Na primer, imamo sredstva da ponovimo eksperiment dva puta ili da imamo samo dve grupe za eksperimentalni tretman nekim faktorom. Ukoliko se ispostavi da su jako različite, a nemamo mogućnosti za dalja ponavljanja, takvo istraživanje nema nikakav značaj jer se ovi podaci ne mogu upotrebiti.

Pri definisanju predmeta istraživanja, još neki kriterijumi koje je preporučljivo da istraživač koristi su:

1. predmet istraživanja treba da bude nedovoljno istražen, jer je jako teško smisliti nov pristup i dobiti nova saznanja o nekoj veoma istraženoj temi (predmetu/pojavi)

2. predmeti istraživanja koji su kontroverza, tj. sporno pitanje u nauci NE SMEJU biti izbor prosečnog ili mladog istraživača

3. treba izbegavati previše široke ili previše uske termine koji će biti istraživani

4. Istraživač mora biti upoznat sa potrebama koje iziskuje istraživanje tog subjekta, jer kasnije može doći u situaciju da shvati da nije u mogućnosti da svoje istraživanje izvede do kraja. Ovo predstavlja gubitak dragocenog vremena i energije, pa je preporučljivo intenzivno pretraživanje literature na temu datog subjekta i upoznavanje sa metodologijom - tehnikama i aparatima koji se u njemu koriste. Pri tome uvek treba pokušati da se proceni mogućnost korišćenja istih takvih metoda na subjektu koji je izabran. U istu svrhu može poslužiti razgovor sa osobom koja ima iskustva u toj oblasti.

5. Veoma bitni kriterijumi za izbor teme (predmeta istraživanja) treba da budu i njegov značaj za nauku (na primer lokalizovana istraživanja areala rasprostranjenja nekih leptira na teritoriji Srbije ako se zna da je ova teritorija unutar njihovog već utvrđenog Evropskog areala; mogu imati značaj u smislu ugroženosti vrste i njihove kategorizacije za Crvene knjige...), troškovi istraživanja, vremenski faktor i slično. Na kraju, pre bilo kojih daljih koraka, istraživač treba da postavi sebi sledeća pitanja:

- (a) da li postoji prostor i oprema u kojoj mogu da završim svoje istraživanje?
- (b) da li se istraživanje nalazi u okviru budžeta koji mogu da priuštim?
- (c) da li su neophodni saradnici na tom istraživanju dostupni i voljni da sarađuju?

6. Formulacija predmeta istraživanja obično treba da bude praćena tzv. preliminarnim istraživanjem, posebno u slučaju ako tako nešto do tada nije rađeno. Ovo se vrši u cilju optimizacije metodologije i aparature sa minimalnim troškovima. Isto tako, ova istraživanja ukazuju na to da li je problem dobro definisan.

Značaj definicije problema ogleda se u izreci da je dobro definisan problem pola završenog posla na njegovom rešavanju. Ovaj postupak omogućava veoma jasnu definiciju i razdvajanje značajnih podataka od onih koji to nisu. U poslednjih nekoliko godina, istraživači ovom koraku ne pridaju veliki značaj što kasnije dovodi do problema u dobijanju podataka (kvalitetu ovih podataka) i njihovoj kasnijoj analizi. Da bi se definicija problema uradila kvalitetno, potreban je sistematičan pristup koji podrazumeva poznavanje osobina naučnog problema koji je pogodan za istraživanje (posebno u slučaju mladog naučnog radnika).

OSOBI NE DOBRO DEFINISANOG NAUČNOG PROBLEMA:

1. Naučni problem koji ste izabrali treba da bude veoma značajan za vas, posebno u smislu da postoji interesovanje o tome, predznanje o njemu i velika želja da se taj problem reši (istraži). Ovo je neophodno jer istraživač mnogo vremena provodi u rešavanju problema, radno vreme ne postoji, već je u pitanju intelektualni rad koji ne prestaje ni nakon izlaska sa posla. Zbog toga je neophodno da istraživač poseduje veliko interesovanje i naklonost ka predmetu kojim se bavi.
2. Predznanje o predmetu. Bez obzira na to na koji način je predznanje formirano (tokom školovanja i samo je teorijsko ili je iskustveno), potrebno je da istraživač bude u određenoj meri upoznat sa predmetom istraživanja. Ukoliko predznanje ne postoji, onda možda taj predmet nije najbolji izbor za istraživača jer će previše vremena posvetiti upoznavanju sa samim predmetom i pitanje je hoće li moći da ga reši na način na koji bi ga rešio istraživač koji ima iskustva sa istim predmetom.
3. Predmet treba da bude od značaja za nauku u smislu doprinosa opštem znanju ili primenljivosti
4. Dostupnost podataka o predmetu istraživanja. Treba proveriti da li postoji mogućnost sakupljanja svih relevantnih podataka u adekvatnom formatu iz izvora koji postoje i da je to moguće uraditi pre početka istraživanja.
5. Etika i moral. Pri formulaciji predmeta istraživanja, potrebno je i imati u vidu ove principe, bez obzira na to da li bi rad trebalo vršiti na ljudskoj ili na životinjskoj populaciji. Isto tako, trebalo bi isključiti svaku mogućnost da neko vaš rad okarakteriše kao plagijat, bez obzira na to što niste imali nameru da to uradite. Pored ovoga, nemoralni naučnici su skloni prisvajanju tuđih ideja, što se apsolutno ne može dokazati i prepušteno je isključivo moralnoj odgovornosti naučnika. Moralni principi naučnog radnika treba da budu jako čvrsti i ovo je glavni korak u kome se ogleda ova osobina naučnog radnika. Pri tome se možda dobija rezultat, ali se u isto vreme stiže nepoštovanje i gubitak poverenja kolega, kao i određeno mišljenje o naučniku u naučnim krugovima što može imati neverovatne posledice po njegov kasniji rad i napredovanje u karijeri.
6. Na istraživanju će raditi dobar tim.

RADNA HIPOTEZA

Nakon definisanja predmeta istraživanja u koje je bila uključena opsežna pretraga literaturnih i drugih podataka, potrebno je uraditi definisanje **hipoteze**, tj. pretpostavke o predmetu istraživanja (o rešenju problema) koja se proverava odabranim metodološkim postupkom. Pri tome se hipoteze formulišu na osnovu postojećih činjenica, npr. znamo da je svetlost neophodna za klijanje semena, pa ćemo kao hipotezu staviti tvrdnju da seme naše istraživane biljke zahteva određeni broj sati svetlosti tokom 24 h kako bi proklijala. Postoji tzv. nulta hipoteza, tj. ona pretpostavka koja tvrdi da razlike između nekih pojava ne postoje, npr. u efektu delovanja različitih tretmana (odrična je i govori suprotno od istraživačevih očekivanja). Nasuprot njoj, postoji alternativna hipoteza koja tvrdi suprotno od nulte. U prethodnom primeru vezanom za seme i klijanje, nulta hipoteza bi bila da svetlost nije neophodna za klijanje semena. Oporogavanje, tj. odbacivanje nulte hipoteze predstavlja pozitivan naučni rezultat koji se proverava statističkom obradom podataka.

