POSTAVLJANJE I TESTIRANJE HIPOTEZA

Hipoteza je precizno formulisana verbalna tvrdnja, pretpostavka o karakteristici jednog skupa ili o odnosu vrednosti posmatrane karakteristike u više skupova.
U statističkim istraživanjima polazi se od dve međusobno isključive, suprotne pretpostavke o ishodu ispitivanja: nulte (H0) i alternativne (Ha) hipoteze.

Nulta hipoteza (Ho) glasi: *Između aritmetičkih sredina dva osnovna skupa, odnosno između aritmetičkih sredina dva uzorka dobijena iz dva osnovna skupa ne postoji značajna razlika. Ako razlika i postoji ona je slučajnog karaktera, odnosno nastala je pod dejstvom slučajnih faktora i uzorci se ponašaju kao da pripadaju istom osnovnom skupu.*

Matematički se definiše: $Ho : \overline{x_1} = \overline{x_2}$ ili $\overline{x_1} - \overline{x_2} = 0$

Alternativna hipoteza (Ha) se pridružuje nultoj hipotezi i tvrdi suprotno: *Između aritmetičkih sredina uzoraka postoji značajna razlika i ona nije slučajnog karaktera, već je nastala pod dejstvom sistemskih ili eksperimentalnih faktora.*

Matematički se definiše: $Ha : \overline{x_1} \neq \overline{x_2}$ ili $\overline{x_1} - \overline{x_2} \neq o$

Pravila za formulisanje hipoteze:

1. Kao Ha treba definisati onaj ishod za koji istraživač predpostavlja da je tačan i istinit.
2. Ho treba da sadrži pretpostavku koja je po istraživaču pogrešna i neistinita i koju on primenom testa treba da ospori.
3. Prvo treba postaviti Ha, a zatim doći do Ho

U istraživanjima se uvek polazi od pretpostavke da je nulta hipoteza istinita i da „razlika“ nije statistički signifikantna (lat. signifikans – znatan, značajan, karakterističan).

Naučni metod, tj. postupak provere pretpostavki predstavlja *testiranje statistističkih hipoteza.* Testira se *isključivo nulta hipoteza*, a alternativna se prihvata ili odbacuje posredno.

Istinitost nulte hipoteze se utvrđuje specifičnim *testovima značajnosti*.

Ako se odgovarajućim statističkim testom, za odgovarajuću verovatnoću i prag značajnosti, *dokaže istinitost nulte hipoteze* ona se ne odbacuje, već se *odbacuje alternativna hipoteza* i zaključuje se: pretpiostavka istraživača, data kroz alternativnu hipotezu, nije tačna, jer dobijena *razlika* između realizovane vrednosti (vrednosti dobijene posle israživanja ili posle eksperimenta) parametra skupa i hipotetične vrednosti *nije statistčki značajna (signifikantna)*. Razlika je posledica slučajnog karaktera i nastala je dejstvom sporednih faktora.

Obrnuto, ako statistički test, za odgovarajuću verovatnoću i prag značajnosti, *ne potvrdi istinitost nulte hipoteze,* onda je odbacujemo i automatski *prihvatamo alternativnu hipotezu* kao istinitu i zaključujemo: *razlika je statistički signifikantna* i verovatno je nastala pod uticajem sistemskih, odnosno eksperimentalnih faktora.

Pri testiranju nulte hipoteze mogu da se naprave dve greške:

1. *greška prve vrste (α-greška)*, koja nastaje kada odbacimo istinitu nultu hipotezu i
2. *greška druge vrste (β-greška)* nastaje kada prihvatimo (ne odbacimo) netačnu nultu hipotezu

Verovatnoća ostvarenja greške prve vrste (α-greške) su *pragovi značajnosti.* Znači, prag ili pragovi značajnosti su verovatnoća rizika da nulta hipoteza bude odbačena iako je istinita. Što je p vrednost manja, to je manje verovatno da je nulta hipoteza tačna, dok velika p vrednost sugeriše prihvatljivost nulte hipoteze. Ta tvrdnja vredi sa verovatnoćom P koja je jednaka razlici 1-p.

