Dr Đurđica Vasilj,
Poljoprivredni fakultet — Zagreb

PRIMJENA NEKIH NEPARAMETRIJSKIH METODA U STATISTIČKOJ OBRADI EKSPERIMENTALNIH PODATAKA

U V O D

Kod nas se u eksperimentalnom radu najviše koriste parametrijske metode, dok su neparametrijske vrlo malo poznate i gotovo se ne upotrebljavaju.

Kod prvih (koje uključuju analizu varijance, »F« test, »t« test, korelaciju, kovarijancu i dr.) zapravo se uvijek ide od pretpostavke da se radi o eksperimentalnim podacima koji se mogu prikazati krivuljom više ili manje sličnoj normalnoj. Pri tome se procjenjuju različiti parametri, a onda se na temelju njih donose zaključci o varijancama, prosjecima i sl.

Zbog toga se te metode obično zovu parametrijske statističke metode.

Međutim, ima mnogo slučajeva kad populacije nisu predstavljene distribucijom koja slijedi normalnu. Vrlo često nije ni moguće odrediti formu distribucije populacije odnosno uzroka.

Upravo za to je okrisno upoznati metode koje se moraju primijeniti u takvim slučajevima kad se radi o uzorcima uzetim iz populacija koje odstupaju od normalne raspodjele varijanata.

Najčešće se te metode zovu neparametrijske statističke metode ili metode neovisne o raspodjeli. Pomoću njih se vrše komparacije između distribucija, a ne između parametara kao kod parametrijskih metoda.

U našim biološkim istraživanjima neobično često imamo uzorke koji nikako ne pripadaju populacijama s normalnom raspodjelom, dakle slučajeve gdje se ne smiju upotrebljavati parametrijske metode. Zbog toga je važno upoznati neparametrijske metode i pravilno ih primijeniti.

Počeci neparametrijskih metoda datiraju još iz 19. stoljeća a najpoznatiji test, koji je najviše upotrebljavana neparametrijska metoda je h^2 test kojeg je još 1900. god. uveo Karl Pearson i primijenio kod usporedbe teoretskih i eksperimentalnih (opaženih) podataka, a kasnije proširio i na problem dvaju uzoraka. Kako je taj test u širokoj upotrebi, to se u ovom prikazu nećemo na nj osvrnuti.

Međutim, koliko se danas zna, pravi početak neparametrijskih testova bio je rad Hotellinga i Pabsta 1936. god.

Kako nema neke posebne literature o neparametrijskoj statistici, nema ni specijalne definicije o tome.

Osim toga, često nije ni lako uočiti problem i dati postupke i metode koji bi bili zaista neparametrijski.

Svrha ovog izlaganja bila bi, upoznati se s osnovnom teorijom o tome i ilustrirati je upotrebom nekoliko neparametrijskih testova.

Za praktičara je gotovo uvijek poželjno učiniti statističke opracije jednostavnima i svestrano upotrebljivim. Međutim, to ima izvjesnih nedostataka, pogotovo kad se neparametrijske metode primijene na one slučajeve gdje se zapravo, trebaju koristiti parametrijske.

Nameće se pitanje: mogu li se donositi zaključci direktno iz podataka, bez formuliranja postavki o formi distribucije?

Odgovor je »da«, iako treba posebno naglasiti da takvi zaključci imaju mnogo nedostataka i manje su vjerodostojni nego oni bazirani na specificiranim distribucijama.

Što dakle preostaje istraživaču kad je suočen s populacijom čija distribucija nije normalna? On može nastaviti dalje tretirati podatke kao da pripadaju populaciji raspoređenoj prema normalnoj rspodjeli, ali test koji primjenjuje smatrati samo aproksimativnim.

Umnogim slučajevima aproksimacija može biti dobra, pogotovo kad se radi s vrijednostima dobivenim na temelju velikih uzoraka. U suprotnom, zaključci doneseni na temelju parametrijskih metoda potpuno su krivi. Primjenom parametrijskih metoda operira se s prosjecima i standardnim devijacijama kao mjerilima centralne tendencije i disperzije. Kod neparametrijskih nailazimo na druga mjerila. Tako je npr. središnja tendencija izražena medijanom, često se upotrebljavaju percentili, quartil itd.

Za testiranje hipoteze o razlici između dva uzorka služe različiti testovi. Uglavnom mogu se podijeliti u 3 skupine prema tome da li su bazirani na:

1) predznacima diferenci
2) rangovima (redosljedu)
3) metodi randomizacije.

Naravno da ti neparametrijski testovi imaju svojih prednosti i mana.

Vrlo su praktični jer mogu koristiti rangove ili predznake diferenci.

Relativno su brzi i jednostavni za upotrebu. Mogu smanjiti posao oko skupljanja podataka. Podatke možemo rangirati (čak ako imamo i podatke koji slijede normalnu distribuciju, možemo koristiti rangove ako planiramo sakupiti mnogo podataka).

Ako raspolažemo podacima u toku više godina ili s više lokacija koji mogu sadržati različite varijance, upotrebom rangova oni mogu biti analizirani neparametrijskim metodma. Najveća slabost i nedostatak neparametrijlskih metoda je u slučaju neistinite nulhipoteze kad je problem pronaći razlike između prosjeka. (Nulhipoteza je postavka da razlike između vrijednosti koje ispitujemo nema.) Ako pokus rezultira u takvim podacima da je nulhipoteza istinita, tada su neparametrijske metode dobre kao i bilo koje druge.

Neparametrijske se metode osim toga ne primjenjuju u slučaju 1 slobodne varijante.