Pri testiranju hipoteze mogu se očekivati dva tipa statističke greške: greška tipa 1 (hipoteza se odbacuje iako treba da bude prihvaćena) i greška tipa 2 (hipoteza se prihvata iako treba da bude odbačena).

DIZAJNIRANJE (PLAN) ISTRAŽIVANJA

Ovaj korak predstavlja najčešće jako zastrašujuć deo istraživačkog rada u kome treba postaviti temelje svih narednih postupaka na osnovu odluka o tome gde, kako, koliko i na koji način će se vršiti istraživanje. Po definiciji, 'Dizajn istraživanja je redosled koraka za sakupljanje i analizu podataka na način kojim će se obezbediti relevantnost podataka uz minimalne troškove procedure'.

Pri planiranju istraživanja, definišu se sve etape od momenta postavljanja hipoteze pa do analize podataka, pri čemu se stvara precizan plan koji se na dalje sledi. Osnovna pitanja na koja plan istraživanja treba da odgovori su:

- šta studija istražuje,
- zašto se to proučava,
- gde će istraživanje biti sprovedeno,
- koji tip podataka je potrebno obezbediti istraživanjem,
- gde se mogu naći ti podaci,
- koliko dugo će trajati istraživanje,
- u kojim periodima,
- da li je kontinuirano ili sezonsko istraživanje u pitanju,
- koji je materijal potreban i koliko,
- kako ću formulisati veličinu uzorka (koju ću statističku metodu koristiti),
- koja mi je oprema potrebna,

- da li je potrebno angažovati druge istraživače na njemu i koliko njih,
- u kojoj etapi sus neophodni drugi istraživači,
- kako će podaci biti obrađeni,
- kako pisati, tj. prezentovati rezultate (u kojoj formi)
- koji je minimalni budžet (troškovi) ovog istraživanja?

Dobar dizajn istraživanja razlikuje se od vrste do vrste istraživanja i karakteriše se osobinama poput efikasan, prilagođen potrebama istraživanja, fleksibilan, ekonomičan, precizan, relevantan itd. Ono što je najvažnije postići je pouzdanost dobijenih podataka (minimalna greška u podacima uz dobijanje što više informacija) i to treba da bude najosnovniji kriterijum pri planiranju istraživačke procedure.

U okviru plana istraživanja, potrebno je izvršiti i detaljno planiranje eksperimenta (ukoliko se radi o ovakvom istraživanju) ili terenskog rada. Kada se planira vaš eksperiment, potrebno je u obzir uzeti principe ustanovljene od strane R.A. Fisher-a koji se bavio poljoprivrednim eksperimentima na velikim zasejanim poljima. Pri tome je shvatio da su dobijeni podaci mnogo relevantniji ukoliko se polja podele na sektore i iz njih vrši uzorkovanje.

Ovi principi su princip ponavljanja, princip slučajnosti, princip lokalne kontrole.

Korišćenje **principa ponavljanja**, kojim tretman vršimo na više od jedne grupe (najmanje tri) povećavamo statističku preciznost (tačnost) rezultata jer umanjujemo efekat koji može imati neka osobina unutar specifične grupe koju koristimo kao uzorak. Pri tome, na primeru eksperimentalne grupe pacova, tretman se mora vršiti na bar tri grupe od 10 pacova i uporediti sa tri kontrolne grupe sačinjene od isto toliko pacova. Pri tome se manjuje verovatnoća da se jave lažni rezultati koji su posledica sastava same grupe, npr. pacovi koji čine jednu grupu su rođaci koji imaju nasledni metabolički problem. Kao posledica urođene bolesti, ovakvi pacovi reaguju drugačije na tretman. Ukoliko bi se koristila samo ova jedna grupa, rezultati istraživanja nekog tretmana dali bi pogrešnu sliku i rezultate o tom tretmanu. Uvođenje još nekih grupa smanjuje mogućnost ovakvih grešaka.

Drugi Fišerov **princip slučajnosti** pruža zaštitu od uticaja spoljašnjih faktora. Npr. ako gajimo neke varijetete biljke na istom polju koje je podeljeno na dve polovine (leva polovina na kojoj se gaji stari varijetet i desna na kojoj je novi varijetet), može se desiti da plodnost zemljišta ili sposobnost zadržavanja (fertilnost zemljišta ili nagib jednog dela je veći pa voda otiče) utiču na prinos zasejanih biljaka, a ne njegova genetička struktura. Iz ovog razloga, zasejavanje po principu slučajnosti u velikoj meri znači za smanjivanje ovakvih uticaja (slučajan raspored zasejanih biljaka u razbacanim kvadratima po površini polja). Ovako se treba voditi i pri uzorkovanju na terenu gde po ovom principu treba uzimati u obzir mikroklimatske karakteristike staništa.

Po **principu lokalne kontrole** treba obezbediti način da se poznati spoljašnji faktor (za koji znamo da unosi varijabilnost u rezultate) izmeri i isključi iz greške eksperimentalnih rezultata. Drugim rečima, isključuje se varijabilnost uzrokovana spoljašnjim faktorom koji ne možemo kontrolisati, ali ga možemo izmeriti i uključiti u analizu.

METODE KORIŠĆENE ZA ISTRAŽIVANJA U BIOLOGIJI

Metod posmatranja. Ovaj metod se koristi za sakupljanje kvalitativnih podataka i to pre svega u studijama ponašanja i sličnih kvalitativnih osobina u biologiji. Često se koristi u društvenim naukama poput sociologije (npr. marketinških istraživanja) i psihologije. Osnova za uspešno istraživanje ovog tip je kvalitetno planiranje istraživanja i toga kada i šta će se posmatrati. Za ovu studiju nije neophodan izlov jedinki niti direktan kontakt sa njima a u isto vreme njime se one posmatraju u prirodnom okruženju, što je od izuzetnog značaja za bihevioralne studije. Ovaj dizajn treba da omogući isključivanje svake pristrasnosti i ometanja posmatranog procesa i to su najveći izazovi ovog metoda. Nedostaci koji takođe mogu da budu značajni su finansijski prohtevi ovakvih studija, ograničenost dobijenih informacija ili nemogućnost sakupljanja informacija iz eksperimentom nepredviđenih razloga. Zbog toga se pri dizajniranju ovakvih studija planira šta će biti posmatrano, kako će se ta osobina beležiti i kako obezbediti preciznost dobijenih informacija

Empirijski metod. Ovaj metod služi za sva istraživanja u kojima ne postoji polazna hipoteza, već se na osnovu njega utvrđuju zakonitosti na osnovu kojih se ona naknadno postavlja. Ovaj metod se veoma često koristi u biomedicini, poljoprivredi, farmaciji, mikrobiologiji a kada se ova istraživanja tiču efikasnosti nekog tretmana na biljke, životinje ili organizme, a izvodi se kao pilot istraživanje koje prethodi eksperimentu.

