Uporaba neparametarske statistike u inženjerskim problemima

Dujman, Josip

SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE

ZAVRŠNI RAD

Josip Dujman

Zagreb, 2017.

SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE

ZAVRŠNI RAD

Mentor:

Doc. dr. sc. Hrvoje Cajner, dipl. ing.

Student:

Josip Dujman

Zagreb, 2017.

Izjavljujem da sam ovaj rad izradio samostalno koristeći znanja stečena tijekom studija i navedenu literaturu.

Zahvaljujem se svome mentoru, doc. dr. sc. Hrvoju Cajneru na pruženoj stručnoj pomoći i savjetima pri izradi završnog rada.

Posebno se zahvaljujem svojoj obitelji na podršci, razumijevanju i povjerenju koju su mi pružali tijekom cijelog studija.

Josip Dujman

SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE
Središnje povjerenstvo za završne i diplomske ispite
Povjerenstvo za završne ispite studija strojarstva za smjerove:
proizvodno inženjerstvo, računalno inženjerstvo, industrijsko inženjerstvo i menadžment, inženjerstvo materijala i mehatronika i robotika

Sveučilište u Zagrebu Fakultet strojarstva i brodogradnje	
Datum	Prilog
Klasa:	
Ur.broj:	

ZAVRŠNI ZADATAK

Student: **Josip Dujman** Mat. br.: **0035197112**

Naslov rada na hrvatskom jeziku: **UPORABA NEPARAMETARSKE STATISTIKE U INŽENJERSKIM PROBLEMIMA**

Naslov rada na engleskom jeziku: **THE USE OF NONPARAMETRIC STATISTICS IN ENGINEERING**

Opis zadatka:

Prilikom testiranja statističkih hipoteza veći dio problema inženjerske prakse rješava se putem metoda iz područja parametarske statistike. Međutim, uporaba parametarske statistike podrazumijeva pretpostavku o normalnoj raspodjeli podataka u osnovnom skupu odnosno promatranom procesu. U slučaju nepoznavanja i značajnog odstupanja varijabli od normalnosti, jedini način ispravnog testiranja hipoteza jest uporaba metoda neparametarske statistike. Osnovni zadatak ovog rada je izraditi presjek najučestalijih metoda neparametarske statistike uz izradu algoritma za testiranje statističkih hipoteza u dostupnom programskom paketu.

U radu je potrebno:

- Dati presjek najčešće korištenih metoda neparametarske statistike u inženjerskoj praksi.
- Sistematizirati osnovne metode po područjima primjene te izraditi hodogram njihovog odabira.
- Izraditi algoritam za testiranje statističkih hipoteza u dostupnom programskom paketu.
- Primijeniti opisane metode kroz razvijeni algoritam na nekoliko različitih primjera iz prakse.

Zadatak zadan:

30. studenog 2016.

Rok predaje rada:
1. rok: 24. veljače 2017.
2. rok (izvanredni): 28. lipnja 2017.
3. rok: 22. rujna 2017.

Predviđeni datumi obrane:
1. rok: 27.2. - 03.03. 2017.
2. rok (izvanredni): 30. 06. 2017.
3. rok: 25.9. - 29. 09. 2017.

Zadatak zadao:

Dr. sc. Hrvoje Cajner, doc.

v.d. predsjednika Povjerenstva:

Izv. prof. dr. sc. Branko Bauer

SADRŽAJ

POPIS SLIKA

POPIS TABLICA

POPIS OZNAKA

Oznaka	Jedinica	Opis
D		Maksimalna devijacija
d_n		Razlike parova
f		Broj frekvencija
$F_0(X)$		Teorijska kumulativna frekvencija
$F(X)$		Kumulativna frekvencija prvog uzorka
$G(X)$		Kumulativna frekvencija drugog uzorka
H		Testni parametar Kruskal-Wallis analize varijance
H_0		Nulta hipoteza
H_1		Alternativna hipoteza
K_D		Brojnik od maksimalne devijacije
k		Stupanj slobode, broj stupaca
N		Broj ukupnih mjerenja svih uzoraka
n		Veličina uzorka
p		Vrijednost vjerojatnosti
Q		Testni parametar Cochran Q testa
R		Suma rangova
r		Broj nizova, broj redaka
$S_n(X)$		Promatrana kumulativna frekvencija
T		Testni parametar Wilcoxon testa rangova zavisnih uzoraka
t		Računska vrijednost t-testa
W		Testni parametar Wilcoxon testa sume rangova
$\bar{x}$		Aritmetička sredina uzorka
z		Računska vrijednost z-testa
α		Granica značajnosti
σ_0^2		Varijanca populacije
σ^2		Varijanca uzorka
χ^2		Hi kvadrat varijabla
μ		Očekivanje populacije

SAŽETAK

U ovom radu su prikazane određene metode neparametarske statistike i njihove primjene u inženjerskoj praksi. Potreba za neparametarskom statistikom proizlazi iz činjenice da dobiveni podatci nemaju uvijek normalnu raspodjelu tako da se metode parametarske statistike ne mogu provoditi. Jednako tako ako mjerenja dobivenih podataka nisu u intervalnoj skali, tj. varijable uzoraka nisu kontinuirane, neparametarska statistika je jedini izbor za testiranje statističkih hipoteza. Zbog toga je u radu sistematizirana podjela najvažnijih neparametarskih testova s obzirom na broj uzoraka, vrstu i broj podataka. Za svaki test su opisane njegove mogućnosti, postupak provođenja, ograničenja te je prikazan po jedan primjer iz inženjerske prakse za svaki test. Uz navedeno izrađen je hodogram odabira neparametarskih testova kako bi se brzo i jednostavno mogao izabrati test za određenu svrhu. Za nekoliko testova su napravljeni obrasci u Microsoft Excelu, gdje se samo upisom podataka dobiva rezultat testa.

Ključne riječi: neparametarska statistika, testiranje statističkih hipoteza, testni parametar, granica značajnosti, hodogram

1. UVOD

Inferencijalna statistika se bavi procjenom parametara populacije i testiranjem hipoteza na temelju rezultata iz uzoraka. Statistički testovi sadrže postupke za testiranjem hipoteza. Oni zahtijevaju određene uvjete o tome kakvi su podatci i kako ih dobiti te određuju koliko velike moraju biti uočene razlike između uzoraka prije nego što kažemo da one predstavljaju stvarne razlike populacije.

Prije nego što se odabere statistički test potrebno je razlučiti kakvog tipa su podatci koje želimo testirati. Rezultati mjerenja mogu biti u nekoliko različitih skala: nominalna, ordinalna i intervalna. Nominalna skala podrazumijeva diskretne varijable gdje se podatci prebrojavanjem razvrstavaju u nekoliko kategorija (da i ne, muško i žensko,...). Svaka kategorija ima svoju frekvenciju koja izražava broj varijabli koje spadaju u tu kategoriju. Uz frekvenciju, može se izračunati i mod. Ordinalna skala nastaje kada imamo neku vezu između različitih kategorija. Ta veza stvara neki redoslijed primjerice jedna kategorija je veća od druge, druga je veća od treće itd. Međutim u ordinalnoj skali se ne zna točna razlika (udaljenost) između svake kategorije. Primjerice možemo reći da je neki objekt jači od drugog ali ne i za koliko. To ima za posljedicu nemogućnost uporabe aritmetičkih operacija na ordinalnoj i nominalnoj skali. Ordinalna skala omogućava izračun medijana. Intervalna skala nastaje kada znamo razliku između svih kategorija. Varijable u intervalnoj skali su kontinuirane jer nastaju mjerenjem. Intervalna skala označava brojevni pravac gdje aritmetičke operacije imaju smisla poput izračuna srednje vrijednosti.

Kako bi se opisao značaj neparametarskih statističkih testova koji su tema ovog rada potrebno je navesti pretpostavke i ograničenja najčešće korištenih statističkih testova, a to su parametarski testovi. Parametarski testovi se vrše nad podatcima koji su isključivo u intervalnoj skali. Oni zahtijevaju da su podatci normalni normalno distribuirani i da populacije moraju imati istu varijancu. Međutim kada se ne mogu ili ne žele prihvatiti takve pretpostavke ili kada rezultati mjerenja nisu u intervalnoj skali u obzir dolaze neparametarski testovi.

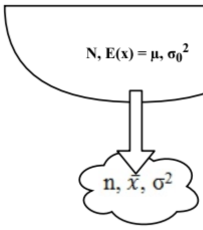
U inženjerskoj praksi neparametarski testovi se trebaju koristit kada se ne mogu ili ne žele raditi točna mjerenja (intervalna skala), već se mjerenja rade u ordinalnoj skali što može pojednostaviti i smanjiti vrijeme mjerenja, a posljedično i troškove ispitivanja. Neophodni su i kada rezultati mjerenja nisu normalno distribuirani ili su uzorci jako mali, a ne znamo točnu distribuciju populacije.

U sljedećim poglavljima razvrstani su najučestaliji neparametarski testovi s obzirom na broj uzoraka i vrstu podataka. Dati su opisi svakog testa, postupak provođenja i jedan primjer iz inženjerske prakse. Na kraju je prikazan hodogram odabira statističkih testova, a u prilogu se nalazi Microsoft Excel datoteka u kojoj su napravljeni obrasci za izračun testova jednostavnim upisom podataka u tablicu.