MEDIJANA I MOD

Jedno od mjerila središnje tendencije je medijana (Med), vrijednost koja se nalazi točno u sredini kad se podaci (varijante) poredaju po veličini. To je ujedno vrijednost od koje je 50% podataka većih, a 50% manjih.

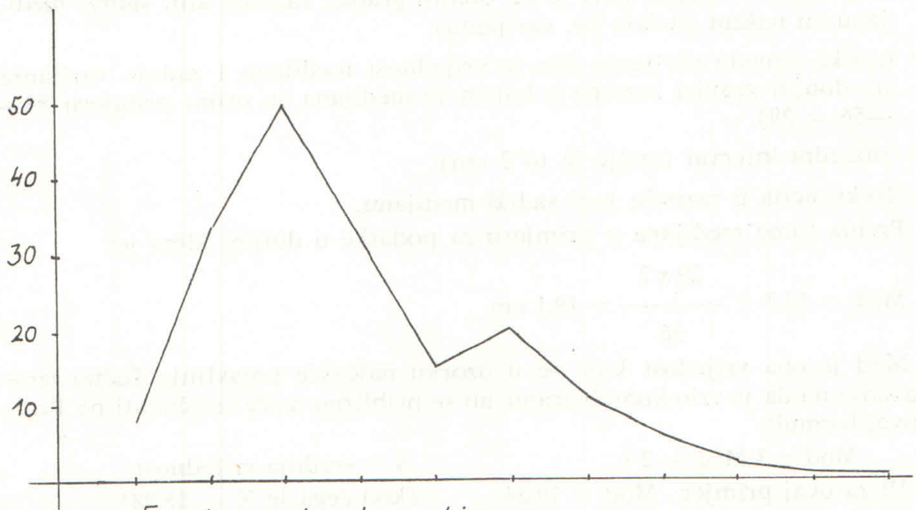
Prema tome, želi li se iz nekih podataka mjerenja izračunati medijana, najprije treba podatke srediti rastućim redoslijedom. Tada se medijana dobije kao vrijednost srednje varijante ili općenito vrijednost $\frac{n + 1}{2}$ varijante.

Slično kao i srednja vrijednost (prosjek) i medijana je mjerilo sredine distribucije. Lako je uočiti da se u slučaju simetrične raspodjele varijanata te dvije vrijednosti (medijana i srednja vrijednost) poklapaju.

Medijana bi se često morala upotrebiti u slučajevima gdje mnogo bolje ilustrira prosjek nego srednja vrijednost. Npr.: u skupini brojeva 1, 3, 4, 5, 8 medijana je 4 (vrijednost srednjeg podatka u nizu podataka svrstanih rastućim redoslijedom), a prosjek je 4,2 ($\frac{21}{5}$ = 4,2). Ako se vrijednosti uzorka nešto promijene pa postoji recimo 1, 3, 4, 5, 24 medijana je 4, a srednja vrijednost 7,4. Četiri podatka su ispod a tek jedan iznad prosjeka. U ovom slučaju srednja vrijednost nije dobar pokazatelj sredine distribucije uzorka (ali treba naglasiti i to, da se s ovakvim uzorkom srednja vriejdnost uopće ne smije niti računati).

Iz ovog se može jasno vidjeti primjena srednje vrijednosti odnosno medijane. U izvjesnim slučajevima medijana bolje ilustrira sredinu raspodjele nego srednja vrijednost i obratno.

Još jedno od mjerila centralne tendencije je i mod (po Pearsonu). To je vrijednost varijante koja se najčešće (s najvećom frekvencijom) pojavljuje u distribuciji. Ilustrirajmo to s primjerom gdje su frekvencije izvjesnih mjerenja (8, 33, 50, 32, 15, 20, 10, 6, 2, 1, 1) prikazane pomoću frekvencionog poligona:



Frekvencioni poligon

Ovdje se frekvencija najprije povećava dok ne dosegne vrh ,pa onda opet pada i ima drugi vrh. U ovom slučaju radi se dakle o bimodalnoj raspodjeli. Dakle, raspodjeli koja nije normalna. Glavna primjena medijane je kod ovakvih raspodjela koje ne slijede normalnu distribuciju. U ovom primjeru (s ukupno 178 varijanata) medijana je vrijednost 89. varijante.

Mod kao mjerilo središnje tendencije, karakterizira skupinu individua. U izvjesnim slučajevima on ima prednost u dnosu na medijanu i srednju vrijednost. Dakle, ovisi o konkretnom slučaju, istraživač će se odlučiti za onu vrijednost (srednja vrijednost, medijana ili mod) koja će mu najbolje predstavljati sredinu distribucije s kojom radi.

Pokažimo na primjeru računanje medijane i moda iz podataka opažanja o duljini klipa jednog linijskog hibrida kukuruza sređenih u varijacioni red:

Razredi za dulj. klipa-cm	11,5 — 13,5	— 15,5	— 17,5	— 19,5	— 21,5	— 23,5	— 25,5	— 27,5	— 29,5
Frekvencije	5	17	34	36	43	27	3	2	2
Kumulativne frekv.	5	22	56	92	135	162	165	167	169

Kako je n = 169 medijana ovog uzorka je vrijednost 85. varijante, koja se nalazi između 56. i 92. u rubrici kumulativnih frekvencija. To znači da se izračuna medijana, potrebno je vrijednosti 56. varijante pribrojiti još vrijednost 29 varijanata (56 + 29 = 85) ili:

$$\text{Med} = x + \frac{g \cdot a}{f}$$

x = vrijednost varijante koja je na donjoj granici razreda koji sadrži medijanu (u našem slučaju 56. varijanta).

g = razlika između varijante čija je vrijednost medijana i zadnje varijante na donjoj granici razreda u kojem je medijana (u ovom primjeru 85—56 = 29).

a = razredni interval (ovdje je to 2 cm).

f = frekvencija u razredu koji sadrži medijanu.