Eksperimentalni metod. Pri korišćenju ovog metoda, osnovno pravilo je da se posmatra uzročno posledični odnos između nekih elemenata prirodne pojave/procesa, uz kontrolisanje što većeg broja ostalih faktora koji bi mogli da utiču na tu pojavu/proces. Pre eksperimenta se postavlja polazna hipoteza koju njime potvrđujemo ili opovrgavamo. Eksperimentom se dobija kvalitativni i kvantitativni opis neke pojave/procesa, nezavisno od vremena kada se oni javljaju u prirodi uz svesno izdvajanje željenih elemenata i zanemarivanje nepoželjnih elemenata te pojave. Eksperiment se vrši tako što se posmatra kvalitativni ili kvantitativni element (osobina, obeležje, karakteristika) koji se još naziva i VARIJABLA. Kvantitativne varijable se izražavaju brojevima i mogu biti kontinuirane (mogu biti bilo koji broj) i diskontinuirane (imaju samo pojedine vrednosti određenih kategorija). Kvalitativne varijable se izražavaju samo opisom - boja dlake, starost, prisustvo bolesti, itd. Razlikujemo zavisne i nezavisne varijable. Zavisna varijabla se još naziva i posledična i to je ona promenljiva koja se menja delovanjem nezavisne varijable (tako da promene u ovoj varijabli zavise od jačine nezavisne varijable). Nezavisna varijabla predstavlja hipotetičku uzročnu varijablu koja je izabrana dizajnom istraživanja kao ona kojom se utiče na nezavisnu varijablu. Pri planiranju eksperimentalnih istraživanja, od osnovnog je značaja dizajnirati eksperiment tako da se jasno odvoji efekat ove varijable od drugih nezavisnih varijabli. Eksperiment je dizajniran tako da se merenjima može utvrditi direktna uzročno posledična veza između ove dve varijable.

Induktivno deduktivni metod. Ovaj metod predstavlja kombinaciju dva suprotna misaona procesa, indukcije i dedukcije. On uključuje fazu sakupljanja činjenica tokom

eksperimenta ili posmatranjem (utvrđivanje svih podataka koji će dati naučnu istinu), a zatim se primenom dedukcije donose zaključci. Prema tome, induktivni zaključak služi kao premisa deduktivnom zaključku i obratno.

UZIMANJE UZORKA

Nakon odabira predmeta istraživanja i njegovog planiranja (dizajniranja), kao i pilot istraživanja, vrši se sakupljanje podataka čijom ćemo obradom doći do naučnih činjenica. Pilot istraživanje je postupak kojim se proverava izvodljivost i svrsishodnost istraživanja. Ono se vrši na jako malom uzorku i dobijaju preliminarni rezultati koji nam potvrđuju da je ono dobro osmišljeno, kao i da su odabrane adekvatne metode za dobijanje podataka. Pored ovoga, pilot istraživanje proverava da li je veličina uzorka dobra za statističku analizu kojom smo isplanirali da obradimo podatke.

Za svaku oblast u biologiji postoje specifične tehnike uzorkovanja (sakupljanja podataka) koje se koriste i koje se međusobno razlikuju u zavisnosti od toga šta se sakuplja. Pošto biologija proučava veoma različite nivoe organizacije počev od subcelularnog pa do nivoa ekosistema, za svaki od njih se primenjuje drugačija metoda uzorkovanja i druga veličina uzorka. Za sakupljanje je neophodno precizno definisati strategiju sakupljanja koja će obezbediti dovoljno materijala za sve buduće etape istraživanja. Na primer, ukoliko se vrši sakupljanje biljaka ili životinja na terenu, treba ga izvršiti na svim mikrostaništima jednog lokaliteta jer razlika u njima može dovesti do varijabilnosti rezultata. Kako bismo izbegli greške tipa 1 i 2, moramo obezbediti dovoljan broj jedinki na kojima će biti izvršena analiza, pa je od izuzetnog značaja i veličina uzorka. Za najtačnija istraživanja neophodno bi bilo sakupiti sve pripadnike (delove) osnovnog skupa koji se istražuje (CENZUS), što je u biologiji često nemoguće. Nije moguće sakupiti sve pripadnike populacije iz više razloga - prvi je neekonomičnost (finansijska i vremenska), zatim moral i etika u nauci ne odobravaju sakupljanje čitavih populacija jer bi samim tim one bile uništene; pokretljivost članova populacije je još jedno ograničenje jer se ne mogu svi uloviti, previše brojna populacija itd. Na primer, nemoguće je izbrojati sve mrave u jednom mravinjaku, sve ribe u jednom jezeru.. Zbog toga se vrši sakupljanje tzv. uzorka, koji je u statistici osnovna veličina koja se koristi za analizu a dobijene činjenice predstavljaju karakteristike cele populacije iz koje je uzet taj uzorak. Da bi uzorak dobro predstavljao osnovnu populaciju, on mora biti REPREZENTATIVAN, tj. on mora oslikavati sve osobine, njihovu varijabilnost i kvantitativnu zastupljenost koje su svojstvene članovima osnovne populacije (uzorak mora predstavljati celu populaciju po osobinama koje poseduje). Osnovna osobina reprezentativnog uzorka je da je uzet nepristrasno (objektivnost), pri čemu su apsolutno svi članovi populacije imali podjedanku šansu da budu uzorkovani, tj. da uđu u sastav uzorka. Zbog toga se u biološkim i drugim istraživanjima koristi nekoliko metoda uzorkovanja koje obezbeđuju reprezentativnost uzorka, a svi se zasnivaju na primeni principa slučajnosti. U slučaju istraživanja u kojima osnovna populacija ne postoji (citološka istraživanja, mikrobiološka istraživanja), ovaj metod se ne može primeniti ali postoje drugi načini za određivanje što veće preciznosti podataka (ponavljanje eksperimenta u kombinaciji sa formiranjem što brojnijih grupa za tretmane).