2. PRESJEK NAJČEŠĆE KORIŠTENIH METODA NEPARAMETARSKE STATISTIKE

2.1. Jedan uzorak

Ovo poglavlje donosi neparametarske statističke testove koji se mogu upotrijebiti za slučaj jednog uzorka. Testovi provjeravaju da li uzorak dolazi iz neke određene populacije. Izvlači se slučajni uzorak i testira hipoteza da je uzorak iz populacije određene distribucije. Slika 1 prikazuje shemu postupka.



Slika 1. Shema uzorkovanja

Iz populacije veličine N, uzima se uzorak veličine n, pri čemu su parametri populacije:

- μ – očekivanje populacije
- σ_0^2 – varijanca populacije

dok su parametri uzorka:

- $\bar{x}$ – aritmetička sredina uzorka
- σ^2 – varijanca uzorka

Stoga nam takvi testovi daju odgovore na pitanja poput: postoji li značajna razlika u lokaciji (centralnoj tendenciji) između uzorka i populacije? Postoji li značajna razlika između izmjerenih frekvencija i očekivanih, između očekivanih proporcija i izmjerenih? Može li se smatrati da uzorak dolazi iz populacije određenog oblika ili forme (primjerice normalna, pravokutna)? Može li se smatrati da je uzorak slučajni uzorak neke poznate populacije? Uobičajeni parametarski test za jedan uzorak je t-test koji govori o razlici srednje vrijednosti uzorka i populacije. t-test pretpostavlja da rezultati mjerenja dolaze iz populacije koja ima normalnu razdiobu i da su rezultati mjerenja u intervalnoj skali. Međutim postoje podatci nad kojima se ne može izvršiti t-test. Ako su pretpostavke i zahtjevi t-testa neprihvatljivi za podatke, ako se želi donijeti zaključak o općenitoj razlici između uzorka i populacije, ako su podatci u rangovima ili u nominalnoj skali, ako nas ne zanima razlika u lokaciji nego se želi otkriti bilo kakva različitost. U takvim slučajevima se trebaju odabrati neparametarski statistički testovi predstavljeni u ovom poglavlju.

2.1.1. Binomni test

Binomni test se koristi kod podataka koji imaju samo dva moguća ishoda, kategorije. Za takve slučajeve svi podaci iz populacije će biti u jednoj od diskretnih kategorija. Primjeri kategorija: ispravan i neispravan, muško i žensko, događaj A i događaj B.

Ako znamo proporciju (vjerojatnost) jedne kategorije *P* tada znamo i proporciju druge kategorije 1-*P*. Ako znamo ili pretpostavimo proporciju (ili vjerojatnost pojave određene kategorije) *P* populacije onda možemo testirati da li naš uzorak dolazi iz takve populacije. Drugim riječima binomni test nam kazuje da li proporcije ili frekvencije izmjerenog uzorka mogu biti iz distribucije s određenom vrijednošću *P*.

Ako znamo vrijednost *P* za populaciju ne može se očekivati da će slučajni uzorak imati istu proporciju *P* kao i populacija. Primjerice ako znamo da mala serija proizvoda ima točno 80 ispravnih i 20 neispravnih proizvoda, slučajnim uzorkovanjem se može dobiti rezultat da je vjerojatnost pojave 80 ispravnih komada 75%, tj. veća ili manja vjerojatnost.

Vjerojatnosti da se pojavi *x* podataka jedne kategorije i *N*-*x* podataka iz druge kategorije je dana u jednadžbi (2.1):

$$P_x = \binom{n}{x} p^x q^{n-x} \tag{2.1}$$

Primjerice ako bacamo kocku 5 puta, koja je vjerojatnost da ćemo dobiti dvije šestice [1]? U ovom slučaju *N* je 5 (broj mjerenja), *x* je 2 (broj šestica), *p* je 1/6 tj. vjerojatnost pojavljivanja šestice i *q* = 1- *p* = 5/6. Vjerojatnost da dobijemo dvije šestice je:

$$P_2 = \binom{5}{2} p^2 q^{5-2} = \binom{5}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^3 = 0{,}16$$

Međutim ako nas zanima vjerojatnost da je *x* veći ili manji od određene vrijednosti onda ćemo sumirati vjerojatnosti da se dogodi *x* i sve manje ili veće od *x*, kao što je prikazano u jednadžbi (2.2):

$$P_x = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} \tag{2.2}$$

Primjerice vjerojatnost da dobijemo dvije ili manje šestica $p(x \leq 2)$ je vjerojatnost da dobijemo nula šestica $p(x = 0)$, jednu šesticu $p(x = 1)$ i dvije šestice $p(x = 2)$. Tada ćemo sumirati vjerojatnosti $p(x \leq 2) = p(x = 0) + p(x = 1) + p(x = 2)$.

$p(x<2) = p(0) + p(1)+ p(2) = 0.40+ 0.40 + 0.16 = 0.96$

Odredili smo vjerojatnost dobivanja dvije ili manje šestica u bacanju kocke koja iznosi 96%.

2.1.2. χ^2 test za jedan uzorak

χ^2 test koristimo kada želimo otkriti postoji li značajna razlika među frekvencijama pojedinih kategorija. Broj kategorija je dva ili više. Testiramo postoje li značajne razlike između promatranih i očekivanih frekvencija. Očekivane frekvencije su definirane u hipotezi H_0. Hipotezu H_0 testiramo na sljedeći način [2]:

$$\chi 2 = \sum_{i=1}^{n} \frac{(f_i - f_{ti})^2}{f_{ti}} \tag{2.3}$$

gdje su:

- f_i - promatrani broj frekvencija u *i*-toj kategoriji
- f_{ti} – očekivani broj frekvencija u *i*-toj kategoriji pod hipotezom H_0
- $\sum$ - suma svih *n* kategorija

Ako je razlika između promatranih i očekivanih frekvencija mala, posljedično će i χ^2 biti mali. Ako je razlika velika izračunati χ^2 prema (2.3) će biti veliki. Prema tome što je χ^2 veći, manja je vjerojatnost da promatrane frekvencije dolaze iz populacije na kojoj se temelji hipoteza H_0. Testni parametar χ^2 prati hi-kvadrat distribuciju s $k = n$ -1, gdje je k stupanj slobode. Tablica B u prilogu prikazuje raspodjelu i njezine kritične vrijednosti.

Ako je vjerojatnost pojavljivanja vrijednosti χ^2 pod hipotezom H_0 jednaka ili manja od prethodno odabrane granice značajnosti α, tada se hipoteza H_0 može odbaciti. Ako nije H_0 se prihvaća.

Primjer 2.1

Želimo testirati hipotezu da 8 radnika naprave jednaku količinu proizvoda u jednom danu, tj. da se broj komada značajno ne razlikuje od radnika do radnika. U tablici 1 su dani broj komada koji je svaki napravio u jednom danu, a očekivani broj prema normi je 50.

1. Nulta hipoteza. -H_0: ne postoji razlika između broja komada među radnicima

2. Statistički test. -χ^2 test je odabran jer se uspoređuju promatrane i očekivane frekvencije.

3. Granica značajnosti. -Neka α bude 0,05.

4. Distribucija uzorkovanja. -Testni parametar χ^2 slijedi hi-kvadrat distribuciju s $k = n$ -1 stupnjeva slobode.

5. Područje odbacivanja. -H_0 će se odbaciti ako je vrijednost χ^2 takva da je njezina vjerojatnost pojavljivanja pod hipotezom H_0 jednaka ili manja od α = 0,05.

Tablica 1. Promatrane frekvencije u jednom danu

Radnik	1	2	3	4	5	6	7	8
Broj komada	47	59	61	45	48	40	55	53

Izračun χ^2 prema jednadžbi (2.3):

$$\chi 2 = \sum_{i=1}^{n} \frac{(f_i - f_{ti})^2}{f_{ti}} = \frac{(47-50)^2}{50} + \frac{(59-50)^2}{50} + \frac{(61-50)^2}{50} + \cdots = 7{,}48$$

Iz tablice B je vidljivo da je za χ^2 = 7,48 i k = 7 vjerojatnost pojavljivanja između 0,5 i 0,8. Pošto je granica značajnosti α = 0,05, ne možemo odbaciti hipotezu H_o.

Ako je k = 1 svaka očekivana frekvencija ne bi smjela biti manja od 5. Kada je k > 1 test se ni bi smio koristiti ako je više od 20% očekivanih frekvencija manje od 5 ili bilo koja frekvencija manja od 1 [1]. Očekivane frekvencije se mogu povećati spajanjem susjednih kategorija samo ako to ima smisla. Primjerice ako postoje kategorije: u potpunosti se ne slažem, djelomično se ne slažem, djelomično se slažem i u potpunosti se slažem, tada možemo spojiti prve dvije i zadnje dvije navedene. Međutim ako imamo samo dvije kategorije s frekvencijama manjim od 5 onda bi se binomni test trebao koristiti [1].