Često na osnovu rezultata statističkih testova donosimo odluke (npr. proizvodnja novog leka, primena nove dijagnostičke metode i sl.). To nas prisiljava da odredimo granični stepen značajnosti (tj. graničnu p vrednost). Kažemo da se uzorci statistički značajno razlikuju ako je dobijena p vrednost testa manja od granične p vrednosti. Ako to nije slučaj, tj. ako je p vrednost veća ili jednaka graničnoj p vrednosti, kažemo da se uzorci ne razlikuju signifikantno. U prvom slučaju nultu hipotezu odbacujemo, a u drugom je prihvatamo. Većina kriterijuma dopušta mogućnost čuvene granične greške od 5%, tj. $p \leq 0,05$ i verovatnoće $P > 95\%$.

Verovatnoća da nećemo napraviti grešku druge vrste (β-grešku) povezana je sa *snaga (power) testa.* Snaga testa *(γ)* je verovatnoća da se odbaci netačna H_0 hipoteze i jednaka je razlici između 1 i verovatnoće greške druge vrste.

Odnos verovatnoća je:

$\beta + \gamma = 1$

$\gamma = 1 - \beta$

Verovatnoća snage testa je utoliko veća ukoliko je verovatnoća greške druge vrste manja.

Snaga istraživanja raste s veličinom uzorka (N) i s veličinom razlike koju smatramo stvarno značajnom, a opada s nivoom značajnosti. Što je snaga testa veća, test je bolji.

Bilo koji tip testiranja statističkih hipoteza može da se izvede u sledećim etapama:

1. Formulisanje nulte i alternativne hipoteze;
2. specifikacija teorijskog rasporeda verovatnoće i izbor odgovarajućeg testa značajnosti i izračunavanje vrednosti testa značajnosti
3. određivanje p vrednosti I odluka o prihvatanju ili odbacivanju nulte hipoteze zasnovane na realizovanoj vrednosti testa značajnosti i određenoj p vrednosti.

Ponekad unapred znamo da neko obeležje populacije A ne može biti prosečno manje nego u poulaciji B, a nismo sigurni da li je veće (ili obrnuto). Npr. ispitujemo novi lek za sniženje krvnog pritiska. Poznato je da taj lek ne može povisiti krvni pritisak, ali nije poznato da li ga smanjeje i koliko. Tada se testiranje radi *jednosmernim (*engl. *one-tailed) testovima.*
Kada se unapred ne može sa sigurnošću odrediti smer neke razlike, primenjuje se *dvosmerni (*engl. *two-tailed)* test. Ako npr. nije određen smer razlike u procentu pušača, tj. da li je procenat pušača kod muškaraca veći ili manji u odnosu na žene u populaciji, primenjuje se dvosmerni test. Ako se pri testiranju dobije pozitivna vrednost testa, znači da je smer razlike pozitivan, tj. u našem primeru bi smo zaključili da je procenat pušača kod muškaraca veći nego kod žena Ako se pri testiranju dobije vrednost testa sa negativnim predznakom, to znači da je smer razlike negativan, tj. u našem primeru bi to značilo da je procenat pušača kod muškaraca manji nego kod žena.
Jednosmerna testiranja treba kritički koristiti. U praksi su retka. Ako se posebno ne istakne suprotno, testiranje se vrši dvosmernim testom.

Testovi na kojima se temelji postupak dokazivanja ili odbacivanja nulte hipoteze nazivaju se testovima značajnosti (signifikantnosti) i globalno mogu da se podele u dve grupe:

- parametrijski
- neparametrijski testovi.