Za adekvatnu primenu principa slučajnosti pr uzorkovanju razvijeno je nekoliko metoda:

1. prost slučajni uzorak. Na terenu se koristi kada je lokalitet sa koga se sakuplja relativno uniforman i ne očekujemo velike varijabilnosti izazvane samim staništem, tj. lokalitetom. Pored ovoga, previše veliki lokalitet ili ograničeno vreme za sakupljanje su razlozi za korišćenje ove metode sakupljanja. Kada se koristi ova metoda, sakuplja se veliki broj uzoraka sa različitih delova staništa/lokaliteta. Za ovo sakupljanje se često koristi metod kvadrata koji se postavlja na lokalitet i sve životinje ili biljke unutar njega se broje, sakupljaju ili mere, u zavisnosti od cilja istraživanja. Sakupljanje biljaka i životinja koje se ne kreću ili se sporo kreću se vrši primenom ove metode, pri čemu se kvadrat određene veličine postavlja na određeni broj mesta na lokalitetu i uzimaju uzorci sa njih sa površine koju kvadrat zauzima. Ovo se ponavlja više puta na jednom lokalitetu kako bi se obezbedili što raznovrsniji uzorci i samim tim kvalitetniji podaci. Najčešća veličina površine kvadrata je 1 m², a svrha je obezbeđivanje uporedivih uzoraka sa površina iste veličine i oblika. Ovi oblici ne moraju biti kvadratni, već se u određenim istraživanjima koriste i okrugli ili pravougaoni oblici sa kojih se uzorkuje, ali je ovo potrebno posebno naglasiti i objasniti pri opisu metodologije. Uporedivost sa drugim podacima je, naravno, znatno manja nego pri korišćenju klasičnog kvadrata, jer razlika u oblicima jedinica uzorkovanja dovodi do varijabilnosti u rezultatima. Nedostatak ove metode je moguća subjektivnost pri izboru mesta na koja će se postaviti kvadrat. Takođe, izbor loše veličine kvadrata je od ogromnog značaja za relevantnost istraživanja, npr. za sloj krošnje šumskog ekosistema nemoguće je koristiti kvadrat veličine 1 m², već se koriste znatno veći kvadrati (bar 20 m²), dok se za npr. zajednicu mahovine na nekoj steni ili perifton na podvodnom predmetu koristi kvadrat veličine 25cm² ili manji. Obično se zabeleži procentualno učešće vrste u kvadratu tako što se proceni površina koju zauzima jedna vrsta.

Međutim, priroda istraživanja često ne dopušta vršenje ovakvog uzorkovanja, pa se moraju koristiti druge metode (uzorkovanje planktona iz vodenih ekosistema ne može koristiti ovu metodu, pa je neophodno uzorkovati standardizovanom zapreminom boce). Ukoliko se istražuje populacija parazitskih osa na drugim insektima, insekt domaćin biće jedna jedinica uzorkovanja (1 kvadrat). Pri tome se poređenje rezultata u poslednjem slučaju ne može izvršiti bez prethodnog računanja broja osa/gramu materijala, jer jedinice uzorkovanja mogu da se razlikuju po veličini (i težini) pa može doći do varijabilnosti rezultata uzrokovanom ovim.

U svojoj najprostijoj formi, slučajno uzorkovanje ovom metodom se postiže tako što se kvadrat baca unutar lokaliteta i materijal se sakuplja sa mesta na koje je on pao po principu slučajnosti. Međutim, ovo nije zadovoljavajuće jer istraživač može da svojim fizičkim ograničenjima ili nesvesnom pristrasnošću pri bacanju onemogući sasvim "slučajno" uzorkovanje. Drugo ograničenje metode je veličina kvadrata, tj. za ovaj metod mogu se koristiti samo mali kvadrati.

Drugi način korišćenja kvadrata za slučajno uzorkovanje je mapiranje lokaliteta i polaganje mreže sa numerisanim kvadratim preko nje, čime se dobija numerisanje svih delova lokaliteta. Svaki broj je kvadrat tačno određene veličine. Zatim se iz Tablice slučajnih brojeva vrši odabir kvadrata koji će biti mesta uzorkovanja. Tablice slučajnih brojeva se koriste i za druga istraživanja, na primer na populaciji ljudi ili životinja iz kojih se mora uzeti reprezentativni uzorak. Za ovakvo uzorkovanje neophodno je da se svim

pripadnicima populacije dodele brojevi, a zatim se iz tablice slučajnih brojeva (kompjuterski generisana) izvuku brojevi koji će biti sakupljeni kao deo uzorka. Odabir brojeva se ne može prepustiti istraživaču jer on pri tome i nesvesno može da bude pristrasan, pa je tablica najbolje rešenje jer se isključuje svaka mogućnost za favorizaciju pojedinih brojeva. Slučajnost se postiže i na taj način što se napravi spisak pripadnika populacije prema starosti, polu, abecednom redosledu prezimena itd. a zatim se izabere broj koji se koristi za biranje jedinki sa tog spiska. Na primer, izabrani broj 12 značiće da će svaka 12. jedinka sa spiska biti uzorkovana.

Sistematski uzorak. Sistematsko uzorkovanje radi se tako što se uzorci uzimaju u fiksiranim intervalima, obično duž neke linije. Ovo normalno uključuje korišćenje transekta gde je linija uzorkovanja podešena tako da prati smenjivanje ekoloških faktora na nekom lokalitetu. Tako se na primer linija transekta može prostirati duž linije koja vodi od centra nekog grada koji je veliki zagađivač (Bor) do najbliže planine kako bi se pratio uticaj koji ima zagađenje vazduha na kvalitativni i kvantitativni sastav populacije lišaja u tom regionu. Linijski transekt je proces u kome se duž linije konopca koji je obeležen na određenim intervalima (npr. 1 m) beleži neka informacija vezana za tu tačku. To mogu biti prisustvo/odsustvo neke vrste, ukupan broj vrsta na toj tački, itd. U ovoj metodi je od izuzetnog značaja kako će se postaviti linija transekta, tj. da prati gradijent određenog faktora koji je od značaja za ispitivanje. Postoji još jedna varijanta ove metode u kojoj se linija na kojoj se beleže podaci proširuje u pojas određene širine, pa se mogu odrediti i učestalost nekih vrsta a ne samo prisustvo ili odsustvo. U ovoj varijanti, na mesto gde je tačka na liniji transekta se postavlja kvadrat određene veličine i životinje se broje i/ili sakupljaju, a ukoliko su biljke u pitanju, određuje se njihova procentualna zastupljenost, određuje ukupna biomasa u kvadratu itd. (zavisno od cilja istraživanja).