2.1.3. Kolmogorov-Smirnov test za jedan uzorak

Kolmogorov-Smirnov test ispituje dolazi li uzorak iz neke određene teoretske distribucije. Test uspoređuje kumulativnu frekvenciju distribucije uzorka s kumulativnom frekvencijom određene teorijske distribucije [3]. Određuje se točka najveće razlike između promatrane i teorijske distribucije te se na temelju distribucije uzoraka određuje je li tako velika razlika slučajna. To jest da li je razlika te veličine posljedica slučajnog uzorkovanja teorijske distribucije.

Neka je $F_0(X)$ potpuno određena kumulativna frekvencija (teorijska), a $S_N(X)$ promatrana kumulativna frekvencija slučajnog uzorka od N mjerenja. Pod hipotezom H_0, da uzorak potiče iz teoretske distribucije, se podrazumijeva da za svaku vrijednost X, $S_N(X)$ treba biti približno jednak $F_0(X)$. To znači da očekujemo da su razlike male i u granicama slučajne pogreške.

Potrebno je izračunati najveću devijaciju. Najveća vrijednost $F_0(X) - S_N(X)$ se zove maksimalna devijacija D, a računa se prema jednadžbi (2.4):

$$D = maksimum\ |F_0(X) - S_N(X)| \tag{2.4}$$

Distribucija uzorkovanja za testni parametar D je dana u tablici C. Može se primijetiti da granica značajnosti određene vrijednosti D ovisi o broju mjerenja N.

Primjer 2.2

Pretpostavimo da želimo istražiti postoji li sklonost kupaca određenoj boji istog proizvoda. Napravili smo anketu s desetero slučajno odabranih ljudi. Svakoj osobi je prikazan isti proizvod, ali svaki u različitoj boji. Pet boja je predstavljeno: crvena, plava, zelena, žuta i ljubičasta. Nakon toga osoba izabire koje boje bi bio proizvod u slučaju kupovine. Ako je boja nebitna pri kupnji proizvoda svaka boja bi trebala biti jednako često izabrana. Ako je boja bitna osobe bi trebale favorizirati jednu od boja.

1. Nulta hipoteza. -Ne postoji razlika u očekivanom broju boja, a svaka uočena razlika je slučajna koja je nastala slučajnim uzorkovanjem iz pravokutne populacije gdje su $f_1 = f_2 = ... = f_5$. H_1 : frekvencije $f_1, f_2, ..., f_5$ nisu sve jednake.

2. Statistički test. -Odabran je Kolmogorov-Smirnov test jer želimo usporediti promatranu distribuciju uzorka, čiji su rezultati u ordinalnoj skali, s teoretskom distribucijom.

3. Granica značajnosti. -Neka je α = 0,01 za N = 10 ispitanika.

4. Distribucija uzorkovanja. -Kritične vrijednosti D su dane u tablici C zajedno s vjerojatnošću pojavljivanja pod H_0.

5. Područje odbacivanja. -Područje odbacivanja se sastoji od vrijednosti D koje su toliko velike da je vjerojatnost njihovog pojavljivanja pod hipotezom H_0 jednaka ili manja od α = 0,01.

Pretpostavimo da jedna osoba izabere plavu, pet osoba izaberu žutu i četiri izaberu ljubičastu boju. Podatci su prikazani u tablici 2.

Tablica 2. Kumulativne frekvencije

	crvena	plava	zelena	žuta	ljubičasta
f	0	1	0	5	4
$F_0(X)$	1/5	2/5	3/5	4/5	5/5
$S_{10}(X)$	0/10	1/10	1/10	6/10	10/10
$\lvert F_0(X) - S_N(X)\rvert$	2/10	3/10	5/10	2/10	0

Zadnji red daje maksimalnu devijaciju D koja iznosi 5/10 tj. 0,5. Tablica C pokazuje da je za N =10, $D \geq$ 0,5 vjerojatnost pojavljivanja pod H_0 p < 0,1. Pošto je vjerojatnost manja od granice značajnosti α = 0,01 odbacujemo H_0 i prihvaćamo H_1. Zaključujemo da naši ispitanici pokazuju značajnu sklonost određenim bojama tj. da je boja bitna pri kupnji toga proizvoda.

2.1.4. Test homogenosti niza

Zaključak o nekoj populaciji, koristeći podatke iz uzorka, je valjan samo ako je uzorak slučajan. Stoga je korisno testirati je li uzorak slučajan. Također koristi se u kontroli kvalitete kod detekcije defekata. Metoda koja će ovdje biti predstavljena se temelji na broju nizova koje uzorak posjeduje. Niz je definiran kao slijed identičnih simbola koji slijede i prethode druge simbole.

Pretpostavimo da uzorak ima niz plusova i minusa kako je zadano:

$$+ + - - - + - - - - + + - +$$

Ovaj uzorak počinje s nizom od 2 plusa, zatim slijedi niz od 3 minusa, pa niz od jednog plusa itd. Ako grupiramo ove nizove vidimo da imamo 7 nizova. r = broj nizova = 7. Ukupan broj nizova u uzorku daje naslutiti je li uzorak slučajan. Premali ili preveliki broj nizova u uzorku daje naslutit da uzorak nije slučajno odabran.

Postupak

Neka je n_1 broj plusova, n_2 broj minusa, a N zbroj svih događaja $n_1 + n_2$. Prvo se na temelju podataka odredi broj nizova r. Ako su n_1 i n_2 manji ili jednaki od 20 koristi se tablica D koja daje kritične vrijednosti r pod hipotezom H_0 za $\alpha = 0{,}05$. Ako je r između kritičnih vrijednosti prihvaćamo H_0, ako je jednak ili veći od njih odbacujemo H_0.

Za velike uzorke tj. kada su n_1 i n_2 veći od 20, tablica D se ne može koristiti već se radi aproksimacija r pomoću normalne distribucije, gdje se izračuna jedinična normalna varijabla z prema jednadžbi (2.5):

$$z = \frac{r - \left(\frac{2n_1n_2}{n_1+n_2} + 1\right)}{\sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)}}} \qquad \textbf{(2.5)}$$

Za velike uzorke vrijednost z, koja je izračunata iz jednadžbe (2.5), je približno normalno distribuirana. Tablica A daje vjerojatnost pojavljivanja z pod hipotezom H_0. Ako je vjerojatnost z jednaka ili manja od granice značajnosti α odbacujemo H_0.

2.2. Slučaj dva zavisna uzorka

Uzorci su zavisni kada uspoređujemo objekte mjerenja prije i nakon nekakvog djelovanja na njih kako bi otkrili ima li to djelovanje neki značaj na ishod mjerenja. Zavisni su i kada na iste objekte mjerenja djelujemo na dva različita načina kako bi otkrili imaju li djelovanja drugačije učinke na ishod mjerenja. Za posljednji način kažemo da su uzorci upareni jer su objekti (članovi) mjerenja jednaki tj. članovi u dva uzorka moraju biti jednaki u smislu da na ishod mjerenja utječu samo različita djelovanja, a ne i različitost objekata (članova) u uzorcima. Broj članova svakog uzorka mora biti jednak.

2.2.1. McNemar test

McNemar test se koristi za provjeru značajnosti promjena. Na uzorke se djeluje te se mjeri stanje prije i poslije. Tako se testira efektivnost djelovanja (tretmana) ako su rezultati mjerenja u nominalnoj skali.

Pretpostavimo da imamo određeni broj ispitanika u anketi čiji su odgovori *da* ili *ne*. Nakon bilježenja odgovora ispitanici su podvrgnuti propagandnom materijalu nakon čega opet daju svoje odgovore.

Neka je A broj ispitanika koji su promijenili svoj odgovor s *da* prema *ne*, a D broj ispitanika koji su promijenili svoj odgovor s *ne* prema *da* kako je prikazano u tablici 3.

Tablica 3. Tablica promjena

		Poslije	
		NE	DA
Prije	DA	A	B
	NE	C	D

Ako u nultoj hipotezi kažemo da propagandni materijal nema utjecaja na odgovore ispitanika tada A treba biti jednak D, što znači da je količina promjena u oba smjera jednaka. Kako A + D predstavljaju ukupan broj promjena, očekivanje nulte hipoteze kaže da je očekivani broj promjena u jednom smjeru ½(A + B). Testni parametar ima hi-kvadrat distribuciju s jednim stupnjem slobode gibanja i računa se prema izrazu u kojem je uključena Yatesova korekcija kontinuiteta [1]:

$$\chi 2 = \frac{(|A - B| - 1)^2}{A + B} \tag{2.6}$$

Izračunata vrijednost $\chi 2$ se uspoređuje s kritičnom vrijednošću $\chi 2$ iz tablice B za određenu razinu značajnosti α i stupnjem slobode $k = 1$. Ako je izračunata vrijednost $\chi 2$ jednaka ili veća od one iz tablice nulta hipoteza se odbacuje, što znači da postoji značajan utjecaj prije i nakon tretmana.

Ako je A = 40 i B = 25, $\chi^2 = 3{,}02$. Za $\alpha = 0{,}05$ i $k = 1$, $\chi 2_{\text{krit}} = 3{,}84$. Prema tome ne možemo odbaciti nultu hipotezu.

Ako je A + B < 25 McNemar test se ne bi trebao koristiti. Umjesto njega se preporuča jednostrani binomni test s $p = 0{,}5$ i $n = A + B$ [1].