Prema tome medijana u primjeru za podatke o duljini klipa je:

$$\text{Med} = 17{,}5 + \frac{29 \times 2}{36} = 19{,}1 \text{ cm}$$

Mod je ona vrijedost koja se u uzorku najčešće pojavljuje Točno izračunavanje moda je vrlo komplicirano, ali se približno može izračunati po Pearsonovoj formuli:

$\text{Mod} = 3\,\text{Med} - 2\,\bar{x}$ $\qquad$ $\bar{x}$ = srednja vrijednost

Ili za ovaj primjer Mod = 19,34 (kod čega je X = 18,98)

Moodov test medijane

Testira razliku između dva uzorka .Uzorci mogu biti sasvim neovisni i različitih veličina (recimo n i m). Postupak je slijedeći:

1) poredati podatke iz obiju uzoraka zajedno, po veličini od najmanjeg do najvećeg.
2) naći medijanu.
3) za svaki uzorak naći vrijednosti veće od medijane.
4) testirati opravdanost sa h^2 i 1 slobodnom varijantom.

Primjer: Iz podataka o težini zrna na klipu kod dviju linija kukuruza (A i B) upotrebom Mood-ovog testa utvrditi će se da li postoji opravdana razlika u težini zrna na klipu između linije A i linije B.

Uzorak A Sample A	Uzorak B Sample B	Zajednički uzorak (A+B) i vrijednosti poredane po veličini— Ranked values of A and B together
47,9	57,6	47,7
52,3	58,7	47,9
50,7	53,2	48,1
56,4	57,0	50,7
48,1	54,3	50,9
50,9	53,7	52,3
47,7	$m = 6$	53,2
$n = 7$	$m_1 = 5$	53,7
$n_1 = 1$		54,3
		56,4
		57,0
		57,6
		58,7

Med = 53,2 a n_1 i m_1 je broj podataka koji su veći od medijane u svakom od uzoraka.

Kriterij za provođenje ovog testa je h^2, a računa se kao:

$$h^2 = \frac{[m_1\ (n-n_1) - n_1\ (m - m_1)]^2\ (m+n)}{m\ n\ (m_1+n_1)\ (m+n-m_1-n_1)} = 6{,}19$$

Iz h^2 tabela (tabela 4) za 1 slobodnu varijantu i $P_{0,05}$ očitano tablični h^2 koji iznosi 3,84. To znači statističku opravdanost, odnosno ukazuje na to da ova 2 uzorka pripadaju dvjema različitim populacijama.

Drugim riječima razlika u težini zrna na klipu je opravdana. Ili, linija B ima veću težinu zrna na klipu od linije A s kojom se uspoređuje.

Test predznaka

Ovo je možda jedan od najjednostavnijih testova. Vrlo je pogodan za podatke koji dolaze u parovima. Jedina postavka na kojoj je temeljen je da su različiti parovi potpuno neovisni (znači da imamo n potpuno neovisnih parova). Tu zapravo promatramo smjer mijenjanja vrijednosti podataka za svaki par pod dvjema različitim okolnostima (razni tipovi tla, gnojidbe, temperature, godine, oborine itd.). Pri tome operiramo samo s predznacima diferenci i zapravo uspoređujemo parove (isti uvjeti za svaki par, a različiti za različite parove). Zbog toga je neobično pogodan za sve one slučajeve gdje imamo uzastopna mjerenja (prije-poslije tretmana) na istom subjektu, te testiranje razlika između dvaju članova iz više pokusa (lokacija) i sl.

Kao kriterij kod ovog testa služi zapravo učestalost pojavljivanja »+« odnosno »—« znakova diference između parova. Ako nema razlike između populacija koje se testiraju, onda je broj »+« jednak broju »—« znakova diferenci. (Diference jednake nuli se isključuju). Tab. 1. daje granične vrijednosti za r tj. predznak koji se pojavljuje manji broj puta.

Pokažimo na primjeru i način izvođenja ovog testa: Kao osnovni podaci poslužit će rezultati farinograma o vrijednostima o upijanju vode kod dviju sorata pšenice (Bezostaja i Ranaja) sijanih u makropokusima na 12 lokacija. Na temelju podataka tih 12 parova vrijednosti (za obe sorte) želi se testom predznaka utvrditi da li postoji razlika u upijanju vode između ovih dviju sorata.

Podaci su:

Lokacija Location		Bezostaja	Ranaja	Predznak diference Sign of difference
1.	Beli Manastir	58,5	59,0	—
2.	Čakovec	62,8	63,0	—
3.	Đakovo	61,8	58,3	+
4.	Našice	58,5	61,0	—
5.	Osijek	58,5	63,5	—
6.	Podrav. Slatina	62,0	54,5	+
7.	Poreč	58,5	60,5	—
8.	Slavonski Brod	57,3	57,8	—
9.	Sl. Orahovica	57,0	58,0	—
10.	Umag	57,0	57,5	—
11.	Valpovo	57,8	60,2	—
12.	Virovitica	57,1	62,0	—

To znači da se kod N = 12 parova »+« znakovi pojavljuju 2, a »—« znakovi 10 puta. Dakle r = 2.