Stratifikovano uzorkovanje se koristi za uzorkovanje iz različitih stratuma (slojeva) koji se nalaze unutar jednog staništa - koji se posmatraju kao zasebne jedinice koje se međusobno razlikuju po mnogim karakteristikama, pa se mora uzeti u obzir uzorak uzet iz svakog sloja kao deo reprezentativnog uzorka. Na primer, ako se vrši uzorkovanje populacije neke vrste sisara, tada su različiti stratumi na primer različite starosne grupe, grupe ženki i mužjaka, grupe bliskih srodnika, grupe određene visine, socijalnog statusa (ljudska populacija)... Jedino na šta treba obratiti posebnu pažnju je procentualna učestalost stratuma unutar osnovne populacije, jer se na osnovu toga procenjuje koliko će uzoraka biti izvučeno iz tog stratuma. Kao rezultat prethodno navedenog, broj uzoraka svakog stratuma zavisice od njegove veličine. U suprotnom, određenom stratumu se može dati veći ili manji značaj od onoga koji on realno ima, a dobijeni rezultati neće biti naučne činjenice.

Veličina uzorka je od izuzetne važnosti pri planiranju eksperimenta i ne postoje opšta pravila kojih se svaka grana biološke nauke mogla pridržavati. Osnovni pokazatelj adekvatne veličine uzorka je pilot studija u koju dobijamo preliminarne rezultate, čijom se obradom utvrđuje greška koja se javlja pri merenju. Ukoliko je ona velika, mora se povećati broj merenja/tretmana/jedinki u populaciji i sl. Najmanji broj ponavljanja eksperimenta je tri, dok je najmanji broj ispitivanih jedinica po nekim statističkim pravilima trideset. ovo se svakako ne može primeniti na sva biološka istraživanja i zbog toga se preporučuje

upotreba uzorka veličina koja se opisuje u metodologiji za datu oblast, kao i pilot studija pre nastavka istraživanja.

Varijabilnost koja može da postoji u rezultatima može biti posledica nehomogenosti materijala na kome se vrše neki tretmani i faktori koji su predmet istraživanja. Te faktore je u cilju što preciznijih rezultata neophodno ukloniti i na taj način utvrditi šta je posledica tretmana (a ne varijabilnosti skupa na kome se tretman primenjuje). Zbog toga pri planiranju eksperimenta moramo izabrati homogeny skup. U slučaju pacova na kojima će biti ispitivan neki tretman, homogenost se postiže odabirom pacove iste težine, pola, starosti, zdravstvenog stanja, a u isto vreme obezbeđuju se identični uslovi za sve grupe (ishrana, stres, svetlost/mrak periodi, prostor itd).

Kako bismo dodatno izdvojili efekat tretmana od varijabilnosti koje se javljaju bez mogućnosti naše kontrole, potrebno je obezbediti bar tri ponavljanja eksperimenta. Što je veći broj ponavljanja, to su rezultati pouzdaniji a standardna devijacija (rasipanje oko srednje vrednosti svih merenja) manja.

Kontrole u eksperimentima. Kako bismo jasno izdvojili efekat tretmana i izbacili svaku mogućnost da se u detektovanu uzročno posledičnu vezu između nezavisne i zavisne varijable umešala neka druga varijabla, neophodno je detaljno kontrolisati druge promenljive (faktori) tokom eksperimentalnog postupka. Zbog toga se uvode nove grupe sa kojima će se porediti eksperimentalne grupe, a nazivaju se kontrolne grupe. Ove grupe se tokom eksperimenta moraju nalaziti u apsolutno istim uslovima kao i tretirana grupa, sa tom razlikom da se na nju ne deluje nezavisnom varijablom (ispitivanim faktorom/tretmanom). Primenom ovakve tehnike izvođenja eksperimenta, kasnijim poređenjem jasnije se izdvaja efekat tretmana od efekta ostalih nezavisnih varijabli na zavisnu promenljivu (tretiranu grupu). Postoje pozitivne i negativne kontrole. **Pozitivne kontrole** pri ispitivanju efekta nekog tretmana bio bi tretman koji zasigurno deluje na način na koji očekujemo da deluje i ispitivan tretman (nezavisna varijabla). Na primer, u ispitivanju efikasnosti nekog novog antipiretika, pozitivna kontrola bio bi paracetamol kojim bi pacovi kontrolne grupe bili tretirani. Ova kontrola služi za kasnije poređenje efekta sa već poznatom supstancom koja izaziva željeni efekat i koristi se u slučaju da takva supstanca postoji (za testiranje antimikrobnog efekta neke supstance pozitivna kontrola bio bi antibiotik koji zasigurno antimikrobno deluje na ispitivanu bakteriju. Kasnijim poređenjem možemo videti da li je efekat ispitivane supstance bio bolji ili ne od već poznatog dejstva antibiotika). Pored ove, postoji i tzv. **negativna kontrola**, kojom se zavisne varijable ne tretiraju nezavisnom varijablom, a svi ostali uslovi su identični sa tretman grupama. Na taj način dobija se informacija o osnovnom stanju grupe, čime se izdvaja čist efekat testiranog tretmana. Eksperimentalni podaci se statistički obrađuju a prezentuju se najpre osnovni elementi deskriptivne statistike poput srednje vrednosti, standardne devijacije i sl., a efekti tretmana između grupa se porede nekim od postojećih statističkih testova.

OBRADA PODATAKA DOBIJENIH ISTRAŽIVANJEM

Osnovna terminologija u statistici za biologe:

Pri izlasku na teren ili radu u laboratoriji, biolozi sakupljaju podatke o predmetu ispitivanja koji se naziva **varijabla**. Varijabla je parametar čija vrednost varira pri merenju koje se ponavlja nekoliko puta (dužina antena kod nekog insekta, pH vrednost nekog vodenog staništa; prinos etarskog ulja iz iste mase biljnog materijala jedne vrste itd.).

Ponavljanja (replikacije) su više puta izvedena merenja na istom uzorku/varijabli (merenja izvedena na istom mestu u različitim danima; nekoliko mesta uzorkovanja iste vrste na isti način; nekoliko homogenih grupa na kojima se ponavlja isti tretman itd.).