2.2.2. Wilocoxon test rangova zavisnih uzoraka

Ovaj test se koristi kada su zavisni (upareni) uzorci u ordinalnoj ili intervalnoj skali i kada ne želimo prihvatiti pretpostavke za *t*-test. Promatra se razlika svakog para te se na temelju toga zaključuje o utjecaju dva različita tretmana. U tablici 4 je dana struktura podataka uparenih uzoraka.

Tablica 4. Struktura podataka [4]

Par	1	2	...	n
Tretman A	X_{11}	X_{12}	...	X_{1n}
Tretman B	X_{21}	X_{22}	...	X_{2n}
Razlika d_n (A – B)	d_1	d_2	...	d_n

pri čemu je $d_n = X_{1n} - X_{2n}$

Prvo treba rangirati razlike d od najmanje prema najvećem gledajući apsolutnu vrijednost razlike tj. ne gledajući oznaku plus ili minus. Rang 1 najmanjem d, rang 2 slijedećem najmanjem itd. Nakon toga se svakom d pridruži oznaka (+/-) tako da se zna koja razlika je negativna, a koja pozitivna.

U slučaju da je razlika jednaka nuli takav par se izbacuje iz analize, a u slučaju da postoji više jednakih razlika d tada svakom od njih dodjeljujemo isti rang na sljedeći način. Ako se u četiri para pojavi razlika -1, -1, +1, +1 svaki d će dobiti rang 2,5 jer je prosjek rangova $\frac{1+2+3+4}{4} = 2{,}5$. Peti po redu d će dobiti rang 5.

Nulta hipoteza kaže da uzorci dolaze iz iste populacije, tj. da su tretmani A i B jednaki. Ako je tako tada suma rangova pozitivnih razlika i suma rangova negativnih razlika treba biti podjednaka. Međutim ako je suma pozitivnih rangova značajno veća od sume negativnih rangova tada zaključujemo da uzorci ne dolaze iste populacije te odbacujemo nultu hipotezu. Testni parametar je suma rangova koja se, ovisno o literaturi, označava s T ili W (u nastavku će se koristiti oznaka T). Za vrijednost T se uzima suma ranga koji je manji po iznosu, dakle ili pozitivni ili negativni rang. Ako je broj članova $N \leq 25$ koristi se tablica E koja daje kritične vrijednosti T s obzirom na granicu značajnosti testa [4]. Ako je izračunati T jednak ili manji od očitanoga za određenu granicu značajnosti, odbacujemo nultu hipotezu.

Za velike uzorke $N > 25$, tablica E se ne može koristiti stoga se statistika T aproksimira normalnom distribucijom $N(\mu, \sigma)$ gdje su

$$\mu = \frac{N(N+1)}{4}$$

$$\sigma = \sqrt{\frac{N(N+1)(2N+1)}{24}}$$

posljedično

$$z = \frac{T-\mu}{\sigma} = \frac{T - \frac{N(N+1)}{4}}{\sqrt{\frac{N(N+1)(2N+1)}{24}}} \tag{2.7}$$

z je normalno distribuirana varijabla čije vrijednosti možemo očitati iz tablice A za određenu granicu značajnosti, usporediti s izračunatom i donijeti odluku o odbacivanju nulte hipoteze.

Primjer 2.3

Dvanaest radnika je izabrano iz jedne tvornice kako bi testirali dvije vrste montažne opreme. Svi radnici znaju raditi na obje vrste opreme. Svakom radniku je mjereno vrijeme obavljanja radnog zadatka na obje vrste opreme. Želimo saznati je li jedna vrsta opreme brža od druge. Podatci su dani u tablici 5.

Tablica 5. Vrijeme potrebno za radni zadatak

Radnik	Oprema 1 [min]	Oprema 2 [min]	Oznaka	Razlika	Rang	Članovi manjeg ranga
1	17	18	-	1	2	2
2	16	14	+	2	6	
3	21	19	+	2	6	
4	14	11	+	3	9,5	
5	18	23	-	5	11	11
6	24	21	+	3	9,5	
7	16	10	+	6	12	
8	14	13	+	1	2	
9	21	19	+	2	6	
10	23	24	-	1	2	2
11	13	15	-	2	6	6
12	18	20	-	2	6	6
						$T = 27$

Prvo se računa razlika između svakog pojedinog para. Zatim se razlike rangiraju prema apsolutnoj vrijednosti. Kako imamo izjednačene vrijednosti tako im dodijelimo srednju vrijednost ranga, u ovom slučaju tri jedinice zauzimaju rangove 1, 2, i 3 stoga im dodjeljujemo srednju vrijednost 2. Zbroje se rangovi koji potječu od negativne T^- i pozitivne T^+ razlike. Minimalna vrijednost sume, koja iznosi 27 se uspoređuje s kritičnom vrijednošću iz tablice E, pri čemu je kritična vrijednost za α=0,05 i n = 12, T_{krit} = 14. Kako je $T^- > T_{krit}$, ne možemo odbaciti nultu hipotezu, što znači da ne možemo tvrditi da jedna vrsta opreme uzrokuje duže ili kraće vrijeme obavljanja radnog zadatka.

2.2.3. Randomizacijski test

Randomizacijski test za zavisne uzorke se koristi kada su varijable u intervalnoj skali i kada je broj članova uzoraka dovoljno mali kako bi se mogao provesti izračun rezultata. Daje točnu vjerojatnost testnog parametra za nultu hipotezu bez pretpostavki o normalnosti i homogenosti varijance.

Postupak

1. Izračunati razliku svakog para d_i

2. Izračunati broj mogućnosti/permutacija d_i s obzirom na znak plus i minus pri čemu je broj mogućnosti jednak 2^N.

3. Izračunati broj mogućih ishoda u području odbacivanja: $\alpha \times 2^N$.

4. Najveće sume razlika $\sum d_i$ se nalaze u području odbacivanja (ekstremni slučajevi). Za jednostrani test suma razlika je ili najveća pozitivna ili negativna vrijednost. Za dvostrani test pola ishoda je u području sume s pozitivnim, a pola s negativnim razlikama.

5. Ako je izračunata suma razlika $\sum d_i$ jedan od uočenih ekstremnih slučajeva, tj. ako je u području odbacivanja, odbacujemo nultu hipotezu.

Primjer 2.4 [1]

Uzmimo u obzir razliku mjerenja osam parova između dva uzorka. Svaki član uzorka je nasumično odabran. Razlike iznose:

+19 +27 -1 +6 +7 +13 -4 +3

Pod nultom hipotezom smatramo da su pozitivne ili negativne razlike ishod slučaja jer smo nasumično odabrali članove uzorka. Jednako tako je +19 mogao biti -19, kao i sve ostale razlike. Niz razlika je mogao primjerice izgledati i ovako:

-19 +27 +1 -6 -7 +13 +4 -3

Prema tome postoji 2^N različitih ishoda, a u ovom slučaju ima 256 mogućih kombinacija. Pod nultom hipotezom smatramo da je vjerojatnost pojavljivanja bilo koje kombinacije jednaka. Za svaki ishod računamo sumu razlika $\sum d_i$. Mnoge sume razlika su kreću oko nule, što i očekujemo ako je nulta hipoteza točna. Sume koje su velike po vrijednosti dolaze od onih kombinacija gdje su gotovo svi znakovi plus ili minus, što znači da srednja vrijednost jednog uzorka premašuje onu od drugog. To znači da je nulta hipoteza netočna. Ako želimo testirati nultu i alternativnu hipotezu moramo uspostaviti zonu odbacivanja koja se sastoji od najvećih suma. Ako odaberemo granicu značajnosti od 0,05 područje odbacivanja će se sastojati od 12 mogućih kombinacija jer je 0,05 x 256 = 12,8. Pod nultom hipotezom vjerojatnost pojavljivanja ovih 12 ekstremnih kombinacija je $\frac{12}{256} = 0{,}047$. Neka je granica značajnosti α = 0,05. Ako je promatrana suma razlika jedna od ovih ekstremnih kombinacija odbacujemo nultu hipotezu.

Za jednostrani test područje odbacivanja se sastoji od 12 najvećih pozitivnih ili negativnih suma. U dvostranom testu područje odbacivanja se sastoji od 6 najvećih pozitivnih i negativnih suma.

Ako želimo napraviti dvostrani test, 6 ekstremnih slučajeva ćemo prikazati u tablici 6.

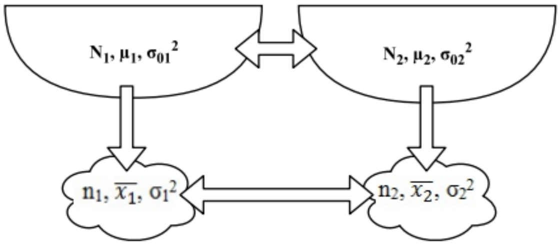
Tablica 6. Ekstremni ishodi za dvostrani test

	Ishodi								$\sum d_i$
1	+19	+27	+1	+6	+7	+13	+4	+3	80
2	+19	+27	-1	+6	+7	+13	+4	+3	78
3	+19	+27	+1	+6	+7	+13	+4	-3	74
4	+19	+27	+1	+6	+7	+13	-4	+3	72
5	+19	+27	-1	+6	+7	+13	+4	-3	72
6	+19	+27	-1	+6	+7	+13	-4	+3	70

Ishod pod brojem 6 je onaj koji smo uočili. Vjerojatnost njegova pojavljivanja je p = 0,047. Kako je $p < \alpha$ odbacujemo nultu hipotezu koja tvrdi da nema razlika između uzoraka.