Tabela 1

Granične vrijednosti »r« za test predznaka
Critical values of »r« for the sign test

N	1%	5%	10%	25%	N	1%	5%	10%	25%
1					26	6	7	8	9
2					27	6	7	8	10
3				0	28	6	8	9	10
4				0	29	7	8	9	10
5			0	0	30	7	9	10	11
6		0	0	1	31	7	9	10	11
7		0	0	1	32	8	9	10	12
8	0	0	1	1	33	8	10	11	12
9	0	1	1	2	34	9	10	11	13
10	0	1	1	2	35	9	11	12	13
11	0	1	2	3	36	9	11	12	14
12	0	2	2	3	37	10	12	13	14
13	1	2	3	3	38	10	12	13	14
14	1	2	3	4	39	11	12	13	15
15	2	3	3	4	40	11	13	14	15
16	2	3	4	5	41	11	13	14	16
17	2	4	4	5	42	12	14	15	16
18	3	4	5	6	43	12	14	15	17
19	3	4	5	6	44	13	15	16	18
20	3	5	5	6	45	13	15	16	18
21	4	5	6	7	46	13	15	16	18
22	4	5	6	7	47	14	16	17	19
23	4	6	7	8	48	14	16	17	19
24	5	6	7	8	49	15	17	18	19
25	5	7	7	9	50	15	17	18	20

U tabeli 1. za 12 parova i uz P = 0,05 kritična vrijednost za r je 2.

Vrijednost r-a koja je jednaka tabličnoj ili manja od nje je signifikantna i obratno. Prema tome u gornjem primjeru postoji opravdana razlika u upijanju vode kod brašna sorata Bezostaja i Ranaja sijanih na 12 lokacija.

Iz primjera je vidljiva jednostavnost i brzina provođenja ovog testa, koji postaje sve točniji što se više povećava broj parova. To je ujedno i njegova najveća prednost koja omogućuje i sakupljanje velikog broja podataka.

Jedinstvena primjena ovog testa je u slučajevima kad se raspolaže s podacima dobivenim bonitiranjem, podacima o zaraženosti nekom bolešću ili štetnikom, podacima u vidu rangova (pri čemu je brojčana vrijednost podatka zamijenjena rangom) i uopće podacima o kvalitativnim svojstvima koja ne mogu biti mjerena nego samo uspoređivana (kao što su okus, kvaliteta i sl.).

Nedostatak ovog testa je da ne vodi računa o veličinama diferenci. Međutim, kompariramo li ga sa »t« testom, efikasnost testa predznaka u odnosu na parametrijski »t« test je oko 67%. To znači da test predznaka zahtijeva 100 parova podataka za dobivanje istovrijednog zaključka kao upotrebom parametrijskog »t« testa sa 67 parova podataka.

Wilcoxonov T-test (rangiranje razlika između mjerenja).

To je zapravo modifikacija testa predznaka, s time što je ovo nešto točniji test jer vodi računa i o veličini diferenci između parova. Pri tome se koriste relativne veličine podataka a ne direktno podaci koji se ispituju. Teoretskim ispitivanjima je dokazano, da je ovaj test tek neznatno manje pouzdan od najvažnijeg parametrijskog testa (t-testa). Za to čak u slučajevima gdje se inače upotrebljava »t« test Wilcoxonov T test sasvim zadovoljava. Želeći dobiti brzu informaciju o opravdanosti razlike između nezavisnih parova, svakako će nam ovaj test biti od velike koristi.

Postupak kod provođenja je:

1) poredati razlike između parova od najmanje do najveće, označavajući najmanju rangom 1 i tako redom.
2) označiti svaki rang predznakom kakav je imala stvarna odgovarajuća diferenca.
3) izračunati sumu pozitivnih i negativnih predznaka.
4) komparirati tu sumu tabličnom vrijednošću (tab. 2).

P r i m j e r (podaci istraživanja prof. Tavčara i dr Kenđelić): Kod 10 parova biljčica kukuruza po jedna je biljka bila podvrgnuta visokoj dozi zračenja gama γ^2 zrakama, te nakon izvjesnog vremena ispitivana razlika u du-

Tabela 2

Granične vrijednosti za Wilcoxon-ov test
Critical values for Wilcoxon,s test

Parovi Pairs n	P .05	.02	.01	Parovi Pairs n	P .05	.02	.01
6	0	—	—	16	30	24	20
7	2	0	—	17	35	28	23
8	4	2	0	18	40	33	28
9	6	3	2	19	46	38	32
10	8	5	3	20	52	43	38
11	11	7	5	21	59	49	43
12	14	10	7	22	66	56	49
13	17	13	10	23	73	62	55
14	21	16	13	24	81	69	61
15	25	20	16	25	89	77	68

ljini korijena između tretirane i netretirane biljke svakog para. Iz podataka o razlikama u porastu između tretiranih i netretiranih biljaka, moguće je primjenom Wilcoxonovog testa utvrditi opravdanost razlike između porasta tretiranih i netretiranih biljaka.

Par Pair	Duljina korijena u mm Root lenght in mm tretirano treated	 netretirano non treated	Razlika Difference	Rang Rank
1	8,0	3,2	4,8	7
2	6,5	2,5	4,0	6
3	2,5	8,2	—5,7	—8
4	6,3	2,6	3,7	5
5	2,0	7,8	—5,8	—9
6	8,5	2,0	6,5	10
7	1,7	4,5	—2,8	—4
8	4,5	2,0	2,5	3
9	4,5	3,0	1,5	1
10	3,1	5,0	—1,9	—2

Suma negativnih rangova je 23, a onih pozitivnih 32. Kriterij kod ovog testa je manja suma (u našem slučaju suma negativnih rangova 23). Prema tabeli 2 za 10 parova, granična (kritična) suma rangova je 8 za P=5% odnosno 3 za P=1%. Budući je ona manja od one u primjeru, to znači da razlika između ovih dviju skupina nije opravdana. Dakle, podvrgavanje biljaka određenoj doza γ zračenja nije utjecalo na razliku u duljini korijena.