Eksplorativne (deskriptivne) analize - predstavljaju se uglavnom grafički, putem histograma ili grafikona i njihova funkcija je da učine vidljivom obrazac, tj. vezu između nekih varijabli, da se ona razjasni i uprosti, kao i da se zbirno predstave podaci dobijeni sakupljanjem tokom istraživanja. Priroda tog odnosa se ovim analizama ne može utvrditi. Za ove analize nemoguće je da se statistički (numerički) potvrdi/opovrgne hipoteza. Takođe, veličina uzorka za ovu analizu je bar $10n$ uzoraka za proučavanje n varijabli. Primer: koja je veza između čistoće vode u Gabrovačkoj reci i prisustva većeg broja vrsta makrozoobentosa - ovo se predstavlja histogramom sa dve y ose, jedna bi bila čistoća vode (merena nekim parametrima tipa prisustvo kiseonika, fosfata i dr.) a druga bi bila broj vrsta u zajednici mzb.

Inferencijalne analize ispituju zakonitosti uzorka koje se projektuju na celu populaciju iz koje je uzorak uzet. Sastoje se u kvantifikaciji zavisne varijable kao funkciju nezavisne varijable. Npr. da li gustina zasejavanje biljaka utiče na njihov prinos ukupne biomase? Rešenje je da se ista biljka poseje na iste kvadrata na homogenim podlogama i da se zasejavanje vrši u različitim gustinama (10 cm između biljaka, 30 cm između biljaka itd.). Nakon vegetacionog perioda, mase biljaka sa svih kvadrata se mere i beleže. Odnos između P i G se određuje linearnom regresijom kao funkcija između ove dve varijable.

Statistički testovi se koriste kako bi se odredile sličnosti i razlike između uzoraka. Njima se testiraju nulta i alternativna hipoteza. Nulta hipoteza predstavlja hipotezu o nepostojanju razlike, dok je alternativna hipoteza iskaz o onome što istraživač veruje da je tačno u slučaju da su uzorački podaci doveli do odbacivanja nulte hipoteze. H_1 se u naučnim istraživanjima često naziva i istraživačkom hipotezom, jer njom istraživač izražava mišljenje koje postupkom testiranja nastoji da potvrdi. To je i razlog zbog kojeg u praksi često prvo postavljamo alternativnu hipotezu, a tek onda nultu. Alternativnu hipotezu, kojom moguća odstupanja stvarne od hipotetičke vrednosti parametra pratimo u oba smera nazivamo dvosmernom ili dvostranom hipotezom. Test koji se primenjuje u ovakvoj situaciji nazivamo dvosmernim ili dvostranim testom (piše se $H_0: \mu = x$; $H_1: \mu \neq x$). Za

razliku od njega, test koji ispituje hipotezu koja se piše $H_0 : \mu \leq x$ i $H_1: \mu > x$, naziva se jednosmerni ili jednostrani test.

Kada koristiti dvosmerne, a kada jednosmerne testove? Ako testirana vrednost parametra ne sme odstupati u bilo kom smeru od standarda prihvaćenog u praksi (hipotetičke vrednosti), ili ako unapred ništa ne znamo o potencijalnom odstupanju parametra od njegove hipotetične vrednosti, onda koristimo dvosmerni test.

Iako naše uverenje formulišemo u obliku alternativne hipoteze, u postupku testiranja proveravamo samo nultu hipotezu. Polazimo od pretpostavke da je H_0 istinita i nastojimo da ovu pretpostavku osporimo. Alternativnu hipotezu ne proveravamo, već je automatski prihvatamo kao istinitu ako podaci ubedljivo "svedoče" protiv nulte hipoteze.

Greške pri testiranju i nivo značajnosti testa. Nulta hipoteza, kao tvrđenje o vrednosti 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 protivrečiti. To znači da kod testiranja, kao i kod ocenjivanja, postoji mogućnost da donesemo pogrešan zaključak. Kako uzorak nikada nije savršeno reprezentativan, moguća su i sledeća dva ishoda: da informacija iz uzorka protivreči istinitoj nultoj hipotezi (greškom prve vrste) ili da se saglasi sa neistinitom nultom hipotezom (greška druge vrste). Zato se pre statističkog testa definiše **nivo značajnosti testa ili nivo signifikantnosti** (naziva se još rizikom greške prve vrste) i predstavlja verovatnoću da ćemo odbaciti istinitu nultu hipotezu. Obeležava se sa α .

Za razliku od njega, verovatnoća da nećemo odbaciti netačnu nultu hipotezu naziva se rizikom greške II vrste i obeležava se sa β . Postupak testiranja sprovodimo tako što unapred fiksiramo rizik greške I vrste, tj. nivo značajnosti α . Pri tome biramo relativno mali nivo značajnosti, ali ne toliko mali da on onemogućava odbacivanje svake nulte hipoteze. Koliko mali? Većina istraživača koristi uglavnom samo dva nivoa značajnosti: 0,05 i 0,01. Ako odaberemo najčešće korišćeni nivo značajnosti $\alpha = 0,05$, onda svesno unapred prihvatamo rizik da ćemo u 5% slučajeva odbaciti H_0 iako je ona istinita. To će se dogoditi ako izaberemo jedan od 5% nereprezentativnih uzoraka (u proseku jedan od dvadeset mogućih uzoraka) koji po svojim karakteristikama značajno odstupaju od karakteristika skupa.

Izbor testa. Pri izboru testa treba da izaberemo onaj koji najviše odgovara empirijskim podacima (tj. test za koji su zadovoljeni polazni uslovi), ali i test koji će za izabrani nivo greške I vrste imati najveću jačinu, i to za unapred izabranu veličinu uzorka, n . Rešenje problema izbora nam daje teorijska statistika: ako je raspored skupa normalan, najveću jačinu imaju parametarski testovi. U suprotnom se koriste neparametarski testovi.

U postupku testiranja polazimo od pretpostavke da je nulta hipoteza istinita. Drugim rečima, pretpostavljamo da uzorak biramo iz skupa u kome je stvarna vrednost parametra jednaka hipotetičnoj vrednosti. Hipotetičku vrednost aritmetičke sredine skupa obeležavamo sa μ_0 , a proporcije skupa sa π_0 .

Statistike testa najčešće imaju sledeći oblik:

$$\text{Statistika testa} = \frac{\text{Ocena parametra} - \text{Hipotetična vrednost parametra}}{\text{Standardna greška ocene}}$$

Tako, na primer, statistika testa za testiranje hipoteze o aritmetičkoj sredini skupa, u zavisnosti od ispunjenosti polaznih uslova, može biti:

$$Z = \frac{\bar{X} - \mu_0}{\sigma_{\bar{X}}} \quad \text{ili} \quad t = \frac{\bar{X} - \mu_0}{S_{\bar{X}}}$$

Ako je nulta hipoteza istinita, tada očekujemo da razlika između ocenjene i hipotetične vrednosti parametra bude mala, pa će i vrednost statistike testa biti mala. U suprotnom, očekujemo relativno veliku (bilo pozitivnu, bilo negativnu) vrednost statistike testa. Vidimo da je značaj statistike testa u tome da ona predstavlja meru usaglašenosti između podataka iz uzorka i nulte hipoteze.