PARAMETRIJSKI (NUMERIČKI) TESTOVI – primenjuju se kada:

- su vrednosti ispitivanog obeležja data numerički, odnosno intervalnom i skalom odnosa
- podaci o pojavi koja se proučava ne odstupaju značajno od normalne raspodele, a za velike uzorke (veće od 30 jedinica) i bez obzira na njihov raspored

Od parametrijskih testova najpoznatiji i teorijski najrazrađeniji i praktično primenljivi su **z-test** i **t-test**. Tu spadaju i regresiona analiza, analiza varijanse itd.
Međutim, kao što se može uočiti iz tabele br. 1, jedan od osnovnih uslova za primenu z testa je da je varijansa osnovnog skupa poznata, a kod osnovnog skupa koji nema normalan raspored, uzorci moraju da budu veći od 30 jedinica. Na drugoj strani t test može da se primeni i kod malih i kod velikih uzoraka i još plus nije uslov da je varijansa osnovnog skupa poznata vrednost. U praksi, pri zaključivanju na osnovu uzorka, varijansa osnovnog skupa nikad nije poznata, pa je praktična primena t testa mnogo značajnija, tako da će u ovom praktikumu samo t test biti obrađen.

NEPARAMETRIJSKI TESTOVI – se primenjuju kada:

- Su vrednosti obeležja date opisno
- Ponekad za numerička obeležja, kada je mali uzorak (manji od 30 jedinica), a podaci su nesimetrično raspoređeni (odstupanje od normalne raspodele testira se posebnim testovima), kada se varijable ne tretiraju kao brojevi s kojima su moguće matematičke operacije, već kao rangirani niz.

Od neparametrijskih testova poznati su: χ^2 test, Mek Nemarin test, Fišerov test itd.

Neparametrijski testovi imaju manju pouzdanost i manju snagu od parametrijskih.

Razlika između parametrijskih i neparametrijskih testova može da se iskaže i na sledeći način:

Parametrijski testovi se bave testiranjem parametara osnovnog skupa, a neparametrijski testiranjem vrednostima neparametara (frekvencija) osnovnog skupa.

Da bi smo lakše prepoznali test koji treba primeniti, potrebno je znati:

1. veličinu uzorka - *mali uzorci* su uzorci manji od 30 (n<30 ili n1+n2<60), a *veliki uzorci* veći od 30 jedinica (n>30 ili n1+n2>60).
2. vrstu rasporeda kod numeričkih obeležja - ukoliko je uzorak veliki koristi se parametrijski test. Ukoliko je uzorak mali treba proveriti tip rasporeda, pa ako je raspored normalan koristi se t test, a ukoliko nije normalan primenjuje se neki od neparametrijskih testova.

Za orjentaciju izbora testa pri testiranju nulte hipoteze o aritmetičkoj sredini osnovnog skupa može da posluži sledeća šema:

Tabela 1. Izbor testa pri testiranju nulte hipoteze

Osnovni skup	Veličina uzorka	Varijansa poznata	Varijansa nepoznata
Normalno raspoređen	n≥30	Z – test	t-test ili z–test
	n<30	Z - test	t - test
Nema normalan raspored	n≥30	Z - test	t - test
	n<30	Neparametrijski testovi	

Da bi smo se orjentisali koji tip parametrijskog ili neparametrijskog testa treba primeniti, potrebno je razlikovati zavisne od nezavisnih, tj. uparene od neuparenih uzoraka.
Nezavisni uzorci su oni uzorci kod kojih jedinice jednog uzorka nemaju nikakav uticaj na izbor jedinica drugog uzorka i metodom slučajnosti se odabiraju iz dva osnovna skupa sa različitim aritmetičkim sredinama. Kod nezavisnih uzoraka jedna grupa ispitanika je kontrolna (npr. grupa hipertoničara iz Niša lečenih novim lekom) a druga eksperimentalna (npr. grupa hipertoničara iz Beograda lečenih placebom). *Zavisni uzorci* su uzorci gde je jedna ista grupa ispitanika i kontrolna i eksperimentalna (npr. merena je jačina stiska šake grupe muškaraca – kontrolna grupa, a zatim su poslati na treniranje odbojke godinu dana nakon čega im je merena jačina stiska šake –nakon čega su oni eksperimentalna grupa.)