Ovdje treba naglasiti da je za razliku od drugih testova (t, F, h^2 i sl.) opravdana ona vrijednost koja je manja od kritične (granične), a ne obratno kao obično.

Još je važno uočiti kako se postupa pri rangiranju diferenci koje imaju istu vrijednost. Recimo da su registrirane dvije diference čija je vrijednost bila 10,2 a po redosljedu rangova one bi trebale doći na 7 mjesto. Pri tome se postupi na taj način da obje dobiju rang 7 1/2 s time da slijedeća veća diferenca (koja bi trebala dobiti rang 8) bude označena kao 9.

Tabela 2 vrijedi za testiranje diference između manje od 16 parova. Za više od 16 parova može se upotrebiti Z, kao

$$Z = \frac{\left|\mu - T\right| - \frac{1}{2}}{\delta}$$

Gdje je T manja suma rangova, $\mu = \dfrac{n\,(n+1)}{4}$, a $\delta = \sqrt{\dfrac{(2n+1)\,\mu}{6}}$

Kao i obično ako je $Z > 1{,}96$ razlika je opravdana uz $P = 5\%$.

Brzina, lakoća izvođenja i relativno visoka točnost (obzirom da se radi o neparametrijskom testu) glavne su prednosti ovog testa.

Mann-Whitney -ev test (ili Wilcoxonov test za dva uzorka).

Wilcoxon je prije opisani test proširio na dva uzorka za koje nisu nužni parovi (dakle potpuno neovisni uzorci i podaci u njima), a Mann i Whitney su ga usavršili za uzorke koji imaju različiti broj mjerenja (ili $n_1 \neq n_2$, odnosno $n_1 \leq n_2$).

Postupak je slijedeći:

1) Poredati (rangirati) podatke po veličini od najmanjeg do najvećeg za oba uzorka zajedno.
2) Pridati rangove manjem uzorku (čiju sumu rangova označiti sa T)
3) Izračunati $T' = n_1 (n_1 + n_2 + 1) - T$
4) Usporediti manju sumu rangova s tabličnim vrijednostima.

Iskoristimo podatke o koeficijentu probavljivosti suhe tvari kod ovaca (skupina A) i junadi (skupina B) hranjenih kukuruznom silažom. Podaci za oba uzorka (skupina A i B) zajedno poredani su po veličini i pridani su im rangovi:

Uzorak (A+B) Sample (A+B)	53,2(A)	53,6(A)	54,4(A)	56,2(A)	56,4(A)	57,8(A)	58,7(B)
Rangovi Ranks	1	2	3	4	5	6	7
	59,2(B)	59,8(B)	69,1(A)	62,5(B)	63,1(B)	64,2(B)	
	8	9	10	11	12	13	

gdje je $n_2 = 7$ (A) i $n_1 = 6$ (B)

Suma rangova za A (veći uzorak) je $1+2+3+4+5+6+10 = 31$

za B (manji uzorak) je $7+8+9+11+12+13 = 60$ (T)

$T' = n_1 (n_1 + n_2 + 1) - T = 6 (6 + 7 + 1) - 60 = 24$

$T' = 24$

Iz tabele 3 tablična vrijednost za $n_1 = 6$ i $n_2 = 7$ uz $P = 5\%$ je 27, a to znači da je razlika u probavljivosti suhe tvari (u kukuruznoj silaži) između ovaca i junadi opravdana. Dakle opravdano veći koeficijent probavljivosti je kod junadi.

Kruskal Wallisov H test

Ovo je test za kompletno slučajni raspored bilo kojeg broja tretiranja, kojim se testiraju razlike između tretiranja.

Postupak:

1) poredati podatke svih tretiranja zajedno, od najmanjeg do najvećeg (rangirati).
2) sumirati rangove za svako tretiranje.
3) izračunati H vrijednost i usporediti s tabličnom vrijednosti.

Tabela 3 Granične sume rangova
Critical values of rank sum

n_2 = veći n = larger n	P	n_1 = manji n = smeller n: 2	3	4	5	6	7	8	9	10	11	12
4	.05			10								
	.01			—								
5	.05		6	11	17							
	.01		—	—	15							
6	.05		7	12	18	26						
	.01		—	10	16	23						
7	.05		7	13	20	27	36					
	.01		—	10	17	24	32					
8	.05	3	8	14	21	29	38	49				
	.01	—	—	11	17	25	34	43				
9	.05	3	8	15	22	31	40	51	63			
	.01	—	6	11	18	26	35	45	56			
10	.05	3	9	15	23	32	42	53	65	78		
	.01	—	6	12	19	27	37	47	58	71		

Pri tom je

$$H = \frac{12}{n(n+1)} \sum \frac{R_i^2}{n_i} - 3(n+1)$$

n = broj podataka opažanja u i-tom tretiranju

i = 1, ... , k broj tretiranja

$n = \Sigma n_i$

R_i = suma rangova

12 i 3 su konstante i neovisne su o veličini eksperimenta

Kod usporedbe faktora H s tablični ma, mogu se koristiti h^2 tabele (jer je H vrlo slično raspoređen kao h^2 (za K—1 slobodnih varijanata. (Za K=2 bolje je upotrebiti Wilcoxonov test, a za veće vrijednosti K dolazi u obzir ovaj test).