Mala vrednost statistike testa nam ne daje osnove da odbacimo nultu hipotezu. Nasuprot tome, njene velike vrednosti nam sugerišu da nulta hipoteza nije tačna. Što je veća vrednost statistike testa, to su uverljiviji dokazi da je nulta hipoteza pogrešna. Ostaje samo da se odredi koje vrednosti statistike testa ćemo smatrati malim a koje velikim, odnosno, kolika treba da bude vrednost statistike testa da bi se nulta hipoteza odbacila.

Postoje dva načina da donesemo konačan zaključak o odbacivanju ili neodbacivanju nulte hipoteze. To su Nimen-Pirsonov (Neyman-Pearson) pristup, i Fišerov pristup, zasnovan na primeni tzv. p-vrednosti. Testiranje hipoteza se u praksi najčešće sprovodi primenom računara, kada ćemo p-vrednost naći u „izlazu“ svakog statističkog softvera. Po definiciji, **p-vrednost je verovatnoća da statistika testa uzme vrednost jednaku ili još ekstremniju od vrednosti koja se upravo realizovala u uzorku, pod uslovom da je nulta hipoteza tačna. Što je manja p-vrednost, jači su dokazi protiv nulte hipoteze.** Pravilo odlučivanja na osnovu p-vrednosti glasi: Ako je p-vrednost manja od nivoa značajnosti α , nulta hipoteza se odbacuje. U suprotnom, kazaćemo da nemamo dovoljno argumenata da odbacimo H_0 .

Ako je p vrednost manja od nivoa značajnosti α , kažemo da je dobijeni rezultat statistički značajan ili signifikantan, i nultu hipotezu odbacujemo sa rizikom α . U suprotnom, kažemo da rezultat nije statistički značajan.

Postoje parametarski i neparametarski testovi, pri čemu su parametarski oni koji se koriste za uzorke koji imaju normalnu distribuciju i izražavaju se parametrima adekvatnim za tu distribuciju poput srednje vrednosti i standardne devijacije (za njih je potreban veći uzorak nego za neparametarke testove). Normalna distribucija je veoma česta u biologiji, jer označava populacije kod kojih su vrednosti oko srednje vrednosti jako česte, a ekstremno

niske i visoke predstavljaju ređu pojavu. Zbog toga njihova raspodela ima oblik zvona (normalna ili Gausova raspodela). Veličina uzorka za ove testove je velika, jer je neophodno bar 30 tačaka (izmerenih vrednosti) po varijabli koja se ispituje. Minimalni zahtevi za neparametarski test su bar 6 tačaka po varijabli.

Za mogućnost obrade podataka statističkim metodama, neophodno je:

1. dobro formulirati biološko pitanje. Ovaj korak podrazumeva definisanje glavnog problema koji se ispituje (koji problem ja istražujem?), kao i preciznu definiciju konkretnog biološkog pitanja koje se postavlja (ja tražim odgovore na koja konkretna pitanja?). Odgovor na prvo pitanje bio bi na pr. uticaj insekticida na insekte, a na drugo - koja veza postoji između doze insekticida i populacije insekata, kao i da li je taj odnos isti za različite vrste komaraca i isti taj insekticid? Pri formulisanju pitanja, neophodno je problem podeliti na više pitanja koja moraju da budu što preciznija i kraća, kao i da se odnose na samo jedan problem.
2. Pri postavljanju problema koji moramo rešiti statističkom obradom, moramo definisati promenljive, tj. varijable i napraviti njihovu listu. Takođe treba definisati u kojim jedinicama se one mogu predstavljati (mm, ml...).
3. Koliko ponovaka će biti i koliki nam je uzorak?
4. Koje će vrednosti biti na preseku varijable i uzorka u tabeli?
5. Kojom metodom ćemo sakupiti te podatke?
6. Podaci koji su nominalni moraju se pretvoriti u numeričke (npr. varijabla je boja pa je plavo-1, crveno-2, crno-3 itd.)
7. podaci trebaju da budu predstavljeni kao tabele u kojima su kolone varijable a redovi uzorci. Ovo je opšte pravilo uređivanja podataka pre statističke obrade jer inače rezultati neće biti dobri.

Koraci u statističkom testu.

- definisanje hipoteze
- izbor nivoa značajnosti
- izbor test statistike
- izračunavanje statistike testa
- statistički zaključak

Deskriptivna statistika.

Aritmetička sredina predstavlja prosek, tj. srednju vrednost, a suštinski to je vrednost oko koje se kreće najviše rezultata u toj pojavi.

Medijana (centralna vrednost) je vrednost koja se u nizu rezultata nalazi tačno u sredini. Deli dobijene rezultate na dve polovine, polovinu sa manjom vrednošću od nje i polovinu sa većom vrednošću od nje.

Dominantna vrednost je ona koja u nizu merenja dominira, tj. najčešće je postignuta.

Mere varijabilnosti. Kod merenja se rezultati grupišu i skupljaju oko jedne srednje vrednosti. Zbog toga je i računamo jer ona predstavlja reprezentativnu vrednost naših rezultata. Pošto rezultati mogu biti gusto ili ne grupisani oko nje, sama srednja vrednost nije adekvatno merilo reprezentativnosti, pa je uvek potrebno znati kako se vrednosti grupišu oko nje.

Raspon je razlika između najvećeg i najnižeg rezultata. Predstavlja nepouzdanu meru varijabilnosti jer bilo koji ekstremni rezultat može značajno povećati raspon, iako grupacija oko srednje vrednosti može ostati relativno nepromenjena. Osnovni nedostatak ove vrednosti je što na nju utiče broj merenja, pa veći broj merenja uvek rezultuje većim rasponom za većinu pojava.

Srednje odstupanje je prosečna veličina odstupanja pojedinačnih rezultata od srednje vrednosti.

Varijansa je mera varijabilnosti nekog parametra unutar populacije i odgovara rasipanju podataka oko srednje vrednosti.

Standardna devijacija je vrednost odstupanja od aritmetičke sredine i njena veličina nam ukazuje da li je aritmetička sredina dobar represent populacije. Zbog toga se uz srednju vrednost uvek navodi i standardna devijacija. Praktična primena SD je u tome da predvidimo na osnovu nje u kom rasponu će se kretati svi rezultati.

Koeficijent varijabilnosti. Služi za određivanje razlike u variranju nekog svojstva u grupama ili u kojoj osobini grupa varira više, a u kojoj manje.