2.3. Slučaj dva nezavisna uzorka

Kada se ne mogu osigurati zavisni uzorci tj. upareni uzorci koriste se nezavisni. Za razliku od zavisnih uzoraka koji dolaze iz iste populacije, nezavisni uzorci se slučajno odabiru iz dvije populacije, kako je prikazano na slici 2.



Slika 2. Shema testiranja uzoraka i populacija

Iz populacija veličine N_1 i N_2 izvlače se uzorci veličina n_1 i n_2 pri čemu su parametri populacija:

- μ – očekivanje populacije
- σ_0^2 – varijanca populacije

dok su parametri uzorka:

- $\bar{x}$ – aritmetička sredina uzorka
- σ^2 – varijanca uzorka

Testovi koji slijede, jednako kao i testovi prikazani u prethodnom poglavlju, ispituju vjerojatnost da dva uzorka dolaze iz iste populacije, tj. testiraju njihovu različitost. Jedna grupa testova ispituje razlike u lokaciji distribucije uzoraka dok druga ispituje bilo kakvu različitost. U prvu grupu spadaju Fisherov test, Medijan test, Wilcoxon test sume rangova, Kolmogorov-Smirnov za dva uzorka (jednostrani) i randomizacijski test. Drugu grupu čine: χ^2 test i Kolmogorov-Smirnov (dvostrani).

2.3.1. Fisherov test

Fisherov test se koristi za analiziranje diskretnih varijabli koje mogu biti u nominalnoj ili ordinalnoj skali. Test ima zahtjevan izračun stoga se koristi za male uzorke. Uzorci koji se uspoređuju se svrstavaju u dvije kategorije što čini 2 x 2 tablicu kontingencije [Tablica 7].

Ispituje se razlikuju li se značajno uzorci u proporcijama koje zauzimaju u određenim kategorijama (kategorija + i -).

Tablica 7. Tablica kontingencije

	+	-	Suma reda
Uzorak 1	A	B	A + B
Uzorak 2	C	D	C + D
Suma stupca	A + C	B + D	N

Test računa točnu vjerojatnost *p* da se određeni set frekvencija pojavi u tablici kontingencije, dok su sume stupaca i redova konstantne, pomoću hipergeometrijske raspodjele (2.8) [6]:

$$p_{uzorka1} = \frac{\binom{A+C}{A}\binom{B+D}{B}}{\binom{N}{A+B}} \tag{2.8}$$

Pretpostavimo da inženjer želi istražiti utječe li duljina toplinske obrade u određenoj fazi proizvodnog procesa na krhkost polietilenskih pločica. Za testiranje inženjer 10 pločica toplinski obrađuje 45 sekundi i 9 pločica obrađuje 60 sekundi. Svaka pločica se testira i klasificira kao krhka ili žilava. Podatci su dani u tablici 8. [5]

Tablica 8. Broj žilavih i krhkih komada

Toplinska obrada	Krhka	Žilava	Ukupno
1	1	9	10
2	6	3	9
Ukupno	7	12	19

Pomoću Fisherovog testa se želi testirati hipotezu da kraća toplinska obrada uzrokuje manju krhkost (jednostrani test), nasuprot nulte hipoteze koja kaže da nema značajnih razlika između tretmana te očekujemo da je 7 krhkih i 12 žilavih komada slučajno raspoređeno u skupine od 10 i 9 komada.

Vjerojatnost da se slučajnim odabirom dobije točno 1 krhki komad se računa pomoći hipergeometrijske raspodjele na slijedeći način:

$$\frac{\binom{7}{1}\binom{12}{9}}{\binom{19}{10}} = \frac{7 * 220}{92378} = 0{,}0167$$

Međutim, ako tvrdimo da toplinska obrada 1 uzrokuje manju krhkost trebamo izračunati još ekstremnije slučajeve koje podupiru tu tezu, što znači da nam treba vjerojatnost da se u grupi od 10 komada (toplinska obrada 1) ne pojavljuje niti jedan krhki komad.

$$\frac{\binom{7}{0}\binom{12}{10}}{\binom{19}{10}} = \frac{1 * 66}{92378} = 0{,}0007$$

Da bi odredili jesu li uočene frekvencije u suprotnosti s konceptom slučajnog odabira trebamo izračunati ukupnu vjerojatnost uočenog i još ekstremnijeg ishoda, a on iznosi p = 0,0167 + 0,0007 = 0,0174. Ako uzmemo da je α = 0,05 tada je $p < \alpha$ te možemo odbaciti nultu hipotezu koja kaže da su razlike slučajne. Ako se želi učiniti dvostrani test vjerojatnost p se mora udvostručiti.

2.3.2. χ^2 test za dva nezavisna uzorka

Koristi se kada su rezultati mjerenja u nominalnoj skali ili točnije rečeno kao frekvencije diskretnih kategorija. Želi se testirati postoji li značajna razlika između dva nezavisna uzorka što znači postoji li značajna razlika između frekvencija (alternativna hipoteza) ili su razlike slučajne (nulta hipoteza). Veličina testa χ^2 daje odnos stvarnih f_{ij} i očekivanih ft_{ij} frekvencija, a računa se prema jednadžbi (2.9):

$$\chi 2 = \sum_{j=1}^{s}\sum_{i=1}^{r}\frac{\left(f_{ij} - ft_{ij}\right)^2}{ft_{ij}} \qquad \textbf{(2.9)}$$

Prvo se dobivene frekvencije napišu u tablici kontingencije sa r redaka i s stupaca. Dakle stupanj slobode je $k = (r - 1)(s - 1)$. Zatim se odrede očekivane frekvencije za svako polje tablice. Nakon toga se izračuna χ^2 prema jednadžbi (2.9) i usporedi s vrijednošću iz tablice B za postavljenu granicu značajnosti. Ako je vjerojatnost pojave χ^2 jednaka ili manja od granice značajnosti odbacujemo nultu hipotezu.

Ograničenja

Očekivane frekvencije ne smiju biti manje od 5 inače se koristi Fisherov test [1].

Primjer 2.5 [5]

Kad se novi proizvod izbacuje na tržište, proizvođaču je važno procijeniti uspješnost proizvoda tokom prvih mjeseci distribucije. Želi se utvrditi uspješnost na tržištu proizvoda u 2 grada. Marketinški odjel tvrtke uzorkuje 200 i 150 potrošača iz dva grada i dobiva sljedeće podatke prema tablici 9:

Tablica 9. Tablica kontingencije s rezultatima ispitivanja

	Nikad čuo za proizvod		Čuo, ali nije kupio		Kupio najmanje jedanput		Ukupno
Grad 1	36	*46,3*	55	*63,4*	109	*90,3*	200
Grad 2	45	*34,7*	56	*47,6*	49	*67,7*	150
Ukupno	81		111		158		350

Očekivane frekvencije su također dane u tablici, pisane u kurzivu, a računaju se pomoću parcijalnih suma. Suma *s*-tog stupca se pomnoži sa sumom *r*-tog retka i podijeli s ukupnim brojem frekvencija. Primjer za ćeliju $f_{11} = \frac{81*200}{350} = 46{,}3$. χ^2 prema (2.9) iznosi 16,98. Kako je $k = (2-1)*(3-1) = 2$, za $\alpha = 0{,}05$, χ^2_{krit} iznosi 5,99. Pošto je $\chi^2 > \chi^2_{\mathrm{krit}}$ odbacujemo nultu hipotezu te možemo zaključiti da je uspješnost na tržištu različita između gradova.

2.3.3. *Medijan test*

Medijan test se koristi kada se želi saznati dolaze li dva uzorka iz populacije s istim medijanom. Nulta hipoteza glasi da dva uzorka dolaze iz populacije s istim medijanom dok alternativna hipoteza može tvrditi da su medijani različiti, što podrazumijeva dvostrani test, ili da je jedan medijan veći od drugog dakle jednostrani test. Test se koristi kada su podatci najmanje u ordinalnoj skali.

Tablica 10. Tablica frekvencija

	Uzorak 1	Uzorak 2
Broj rezultata koji su iznad medijana	A	B
Broj rezultata koji su ispod medijana	C	D

Postupak

1. Pronaći zajednički medijan svih rezultata (n_1 + n_2)

2. Napraviti tablicu frekvencija 2 x 2 po predlošku iz tablice 10, gdje se u prvi redak piše broj rezultata koji su veći od medijana, a u drugi redak broj rezultata manjih od zajedničkog medijana.

3. Izvršiti χ^2 test, prema jednadžbi (2.10), na dobivenoj tablici (ako su n_1 + n_2 < 20 koristiti Fisherov test) za stupanj slobode k = 1 [1].

$$\chi 2 = \frac{N\left(|AD - BC| - \frac{N}{2}\right)^2}{(A+B)(C+D)(A+C)(B+D)} \qquad \textbf{(2.10)}$$

Primjer 2.6

Proizvođač mobilne opreme želi ispitati dvije vrste baterija. Baterije se stavljaju pod istu razinu opterećenja i mjeri se porast temperature nakon 15 minuta rada. Rezultati mjerenja su dani u tablici 10.