Uzmimo za ilustraciju primjer sa 6 (A, B, C, D, E i F) tretiranja od kojih je svako imalo po 5 podataka.

Već sređeni podaci po veličini za sva tretiranja skupa (rangirani) su:

Podaci opažanja Observed values	9,1(C)	11,6(E)	11,8(E)	11,9(C)	14,2(E)
Rangovi — Ranks	1	2	3	4	5
	14,3(E)	14,4(E)	15,8(C)	16,9(F)	17,0(C)
	6	7	8	9	10
	17,3(F)	17,7(B)	18,6(D)	18,8(D)	19,1(F)
	11	12	13	14	15
	19,4(C)	19,4(A)	19,4(F)	20,5(D)	20,7(D)
	17	17	17	19	20
	20,8(F)	21,2(D)	24,3(B)	24,8(B)	25,2(B)
	21	22	23	24	25
	27,0(A)	27,9(B)	32,9(A)	32,6(A)	33,0(A)
	26	27	28	29	30

Sume rangova za svako od 6 tretiranja su:

R(A) = 130
R(B) = 111
R(C) = 40
R(D) = 88
R(E) = 23
R(F) = 73

$$H = \frac{12}{30 \times 31} \times \frac{130^2 + \dots\dots\dots\dots + 73^2}{5} - 3 \times 31 = 21.64$$

U h^2 tabelama (Tab. 4) za 5 slobodnih varijanata tablična vrijednost uz P = 5% je 11,1, a za P = 1% je 15,1.

Kako je dobivena vrijednost za H (21,64) veća, to znači da su opravdane razlike između tretiranja.

Test sume kvadranata

Olmstead i Tukey su razvili ovaj neparametrijski test kojim se testira pretpostavka da su dvije kontinuirane varijable neovisne (drugim riječima da nema korelacije među njima). Test je neobično praktičan u slučaju da su podaci sređeni u vidu dijagrama.

Postupak:

1) naći i ucrtati medijanu za svaku varijablu.
2) zatim sumirati podatke odozgo prema dolje (koristeći os y) sve dok ne bi trebalo preći vertikalnu medijanu. Označiti taj zbroj podataka predznakom odgovarajućeg kvadranta.
3) ponoviti taj postupak u smjerovima s desna na lijevo, odozdo prema gore i s lijeva na desno.
4) izračunati sumu kvadranata i usporediti s tabličnom vrijednosti.

Tabela 4 X^2 vrijednosti
Values of X^2

df	P		
	.500	.100	.010
1	.455	2.71	6.63
2	1.39	4.61	9.21
3	2.37	6.25	11.3
4	3.36	7.78	13.3
5	4.35	9.24	15.1
6	5.35	10.6	16.8
7	6.35	12.0	18.5
8	7.34	13.4	20.1
9	8.34	14.7	21.7
10	9.34	16.0	23.2

P r i m j e r: poslužit će primjer iz Steel i Torril-a (Principles and procedus in Statistics) o ispitivanju ovisnosti broja ovuliranih folicula i broja iznesenih jaja kod fazana. Dakle, radi se o utvrđivanju postojanja korelativne veze između te dvije varijante. Podaci su:

Broj iznesenih jaja	39,	29,	46,	28,	31,	25,	49,	57,	51,	21	42,	38,	34,	47
Broj ovuliranih folicula	37,	34,	52,	26,	32,	25,	55,	65,	44,	25,	45,	26,	29,	30

Predstave li se ovi podaci dijagramom i ucrtaju se medijane za svojstvo x i svojstvo y (koje iznose med x = 33, a med y = 38,5) dobijemo:

Sume po kvadrantima i ukupna suma su:

Smjer		Broj podataka
odozgo — dolje	— — — —	3
desno — lijevo	—. —. —. —	6
odozgo — gore	. . . .	3
lijevo — desno	xxxxxxxxxxxxx	5

Suma kvadranata = 17

Iz tabele 5 očita se tablična vrijednost (ili granična suma) uz P = 5% a ta je 11. Kako je suma kvadranata u gornjem primjeru 17, dakle veća od tablične, to znači opravdanost ili signifikantnost. Drugim riječima, postoji korelacija između broja ovuliranih folicula i broja snešenih jaja kod fazana.