Studentov t - raspored

T distribuciju je otkrio 1908. Vilijam Goset, hemičar i statističar zaposlen u Ginisovoj pivskoj kompaniji. Smatrao se studentom koji još uvek uči statistiku, tako da je svoje radove potpisivao pod pseudonimom „student“, ili je možda koristio pseudonim zbog Ginisovih ograničenja o tajnama zanata.
Aritmetičke sredine jednakih uzoraka dobijenih iz istog osnovnog skupa raspoređuju se oko aritmetičke sredine osnovnog skupa u vidu normalne distribucije frekvencije. Raspored nije uslovljen unutrašnjim rasporedom jedinica osnovnog skupa, mada se mogu izdvojiti nekoliko oblika normalnosti:

1. Ako je raspored jedinica unutar osnovnog skupa normalan, raspored aritmetičkih sredina uzoraka oko aritmetičke sredine osnovnog skupa ima sve osobine Gausovog normalnog rasporeda, bez obzira na veličinu uzorka.
2. Iste osobine pokazuje i distribucija aritmetičkih sredina jednakih uzoraka dobijenih iz osnovnog skupa koji ne pokazuje normalan raspored, ako su uzorci veliki. Obično ovaj uslov zadovoljavaju uzorci veći od 30 jedinica (n>30).
3. Aritmetičke sredine jednakih malih uzoraka (n<30) dobijenih iz skupa koji ne pokazuje normalan raspored, takođe čine jedan oblik normalnog rasporeda oko aritmetičke sredine osnovnog skupa, poznat pod nazivom **Student-ov t raspored**.

Kod malih uzoraka simbol z-vrednost se zamenjuje simbolom t-vrednost. Njegova formula je:

$$t = \frac{\bar{x}_{uz} - \bar{x}_{os}}{SG} = \frac{\bar{x}_{uz} - \bar{x}_{os}}{\frac{SD_{uz}}{\sqrt{n-1}}}$$

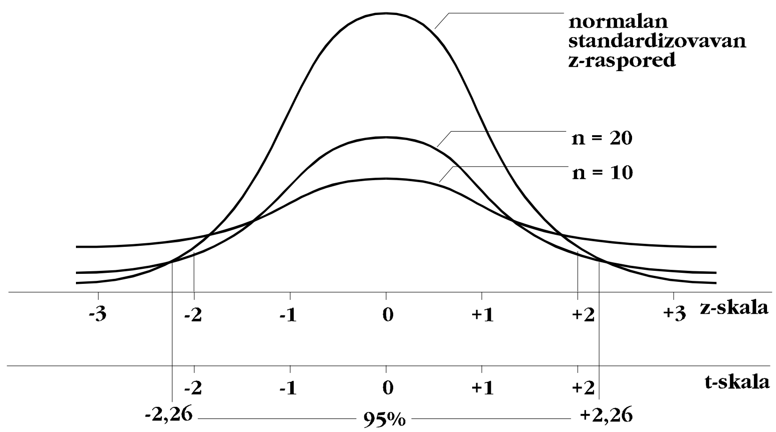
Suštinska razlika između t-vrednosti i z-vrednosti je u tome što u izrazu:

$$t = \frac{\bar{x}_{uz} - \bar{x}_{os}}{\frac{SD_{uz}}{\sqrt{n-1}}}$$

slučajno promenljive predstavljaju i $\bar{x}_{uz} - \bar{x}_{os}$ i $\frac{SD_{uz}}{\sqrt{n-1}}$.