Tablica 11. Temperatura mjerenja u [˚C]

Baterija 1	25	27	29	31	30	26	24	32	33	38	24	27	29	33
Baterija 2	31	33	32	35	34	29	38	35	37	30	33	29	36	38

Zajednički medijan iznosi 31. Tablica 12 daje frekvencije onih vrijednosti koje su manje ili veće od medijana. Kako imamo dva rezultata koja su jednaka 31, možemo ih dodijeliti rezultatima koje su ili iznad ili ispod medijana, u ovom slučaju neka budu ispod.

Tablica 12. Frekvencije rezultata

	Baterija 1	Baterija 2
Iznad medijana	4	10
Ispod medijana	10	4

Prema (2.10) χ^2 = 3,57. Iz tablice B za α = 0,05 i k = 1, χ^2_{krit} = 3,841. Kako χ^2 nije veći od χ^2_{krit} ne možemo odbaciti nultu hipotezu, te zaključujemo da uzorci dolaze iz populacije s istim medijanima.

2.3.4. Wilcoxon test sume rangova

Ovaj test koristimo ako imamo mjerenja u ordinatnoj ili intervalnoj skali dva nezavisna uzorka i želimo testirati njihovu različitost. Test vrlo sličan ovome se zove Mann-Whitney U-test i izvodi se u malo drugačijoj formi, ali rezultati su jednaki. Hipoteze koje se mogu testirati možemo postaviti na slijedeće načine:

H_0: dvije kontinuirane distribucije su jednake

H_1: jedna distribucija ima vrijednosti koje su sustavno veće

Ili pak možemo pobliže testirati parametre distribucije poput *t*-testa, ali ne želimo pretpostaviti normalnost distribucije, te hipoteze mogu biti slijedeće:

H_0: $\mu_1 = \mu_2$

H_1: $\mu_1 > \mu_2$ ili $\mu_1 < \mu_2$ ili $\mu_1 \neq \mu_2$

Jednako vrijedi i za testiranje medijana. Postupak ovog testa je kako slijedi. Neka su X_{11}, X_{12}, ..., X_{1n1} *i* X_{21}, X_{22}, ..., X_{2n2} su članovi dva nezavisna uzorka veličine $n_1 \leq n_2$ i želimo testirati hipoteze [4]:

H_0: $\mu_1 = \mu_2$

H_0: $\mu_1 \neq \mu_2$

Prvo posložimo sve članove u rang od najmanjeg do najvećeg i pripišemo im rang. Ako su dva mjerenja jednaka koristimo srednju vrijednost ranga kako je opisano u prijašnjem Wilcoxon testu za zavisne uzorke. Neka je W_1 suma rangova manjeg uzorka,a W_2 suma rangova većeg. Računa se prema jednadžbi (2.11):

$$W_2 = \frac{(n_1 + n_2)(n_1 + n_2 + 1)}{2} - W_1 \tag{2.11}$$

Ako se očekivanja uzoraka ne razlikuju, očekivat ćemo da će suma rangova uzoraka biti približno jednaka. Suprotno, ako se sume rangova značajno razlikuju, očekivat ćemo da očekivanja uzoraka nisu jednaka. Tablica F u prilogu daje kritične vrijednosti *W* za granice značajnosti. Nulta hipoteza se odbacuje ako su W_1 ili W_2 jednaki ili manji tabličnim kritičnim vrijednostima.

Primjer 2.7

Mjeri se srednje aksijalno naprezanje strukturalnog dijela zrakoplova. Taj dio se izrađuje od dvije legure: legure 1 i legure 2. Deset uzoraka svake legure se testira i mjeri njihovo aksijalno naprezanje. Podatci uzoraka su prikazani u tablici 13 [4]. Ako uzmemo granicu značajnosti od 0,05 želimo testirati hipotezu da su srednje vrijednosti distribucije stresa jednake.

Tablica 13. Aksijalno naprezanje [psi]

Legura 1		Legura 2	
3238	3254	3261	3248
3195	3229	3187	3215
3246	3225	3209	3226
3190	3217	3212	3240
3204	3241	3258	3234

Nakon rangiranja rezultata dobije se sljedeća tablica 14.

Tablica 14. Rangirani iznosi naprezanja

Legura	Aksijalno naprezanje	Rang
2	3187	1
1	3190	2
1	3195	3
1	3204	4
2	3209	5
2	3212	6
2	3215	7
1	3217	8
1	3225	9
2	3226	10
1	3229	11
2	3234	12
1	3238	13
2	3240	14

1	3241	15
1	3246	16
2	3248	17
1	3254	18
2	3258	19
2	3261	20

Suma rangova za leguru 1 je

$$W_1 = 2 + 3 + 4 + 8 + 9 + 11 + 13 + 15 + 16 + 18 = 99$$

a za leguru 2 je

$$W_2 = \frac{(10 + 10)(10 + 10 + 1)}{2} - 99 = 111$$

Iz tablice F možemo pročitati da je za α = 0,05 i $n_1 = n_2 = 10$ kritična vrijednost W = 78. Ako su W_1 ili W_2 manji ili jednaki od 78 odbaciti ćemo hipotezu H_0. Pošto nisu manji ili jednaki ne možemo je odbaciti stoga ćemo zaključiti da je srednje aksijalno naprezanje jednako u obje legure. Drugim riječima ne postoje podatci koji govore da je jedna legura bolja od druge za tu svrhu.

Ako su n_1 i n_2 veći od 25 tada se distribucija W_1 može približno aproksimirati normalnom distribucijom sa očekivanjem μ i varijancom σ^2 [4]:

$$\mu_{W_1} = \frac{n_1(n_1 + n_2 + 1)}{2}$$

$$\sigma^2_{W_1} = \frac{n_1 n_2(n_1 + n_2 + 1)}{12}$$

Iz toga možemo izračunati z kao novu veličinu testa prema jednadžbi (2.12):

$$z = \frac{W_1 - \mu_{W_1}}{\sigma_{W_1}} \qquad \textbf{(2.12)}$$

Usporedbom izračunate s vrijednošću iz tablice A za određenu granicu značajnosti određujemo, ovisno je li jednostrani ili dvostrani test, treba li odbaciti nultu hipotezu.

2.3.5. Kolmogorov-Smirnov test za dva uzorka

Testira jesu li dva nezavisna uzorka izvučena iz iste populacije (ili populacija s istom distribucijom). Dvostrani je osjetljiv na razlike u lokaciji (centralnoj tendenciji). Jednostrani test se koristi kako bi otkrili jesu li vrijednosti iz jedne populacije stohastički veće od vrijednosti uzorka iz druge populacije [7]. Kao i test za jedan uzorak i ovaj se temelji na sličnosti dvaju kumulativnih frekvencija. Test za jedan uzorak se bavi sličnošću distribucije uzorka i neke teoretske distribucije dok test za dva uzorka otkriva sličnost/razliku između dva uzorka.

Ako su dva uzorka izvučena iz iste populacije onda im kumulativne frekvencije moraju biti približno jednake, a razlike su samo slučajne devijacije. Ako se kumulativne frekevencije dva uzorka previše razlikuju u bilo kojoj točki to nam sugerira da uzorci dolaze iz različitih populacija. Stoga je dovoljno velika devijacija između kumulativnih frekvencija dva uzorka dokaz za odbacivanjem nulte hipoteze.

Neka je *F(X)* kumulativna frekvencija prvog uzorka *n*, a *G(X)* kumulativna frekvencija drugog uzorka *m*. Razliku između veličine frekvencija u određenom intervalu se označava s *D* i računa prema izrazu (2.13) [7]:

$$D = maksimum|F(X) - G(X)| \tag{2.13}$$

Značajnost razlike *D* možemo izračunati pomoću jednadžbe (2.14) [1]:

$$\chi^2 = 4D^2\frac{nm}{n+m} \tag{2.14}$$

Distribucija veličine χ^2 prati hi-kvadrat distribuciju s dva stupnja slobode. Značajnost određujemo usporedbom s kritičnom vrijednosti χ^2 iz tablice B. Nulta hipoteza se odbacuje ako je vjerojatnost pojavljivanja izračunate χ^2 pod nultom hipotezom manja ili jednaka granici značajnosti testa.

Ako su $N = n = m < 40$ tada se značajnost razlike određuje brojnikom K_D pri čemu je K_D brojnik od *D* [1]. K_D se uspoređuje s kritičnim vrijednostima iz tablice G te se na temelju toga odbacuje ili ne odbacuje nulta hipoteza.

Primjer 2.8

Uzmimo dva uzorka, veličine 44 i 54 člana. Svakom članu je mjereno jedno svojstvo i ocjenjivano je na skali od 1 do 20. Interval za kumulativnu frekvenciju neka bude 2. Rezultati mjerenja su dani u tablici 15.