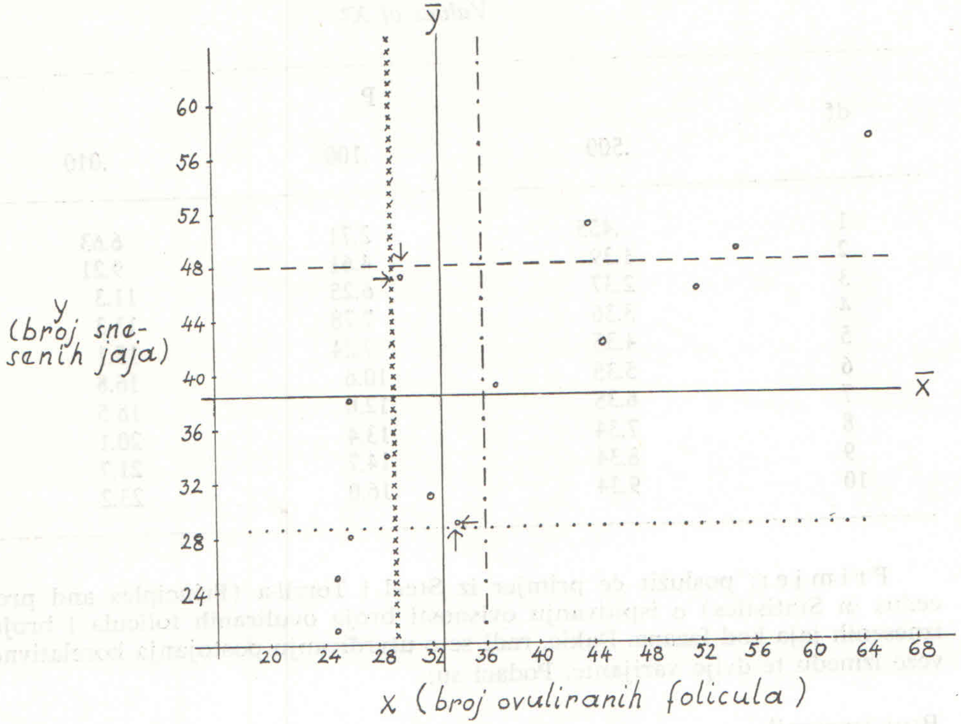


Tabela 5 Granice opravdanosti za sumu kvadranata
Significance levels for quadrant sums

Granica opravdanosti Significance level	Suma kvadranata Magnitude of quadrant sum
.10	9
.05	11
.02	13
.01	14 — 15
.005	15 — 17
.002	17 — 19
.001	18 — 21

Obzirom na jednostavnost izvođenja ovog testa i brzo dobivanje informacije o postojanju ili nepostojanju korelativne veze između dvije varijable, on je vrlo pogodan i u svakom od slučajeva kod kojih se inače mogu primijeniti i parametrijske metode.

Spearmanov koeficijent korelacije (r_s).

Ovaj se test primjenjuje za utvrđivanje postojanja korelativne veze između dvaju svojstava koja su izražena podacima u vidu rangova. Pri tome podaci mogu biti ili sakupljeni u obliku rangova ili pak u nekoj drugoj skali, a onda kasnije rangirati.

Postupak računanja Spearmanovog koeficijenta korelacije je:

1) Poredati po veličini podatke za svaku varijablu (rangirati) ukoliko podaci nisu sakupljeni u vidu rangova.
2) Naći razlike u rangovima za svaki par.
Neka je d_i razlika za i-ti par.
3) Procijeniti korelacioni koeficijent pomoću Spearmanovog koeficijenta korelacije r_s

$$r_s = 1 - \frac{6 \Sigma d^2_i}{n (n-1) (n+1)}$$

4) Ako je broj podataka velik, procjena korelacionog koeficijenta može biti testirana pomoću $t = r_s \sqrt{\frac{n-2}{1-r^2_s}}$, čija se raspodjela poklapa s raspodjelom Studentovog »t« za n—2 slobodne varijante (pri čemu je n broj diferenci između parova).

Kao uostalom sve neparametrijske metode, tako se i ova neobično brzo i jednostavno provodi, a na primjeru će se najbolje vidjeti jedna od njezinih praktičnih upotreba.

Vrlo često je u istraživačkom radu vrlo važno utvrditi pouzdanost osoba (tehničara i sl.) koje sakupljaju eksperimentalne podatke. Povjerimo li izvjesna mjerenja dvojici tehničara, kako utvrditi da podaci neće biti opterećeni njihovom subjektivnošću? Pri tome nam može pomoći upravo računanje Spearmanovog koeficijenta korelacije. Evo jednog primjera iz osobnog iskustva.

P r i m j e r: u okviru istraživanja na kukuruzu (M. Kump, i Đ. Vasilj) opaženo je da jedna linija ima izrazito priklonjen list stabljici i da to svojstvo unosi u jednostruke hibride. Kako je prema novijim istraživanjima utvrđeno, hibrid koji ima list više priklonjen stabljici bolje koristi osvjetljenje u gustom sklopu, zbog toga je bilo interesantno ispitivati spomenutu liniju i unošenje tog svojstva (lista priklonjenog stabljici) u hibride. To je naravno zahtijevalo mnoštvo podataka o kutu između lista i stabljike (na koji smo način izrazili to svojstvo) Mjerenja su bila povjerena dvojici tehničara, a vršila su se pomoću za tu svrhu izrađenog kartonskog kutomjera, s pomičnom kazaljkom. Nakon prvog dana mjerenja, usporedivši vrijednost za kuteve koje su oba tehničara dobili mjerenjem istih biljaka i izračunavši korelacioni koeficijent između njihovih mjerenja, utvrdili smo da se slobodno možemo pouzdati u podatke sakupljene na taj način. Evo za što:

Redni broj mjerenja No. of measurement	Vrijednosti za kut između lista i stabljike (u stupnjevima) — Values for leaf angle Tehničar A Technician A	Tehničar B Technician B
1	20,0	27,5
2	34,5	42,5
3	32,0	32,5
4	30,0	32,5
5	32,5	45,0
6	20,0	27,0
7	27,5	35,0
8	25,0	32,0
9	22,5	35,5

Pripišemo li odgovarajuće rangove pojedinim mjerenjima svakog tehničara dobijemo.

Redni broj mjerenja No. of measurments	Rangovi za A Ranks of A	Rangovi za B Ranks of B	Razlika Difeference d	d^2
1.	1,5	2	—0,5	0,25
2	9	8	1	1,00
3.	7	4,5	2,5	6,25
4.	6	4,5	1,5	2,25
5.	8	9	—1	1,00
6.	1,5	1	0,5	0,25
7.	5	6	—1	1,00
8.	4	3	1	1,00
9.	3	7	—4	16,00
				d^2=29

Napomena: mjerenja označena rangom 1,5 ili 4,5 su sasvim ista, pa su ako se poredaju po veličini — rangiraju ,dobila i iste vrijednosti onim redosljedom kojim dolaze).