U suštini standardna greška predstavlja ovde slučajno promenljivu jer zavisi od veličine uzorka, samim tim t-vrednost zavisi od veličine uzorka, dok to nije bio slučaj sa z-vrednošću jer je standardna greška bila konstantna pa z-vrednost nije zavisila od veličine uzorka.

Aritmetičke sredine jednakih uzoraka, ali manjih od 30 jedinica raspoređuju se oko aritmetičke sredine osnovnog skupa, po posebnom t-rasporedu, tako da t-vrednost zavisi od veličine uzorka. Postoje različite vrste t distribucija, to je kllasa distribucija. Kada se govori o određenoj t distribuciji, mora se odrediti stepen slobode. Krivine gustine t su simetrične i u obliku zvona kao normalna distribucija, a vrh im je na 0. Ipak, širenje je veće nego kod standardne normalne distribucije. Što su veći stepenislobode, to je t raspored bliži normalnom rasporedu.



Za male uzorke za ocenu intervala ne mogu da se primene tablice normalnog rasporeda već posebne Student-ove tablice t-rasporeda.
Širina krive, koja je pri dnu šira od krive normalnog rasporeda, uslovljava da se na nju ne može da primeni sigma tri pravilo, tj. da se u intervalu t= 0 ± 1,96 SG nalazi 95% t-vrednosti.
Ovaj procenat t-vrednosti se nalazi u znatno širem intervalu, a veličina intervala zavisi i od veličine uzorka: što je uzorak manji to je širina intervala veća.

- za uzorak od 2 jedinice u intervalu: $\bar{x}_{os} \pm 12{,}71 SG$
- za uzorak od 4 jedinice u intervalu: $\bar{x}_{os} \pm 3{,}18 SG$
- za uzorak od 10 jedinica $\bar{x}_{os} \pm 2{,}26 SG$
- za uzorak od 20 jedinica $\bar{x}_{os} \pm 2{,}09 SG$
- za uzorak od 30 jedinica $\bar{x}_{os} \pm 2{,}04 SG$

Na osnovu iznetih osobina Studentovog t-rasporeda urađene su posebne tablice za granične t-vrednosti za odgovarajuću verovatnoću i broja stepena slobode

Kod uzoraka većih od 30 jedinica i za odgovarajuću verovatnoću t-vrednost se izjednačuje sa z-vrednošću a Studentov t-raspored zadobija oblik standardizovanog normalnog rasporeda.

Iz ovog zaključka proizilazi drugi veoma važan zaključak da tablice Studentovog t-rasporeda mogu da se primenjuju za ocenu intervala pouzdanosti aritmetičke sredine osnovnog skupa i na osnovu malih uzoraka (n<30 jedinica) i na osnovu velikih uzoraka (n>30).

Razlike između Z i t – vrednosti

- Udaljenost slučajno promenljive od $\overline{X}$ izražena u jedinicama SD
- Ako je $\overline{X}$=0 a SD=1, Z se raspoređuje oko Z=0 u vidu standardizovanog normalnog rasporeda
- U intervalu Z=0 ± 3SD nalazi se 99,73% *svih* Z vrednosti
- Za odgovarajuću P, Z se određuje na osnovu tablica normalnog rasporeda.
- Z=1,64 za P=0,90
  Z=1,96 za P=0.95
  Z=2,58 za P=0.99
- Testovi zasnovani na Z vrednosti primenjuju se samo za velike uzorke n>30.

- Odnos između razlike i SG te razlike
- t-vrednost se oko t=0 ne raspoređuje u vidu standardizovanog normalnog rasporeda, već u vidu specifičnog t-rasporeda
- U intervalu t=0 ± 3SD nalazi se *manje* od 99,73% t-vrednosti.
- Za odgovarajuću P i SS t se određuje na osnovu studentovih tablica t-rasporeda.
- Sa povećanjem n, t se približava Z i sa n>30 izjednačava se sa Z.
- Testovi zasnovani na t-vrednosti mogu se primenjivati i na male i na velike uzorke.