Tablica 15. Rezultati grupirani u intervale

Veličina svojstva	Uzorak 1	Uzorak 2
0-2	11	1
3-5	7	3
6-8	8	6
9-11	3	12
12-14	5	12
15-17	5	14
18-20	5	6
Ukupno	44	54

To znači da 11 članova iz uzorka 1 ima vrijednost mjerenog svojstva od nula do 2 itd. Zatim se podatci pišu u tablicu kumulativne frekvencije na sljedeći način [Tablica 16]:

Tablica 16. Tablica kumulativne frekvencije

	0-2	3-5	6-8	9-11	12-14	15-17	18-20
F(X)	11/44	18/44	26/44	29/44	34/44	39/44	44/44
G(X)	1/54	4/54	10/54	22/54	34/54	48/54	54/54
D	0,232	0,335	0,406	0,252	0,143	0,182	0

Potrebno je pronaći maksimalnu razliku koja iznosi D = 0,406. Pomoću nje izračunamo χ^2

$$\chi^2 = 4 * 0{,}406^2 \frac{44 * 54}{44 + 54} = 15{,}97$$

Pogledom u tablicu B vidimo da je χ^2 veći od kritične vrijednosti za granicu značajnosti od α = 0,05 te odbacujemo nultu hipotezu i zaključujemo da postoji značajna razlika u mjerenom svojstvu između dva uzorka.

2.3.6. Randomizacijski test za dva nezavisna uzorka

Randomizacijski test za dva nezavisna uzorka se koristi kada su varijable u intervalnoj skali i kada je broj članova uzoraka dovoljno mali kako bi se mogao provesti izračun rezultata. Ispituje jesu li srednje vrijednosti dva uzorka značajno različite. Daje točnu vjerojatnost pojave statistike za nultu hipotezu bez pretpostavki o normalnosti i homogenosti varijance.

Provodi se slično kao i randomizacijski test za zavisne uzorke, međutim ovdje uzorci ne moraju biti jednake veličine.

Postupak

1. Odrediti broj mogućih ishoda u području odbacivanja prema jednadžbi (2.15)

$$\alpha \binom{n_1 + n_2}{n_1} \quad \textbf{(2.15)}$$

2. Ekstremni ishodi su oni koji imaju najveću razliku između sume članova uzorka A i sume članova uzorka B. Za jednostrani test svi ishodi su u jednom smjeru, dok su za dvostrani test pola ishoda u jednom, a pola u drugom smjeru.

3. Ako su uočeni ishodi jedni od onih ekstremnih u području odbacivanja, nulta hipoteza se odbacuje uz granicu značajnosti α.

Primjer 2.9 [1]

Neka uzorci A i B imaju sljedeće podatke [Tablica 17]:

Tablica 17. Podatci uzoraka

Uzorak A	0	11	12	20	
Uzorak B	16	19	22	24	29

Kako su $n_1 = 4$ i $n_2 = 5$ broj ishoda u području odbacivanja iznosi:

$$\binom{4+5}{4} = 126$$

Ako je $\alpha = 0{,}05$ područje odbacivanja se sastoji od $\alpha * 126 = 0{,}05 * 126 = 6{,}3$ ekstremna slučaja. U tablici 18 su prikazani takvi ishodi.

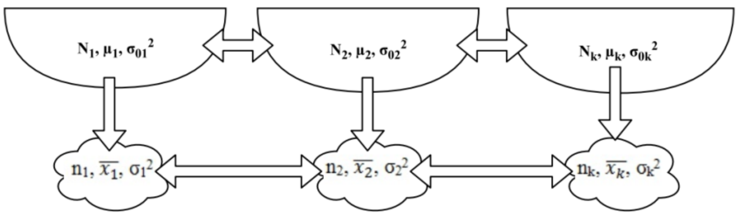
Tablica 18. Ekstremni ishodi

Mogući ishodi za B uzorak					Mogući ishodi za A uzorak				$\sum B$-$\sum A$
19	20	22	24	29	0	11	12	16	75
16	20	22	24	29	0	11	12	19	69
16	19	22	24	29	0	11	12	20	67
16	19	20	24	29	0	11	12	22	63
12	20	22	24	29	0	11	16	19	61
16	19	20	22	29	0	11	12	24	59

Kako su uočeni rezultati jedni od onih koji se nalaze u području odbacivanja, odbacujemo nultu hipotezu za granicu značajnosti α = 0,05.

2.4. Slučaj *k* zavisnih uzoraka

U ovom poglavlju su prikazani testovi za ispitivanje značajnosti razlika između tri ili više uzoraka. Nulta hipoteza glasi da su svi uzorci izvučeni iz iste populacije ili iz identičnih populacija. Na usporedbi *k* uzoraka donosimo zaključke o populacijama kako je prikazano na slici 3.



Slika 3. Shema testiranja k zavisnih uzoraka

Iz populacije veličine N_k uzima se uzorak veličine n_k, pri čemu su parametri populacije:

- μ_k – očekivanje populacije
- σ_{0k}^2 – varijanca populacije

dok su parametri uzorka:

- $\overline{x_k}$ – aritmetička sredina uzorka
- σ_k^2 – varijanca uzorka

Parametarski testovi za testiranje više uzoraka koji dolaze iz identičnih populacija su analiza varijance ili *F*-test. Međutim on pretpostavlja da su uzorci slučajno odabrani iz normalne populacije, da populacije imaju istu varijancu i da se na rezultate mjerenja može primijeniti svojstvo aditivnosti. Uz sve to *F*-test zahtijeva intervalnu skalu mjerenja.

Ako za podatke koje želimo testirati ne možemo prihvatiti navedene pretpostavke i ako nisu u intervalnoj skali koriste se neparametarski testovi.

U slučaju k zavisnih uzoraka, uzorci su jednake veličine i upareni su na način kako je objašnjeno u potpoglavlju 2.2 za dva zavisna uzorka. Dakle jedan objekt može biti izložen trima ili više tretmana kako bi se testirala njihova učinkovitost.

2.4.1. Cochran Q test

Cochran *Q* test je produženi McNemar test jer gdje McNemar test ispituje značajnost razlika između dva seta frekvencija ili proporcija Cochran *Q* testira tri ili više. Jednako tako se koristi za mjerenja u nominalnoj ili dihotomnoj ordinalnoj skali.

Postupak provođenja

1. Za dihotomne podatke dodijelimo nulu jednom svojstvu i jedinicu drugom svojstvu.

2. Podatke upišemo u tablicu veličine k x N, gdje je k broj stupaca (uzoraka), a N broj redaka (broj članova uzorka).

3. Izračuna se vrijednost statistike Q pomoću jednadžbe (2.16) [8]:

$$Q = \frac{(k-1)\left[k\sum_{j=1}^{k} G_j^2 - \left(\sum_{j=1}^{k} G_j\right)^2\right]}{k\sum_{i=1}^{N} L_i - \sum_{i=1}^{N} L_i^2} \tag{2.16}$$

G_j – zbroj jedinica u j-tom stupcu

L_i – zbroj jedinica u i-tom redu

4. Značajnost veličine Q se određuje pomoću tablice B jer je Q približno distribuiran kao hi-kvadrat za stupanj slobode k – 1. Ako je vjerojatnost pojavljivanja veličine jednake ili ekstremnije od Q manja i jednaka od α, tada odbacujemo nultu hipotezu.

Primjer 2.10

Pretpostavimo da imamo tri verzije nekog proizvoda i želimo testirati jesu li jednako privlačni kupcima. Iz populacije kupaca uzimamo uzorak od N = 10 ispitanika i pitamo ih za svaku varijantu proizvoda da li im se sviđa pri čemu su odgovori *da* i *ne*. Podatci su upisani u tablici 19 (jedinica znači *da*, a nula znači *ne*). Granica značajnosti neka je α = 0,05.

Tablica 19. Odgovori prema varijantama

Varijanta			L_i	L_i^2
1	2	3		
0	0	0	0	0
1	1	0	2	4
0	1	0	1	1
0	0	0	0	0
1	0	1	2	4
1	1	0	2	4
1	1	1	3	9
0	1	0	1	1
1	0	0	1	1
0	0	0	0	0
$G_1 = 5$	$G_2 = 5$	$G_3 = 2$	$\sum_{i=1}^{10} L_i = 12$	$\sum_{i=1}^{10} L_i^2 = 24$

Uvrštavanjem podataka u jednadžbu (2.16) dobijemo vrijednost $Q = 3$. Iz tablice B očitana je kritična vrijednost $Q_{\alpha=0,05} = 5{,}99$. Pošto je $Q < Q_{\alpha=0,05}$ ne možemo odbaciti nultu hipotezu, dakle nema razlike u privlačnosti različitih varijanti proizvoda.

2.4.2. Friedmanova analiza varijance s dva promjenjiva faktora

Ovaj test se koristi kada su podatci u ordinalnoj skali, gdje se pod nultom hipotezom testira da li *k* uzoraka dolazi iz iste populacije. Uzorci se upareni stoga je svaki uzorak iste veličine.

Postupak

1. Napisati podatke u tablici s *k* stupaca i *N* redova (parova).

2. Rangirati podatke u svakom redu od 1 do *k*.

3. Zbrojiti rangove u svakom stupcu zasebno: R_j.

4. Izračunati veličinu testa χ_r^2 pomoću jednadžbe (2.17) [9]:

$$\chi_r^2 = \frac{12}{Nk(k+1)} \sum_{j=1}^{k} \left(R_j\right)^2 - 3N(k+1) \qquad \textbf{(2.17)}$$

5. Postoje dvije metode za određivanje kritične vrijednosti χ_r^2:

- Tablica H za točnu vjerojatnost ako vrijednost jednako velika kao χ_r^2 pri k = 3, N = 2 do 9 i k = 4, N = 2 do 4 [1]
- Ako su N i k veći od onih u tablici H, tada se za χ_r^2 koristi aproksimacija hi-kvadrat distribucije s k – 1 stupnjeva slobode, dakle tablica B

6. Ako je vjerojatnost da je vrijednost χ_r^2 manja ili jednaka od vrijednosti χ_r^2 za zadanu granicu značajnosti, odbacujemo nultu hipotezu.