Pomoću formule za r_s utvrdimo da li postoji korelacija između mjerenja ove dvojice tehničara. Dakle:

$$r_s = \frac{6\ \Sigma\ d^2}{n\ (n-1))\ \ (n+1)} = 1 - \frac{6 \times 29}{9x8x10} = 0,75$$

Kako r_s = 1 označuje potpunu korelaciju, to r = 0,75 znači, da postoji vrlo jaka korelacija između vrijednosti koje su dobila ova dvojica tehničara.

To nam je bio indikator da se posao oko sakupljanja podataka o otklonu lista od stabljike može povjeriti ovoj dvojici tehničara,, jer su vrijednosti njihovih izmjera u jakoj korelaciji; to znači manjim vrijednostima jednog odgovaraju i manje vrijednosti drugog i obrnuto. Dakle, ovdje se vidi primjena ove brze i praktične metode koju svaki od nas može danomice upotrijebiti.

Naravno, da smo dobili r_s vrlo mali ili skoro nula, to bi bio znak da im se ne može povjeriti prikupljanje podataka na određenom materijalu, jer će se podaci razlikovati, a zaključci dobiveni na temelju takvih podataka bili bi krivi. Na isti bi način izračunali r_s iz podataka koji su već sakupljeni u vidu rangova (što je često slučaj kog bonitiranja, različitog procjenjivanja i sl.).

Na primjer, iz podataka dvaju ocjenjivača o općem utisku o kvalitetu ploda (ocjenjivanom ocjenama od 1 do 3) za isti uzorak jabuke sorte Delišez, primjenom ovog jednostavnog testa može se utvrditi koliko se cjene ovih ocjenjivača podudaraju.

Ocjenjivač br. 1 Judge 1	Ocjenjivač br. 2 Judge 2	Razlika Difference d	d^2
1	1	0	0
1	2	—1	1
1	1	0	0
1	1	0	0
2	1	+1	1
2	2	0	0
1	2	—1	1
2	1	+1	1

$$r_s = 1 - \frac{6\ \Sigma\ d^2_i}{N\,(N^2 - 1)} = 1 - \frac{6 \times 4}{8 \times 63} = 0{,}952$$

Kako r_s = 0,95 ukazuje na jaku korelaciju, to možemo zaključiti da se ocjene ovih ocjenjivača podudaraju, odnosno da su oba dala u većini slučajeva iste ocjene ispitivanom uzorku jabuka sorte Delišez.

ZAKLJUČAK

Ovdje su iznijete samo neke od niza neparametrijskih metoda i mogućnost njihove upotrebe. Svrha je bila ukazati na njihovo postojanje i način primjene.

Kako je već spomenuto, te metode imaju svojih prednosti i mana. Jedna od njihovih najvećih prednosti je brzina i lakoća kojom se izvode, zahtijevaju malo ili gotovo ništa računanja, a mogu se upotrebiti za velik broj slučajeva, jer ne zahtijevaju nikakve postavke (kao što je to slučaj s parametrijskim metodama). Za to će se uvijek moći upotrijebiti:

1) Kad se žele dobiti brzi ili preliminarni rezultati.
2) Kad su podaci sakupljeni u vidu relativnih (ili komparativnih) a ne apsolutnih vrijednosti, kao npr. veći — manji — najmanji, da — ne i sl.

3) Općenito, one se moraju upotrijebiti uvijek u slučajevima kad se nema uvida o načinu na koji je populacija distribuirana.

Neosjetljivost odnosno manja preciznost, jedna je od najvećih nedostataka neparametrijskih metoda. One naime najčešće koriste relativne vrijednosti, rangove i sl. pa zapravo zanemaruju apsolutne vrijednosti i njihove razlike, a to uzrokuje manju preciznost.

Ipak, znajući kad ih se može i treba koristiti, one mogu biti od neobične koristi. To je ujedno bila i svrha ovog izlaganja.

THE USE OF SOME NONPARAMETRIC METHODS IN THE ANALYSES OF EXPERIMENTAL DATA

by
Dr. Đurđica Vasilj

Summary

We usually assume that the populations we deal with, have some known form. The methods which deal with distributions without specifying the form of the distributions are called nonparametric methods.

Parametric methods, such as analysis of variance, regression and correlation, »F« and »t« tests etc., are well known. But distribution-free or nonparametric methods are not in the use as much as they coulde be.

A considerable amount of collected data is such that the underlying distribution is not easily specified. To handle such data, we need distribution bree statistics.

Therefore, the purpose of this paper is to acquaint some nonparametric methods, to give examples where these methods can be used, and to point out their advantages and disadvantages.

As these nonparametric techniques are very quick, they are very useful in preliminary data analyses.

In parametric methods certain assumptions have to be satisfied. Nonparametric methods do not need any assumption about distributions.

It is very important to know the assumptions, but also to have a grasp of the consequences of violations of assumptions.

LITERATURA:

W. J. Dixon and F. J. Massey: »Introduction to statistical analysis« Mc Graw-Hill Book Company, Inc. 1951.

R. D. Remington and M. A. Schork: »Statistics with applications to the biological and health Sciences« — Prentice — Hall, Inc. 1970.

G. W. Snedecor and W. G. Cochran: »Statistical methods« — The Iowa State University Press, Ames, Iowa, 1967.

R. G. D. Steel and J. H. Torrie: »Principles and procedures of statistics« Mc Graw—Hill Book Company, Inc. 1960.

A. Tavčar: »Biometrika u poljoprivredi«.