U statistici je način opisivanja izvesnosti nekog procesa predstavljen terminom verovatnoća. verovatnoća da će se desiti događaj A se piše kao $p(A)$, a može imati vrednost od 0-1. Kada je $p=0$, događaj se ne može desiti a ako je $p=1$, izvesno je da će se desiti. P vrednost je mera koliko dokza imamo protiv nulte hipoteze. Što je ona manja, to su jači

dokazi protiv nulte hipoteze. Ako je $p < 0,01$ to je veoma jak dokaz protiv H_0 , ako je $0,01 < p < 0,05$, onda je dokaz protiv H_0 umeren, a ako je $0,10 < p$, onda nema realnih dokaza protiv H_0 .

EKSPLORATIVNA ANALIZA U BIOLOGIJI. Koristi se kada postoji više varijabli koje opisuju neku pojavu i koje se moraju grupisati po karakteristikama i međusobnoj zavisnosti, iako se ovom analizom ne opisuje karakter tog odnosa. Na primer, odgovor vodenih vrsta na uslove sredine, kojih ima mnogo bio bi proučavan ovom metodom. Pri tome bi se izvršila grupacija faktora sredine i određenih parametara tih organizama koji se odnose na njihov odgovor na te uslove, na primer: preživaljavanje, veličina tela, sposobnost reprodukcije itd, dok bi faktori bili protok vode, prisustvo gasova, teških metala, količina svetlosti i drugi.

Postoje dve osnovne eksplorativne analize i to PCA (Principal component analysis) i CA (Correspondance analysis, COA). Obe metode imaju funkciju da vizuelno sumarizuju sve analizirane varijable i prikažu njihove međusobne odnose.

Pravilo pri formiranju tabele podataka koji će biti analizirani je da su uzorci raspoređeni u redove, a varijable koje se ispituju u kolone! Razne metode multivarijantne analize su klasifikovane u tri osnovne grupe:

1. analiza koja ispituje samo jednu tabelu podataka
2. an. koja istovremeno ispituje dve tabele sa podacima u isto vreme i upoređuje ih
3. an. koja istovremeno analizira i poredi više tabela sa podacima

U prvu grupu spadaju PCA I CA, gde se ispituju podaci iz samo jedne tabele i predstavlja najprostiji oblik eksplorativne analize. One se ujedno i najčešće koriste u biologiji.

STATISTIČKI TESTOVI ZA POREĐENJE UZORAKA

Izbor test statistike se vrši na osnovu tipa podataka (numerički ili ne, kontinuirani ili diskontinuirani itd), broja uzoraka/grupa i njihove raspodele (normalna ili ne).

Parametarski testovi

Ovi testovi su bazirani na ocenama jednog ili više populacionih parametara (srednja vrednost, SD) dobijenih na osnovu podataka iz uzorka. Ovi testovi se koriste za testiranje hipoteza o populacionim parametrima, npr. o jednakosti aritmetičkih sredina dve populacije. U njima se pretpostavlja da osnovni skup ima normalnu raspodelu (najveći broj podataka je grupisan oko srednje vrednosti, a rasipanje podataka je malo), a broj uzoraka

ne sme biti manji od 30. Pretpostavka koja je takođe uslov za izvođenje parametarskih testova je da su populacije sa istim varijansama u pitanju. Normalnost se može testirati određenim statističkim testovima npr. Kolmogorov Smirnov test, Shapiro Wilk test, a ako je $p < 0,05$ u ovim testovima, empirijska raspodela značajno odstupa od normalne raspodele.

U najpoznatije parametarske testove, koji testiraju PARAMETRE populacije spadaju T test (srednja vrednost), F test i ANOVA (SD- varijansa).

T test je parametarski test koji se koristi kada nije poznata SD populacije, a zna se da je ona normalno distribuirana i ima najmanje 30 uzoraka. Pored uslova normalnosti za ovaj test uveden je i dopunski uslov da su varijanse dva skupa među sobom jednake. Ovaj uslov se u statistici naziva pretpostavkom homogenosti varijansi. T test može koristiti i kada je pretpostavka homogenosti varijansi narušena, ali pod uslovom da oba uzorka imaju jednak broj elemenata. Zbog toga postoji preporuka istraživačima da istraživanje sprovede na uzorcima jednake veličine. Koristi se za upoređivanje srednjih vrednosti dva uzorka i to u situacijama gde postoje dve varijable, pri čemu nezavisna varijabla ima dve kategorije (npr. uticaj svetlosti kao varijabla se opisuje kao prisustvo/odsustvo) i obe utiču nekako na zavisnu varijablu (brzina klijanja). Ovim testom se vrši upoređivanje efekta ta dve kategorije na zavisnu varijablu. Nulta hipoteza bi u ovom slučaju bila da nema razlike između ovih tretmana i ona se testira t testom kako bi se odredila P da su uzorci iz iste populacije, tj. da nema statistički značajne rezlike između njih.

Najpoznatiji parametarski test je ANOVA (analiza varijanse), koja je uvedena od strane Fishera 1923. godine kao metod razdvajanja efekta različitih tretmana u naučnom eksperimentu. Ovim testom se ispituje da li postoje razlike između srednjih vrednosti više populacija ispitivanjem kolčine varijacije između tih populacija. Uslovi za izvođenje ovog testa su isti kao i kod T testa. Anova može biti jednofaktorska, dvofaktorska ili višefaktorska.

Jednofaktorska ANOVA bavi se ispitivanjem uticaja jedne nezavisne promenljive (faktor uticaja koji mi kontrolišemo pri eksperimentu) na jednu zavisnu promenljivu. Faktor uticaja ima tri ili više kategorija (grupa- npr. tri koncentracije NaCl kojima tretiramo biljke). Veličina kategorija može biti različita ali je veći broj unutar kategorije bolji za relevantnost rezultata.

Za razliku od nje, višefaktorska ANOVA prati više od jednog faktora koji imaju tri ili više kategorija.

Neparametarski testovi ne zahtevaju poznavanje oblika raspodele u osnovnom skupu, kao ni normalnost te raspodele kao što je slučaj sa parametarskim testovima. Primenuju se u svim uslovima u kojim se ne mogu primeniti parametarski testovi. Za svaki parametarski test postoji bar jedan ekvivalentan neparametarski test.

Mann Whitney test predstavlja neparametarski pandan t testu i koristi se za upoređivanje medijana dva uzorka ali kada se ne zna raspodela i broj uzoraka je manji od 30.

Wilcoxon-ov test predstavlja pandan uparenom t testu, a Kruskal Wallas- ov test neparametarski ANOVA test.