Primjer 2.11

Želimo testirati utjecaj 4 vrste obrade na svojstvo materijala pri čemu ispitujemo samo tri komada toga materijala. Svojstvo materijala se opisuje skalom od 1 do 10 koja može biti ordinatna ili intervalna. Neka je u ovom slučaju ordinatna dakle cjelobrojna varijabla (diskretna). Rezultati su dani u tablici 20.

Tablica 20. Rezultati mjerenja svojstva materijala

	Vrsta obrade			
Komad	I	II	III	IV
1	9	4	1	7
2	6	5	2	8
3	9	1	2	6

Prvo moramo rangirati vrijednosti za svaki red i sumirati rangove za svaki stupac kako je dano u tablici 21.

Tablica 21. Sumirani rangovi stupaca

	Vrsta obrade			
Komad	I	II	III	IV
1	4	2	1	3
2	3	2	1	4
3	4	1	2	3
R_j	11	5	4	10

Ako je nulta hipoteza točna onda očekujemo da će suma rangova svih stupaca biti približno jednaka i da su sve razlike slučajne, što bi značilo da vrsta obrade ne utječe na svojstvo materijala. Izračun prema jednadžbi (2.17) daje rezultat $\chi_r^2 = 7{,}4$. Kako su k i N relativno mali, kritične vrijednosti χ_r^2 ćemo očitati iz tablice H. Vjerojatnost da je $\chi_r^2 \geq 7{,}4$ iznosi 0,033. Prema tome nultu hipotezu možemo odbaciti pri granici značajnosti od 0,033 i reći da vrsta obrade utječe na svojstvo materijala.

2.5. Slučaj *k* nezavisnih uzoraka

Testovi u ovom poglavlju analiziraju podatke iz k nezavisnih uzoraka. Oni ispituju postoji li značajna razlika između tri ili više uzoraka, tj. testira se nulta hipoteza koja glasi da k nezavisnih uzoraka dolazi iz iste populacije ili iz k identičnih populacija. Naime vrijednosti iz svih uzoraka se razlikuju, a cilj se odrediti je li ta razlika značajna ili je posljedica slučajnog uzorkovanja iz iste populacije.

2.5.1. χ^2 test za k nezavisnih uzoraka

Kada su podatci u obliku frekvencija nekih diskretnih kategorija (nominalnih ili ordinarnih) koristi se χ^2 test kako bi se odredilo razlikuju li se k uzorci značajno. Postupak je jednak kao i kod χ^2 testa za dva nezavisna uzorka.

Prvo se dobivene frekvencije napišu u tablici kontingencije veličine r x s, pri čemu r označava broj redaka, a s broj stupaca. Dakle stupanj slobode je $k = (r - 1)(s - 1)$. Zatim se odrede očekivane frekvencije za svako polje tablice. Nakon toga se izračuna χ^2 prema jednadžbi (2.18) i usporedi s vrijednošću iz tablice B za postavljenu granicu značajnosti. Ako je vjerojatnost pojave χ^2 jednaka ili manja od granice značajnosti odbacujemo nultu hipotezu.

$$\chi 2 = \sum_{j=1}^{s} \sum_{i=1}^{r} \frac{\left(f_{ij} - ft_{ij}\right)^2}{ft_{ij}} \qquad \textbf{(2.18)}$$

f_{ij} – stvarne frekvencije

ft_{ij} – očekivane frekvencije

Ograničenja

Test zahtijeva da očekivane frekvencije u svakoj ćeliji tablice ne smiju biti premale. Više od 20% ćelija ne smije imati broj frekvencija manji od 5, a niti jedna ćelija ne smije imati frekvenciju manju od 1. Susjedne kategorije se mogu spojiti ako to ima smisla, kako bi se povećao broj frekvencija u kategoriji.

Primjer 2.12

Uzmimo podatke iz primjera 2.5 i nadodajmo rezultate ispitivanja iz još jednog grada. Upišimo frekvencije u tablicu kontingencije zajedno s očekivanim frekvencijama kako je dano u tablici 22.

Tablica 22. Tablica kontingencije dobivenih frekvencija

	Nikad čuo za proizvod		Čuo, ali nije kupio		Kupio najmanje jedanput		Ukupno
Grad 1	36	*41,5*	55	*58,2*	109	*100,3*	200
Grad 2	45	*31,2*	56	*43,6*	49	*75,2*	150
Grad 3	54	*62,3*	78	*87,2*	168	*150,5*	300
Ukupno	135		189		326		650

Očekivane frekvencije su izračunate metodom parcijalnih suma jednako kao u primjeru za dva nezavisna uzorka. χ^2 prema jednadžbi (2.18) iznosi 100,7. Kako je k = (3 – 1)*(3 – 1) = 4, za α = 0,05, χ^2_{krit} iznosi 9,49. Pošto je $\chi^2 > \chi^2_{\text{krit}}$ odbacujemo nultu hipotezu te možemo zaključiti da je uspješnost na tržištu različita između gradova.

2.5.2. *Ekstenzija medijan testa*

Test ispituje je li k nezavisnih uzoraka izvučeno iz iste populacije ili iz populacija s jednakim medijanom. Koristi se kada su varijable mjerene u najmanje ordinalnoj skali.

Postupak

1. Odredi se medijan rezultata za svaku k-tu grupu

2. Za svaku grupu se nađe broj slučajeva (frekvencija) koji prelaze vrijednost medijana i koji su manji od vrijednosti medijana. Rezultati (frekvencije) se pišu u k x 2 tablicu.

3. Koristeći podatke iz te tablice, izračuna se veličina testa χ^2 , koja je aproksimirana hi-kvadrat distribucijom za k – 1 stupanj slobode. Izračun χ^2 je da u jednadžbi (2.19)

$$\chi 2 = \sum_{j=1}^{s} \sum_{i=1}^{r} \frac{\left(f_{ij} - ft_{ij}\right)^2}{ft_{ij}} \quad \textbf{(2.19)}$$

f_{ij} – stvarne frekvencije

ft_{ij} – očekivane frekvencije (vrijednost medijana)

4. Nakon svega treba odrediti značajnost za χ^2 pomoću tablice B. Ako je vjerojatnost da je vrijednost χ^2 manja ili jednaka od granice značajnosti, odbacujemo nultu hipotezu.

2.5.3. Kruskal-Wallis analiza varijance s jednim promjenjivim faktorom

Kruskal-Wallis test je posebno koristan za ispitati je li broj od k nezavisnih uzoraka iz različitih populacija. Drugim riječima test ispituje nultu hipotezu koja tvrdi da broj od k uzoraka dolazi iz iste populacije ili iz identičnih populacija s obzirom na srednju vrijednost. Pretpostavka je da su varijable kontinuirano distribuirane i da su varijable najmanje u ordinalnoj skali.

Postupak

1. Članovi svih k uzoraka se rangiraju u jednu seriju tako da im se dodijele rangovi od 1 do N.

2. Za svaki k-ti uzorak se izračuna suma rangova R_j.

3. Veličina testa H se računa prema jednadžbi (2.20) [9]:

$$H = \frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} - 3(N+1) \tag{2.20}$$

pri čemu je

n_j – broj članova u j-tom stupcu (uzorku)

$N = \sum n_j$, ukupni broj članova svih uzoraka

R_j – suma rangova u j-tom stupcu

4. Način za procjenu značajnosti izračunate vrijednosti H ovisi o broju uzoraka k i veličini svakog uzorka:

   a. ako je $k = 3$ i $n_1, n_2, n_3 \leq 5$ koristi se tablica I
   b. u ostalim slučajevima značajna vrijednost se očitava iz tablice B

5. Ako je vjerojatnost da je izračunata vrijednost H jednaka ili manja od granice značajnosti α, nulta hipoteza se odbacuje.

Primjer 2.13

Mjeri se tlačna čvrstoća betona za četiri različite tehnike miješanja betona. Testirajmo hipotezu da tehnika miješanja utječe na tlačnu čvrstoću betona nasuprot hipotezi da ne utječe, tj. da je srednja vrijednost čvrstoće jednaka za sve četiri vrste betona. Podatci su dani u tablici 23. Iako je varijabla kontinuirana te lako moguće i normalno distribuirana ispitivanje se može vrlo lako rješiti ovim neparametarskim testom bez provjere i/ili pretpostavke o normalnosti podataka.

Tablica 23. Tlačna čvrstoća betona s obzirom na tehniku miješanja [4]

Tehnika miješanja	Tlačna čvrstoća [psi]			
1	3129	3000	2865	2890
2	3200	3300	2975	3150
3	2800	2900	2985	3050
4	2600	2700	2600	2765

Članovi svih *k* uzoraka se rangiraju u jednu seriju te se za svaki *k*-ti uzorak izračuna suma rangova R_j kako je prikazano u tablici 24.

Tablica 24. Suma rangova uzoraka

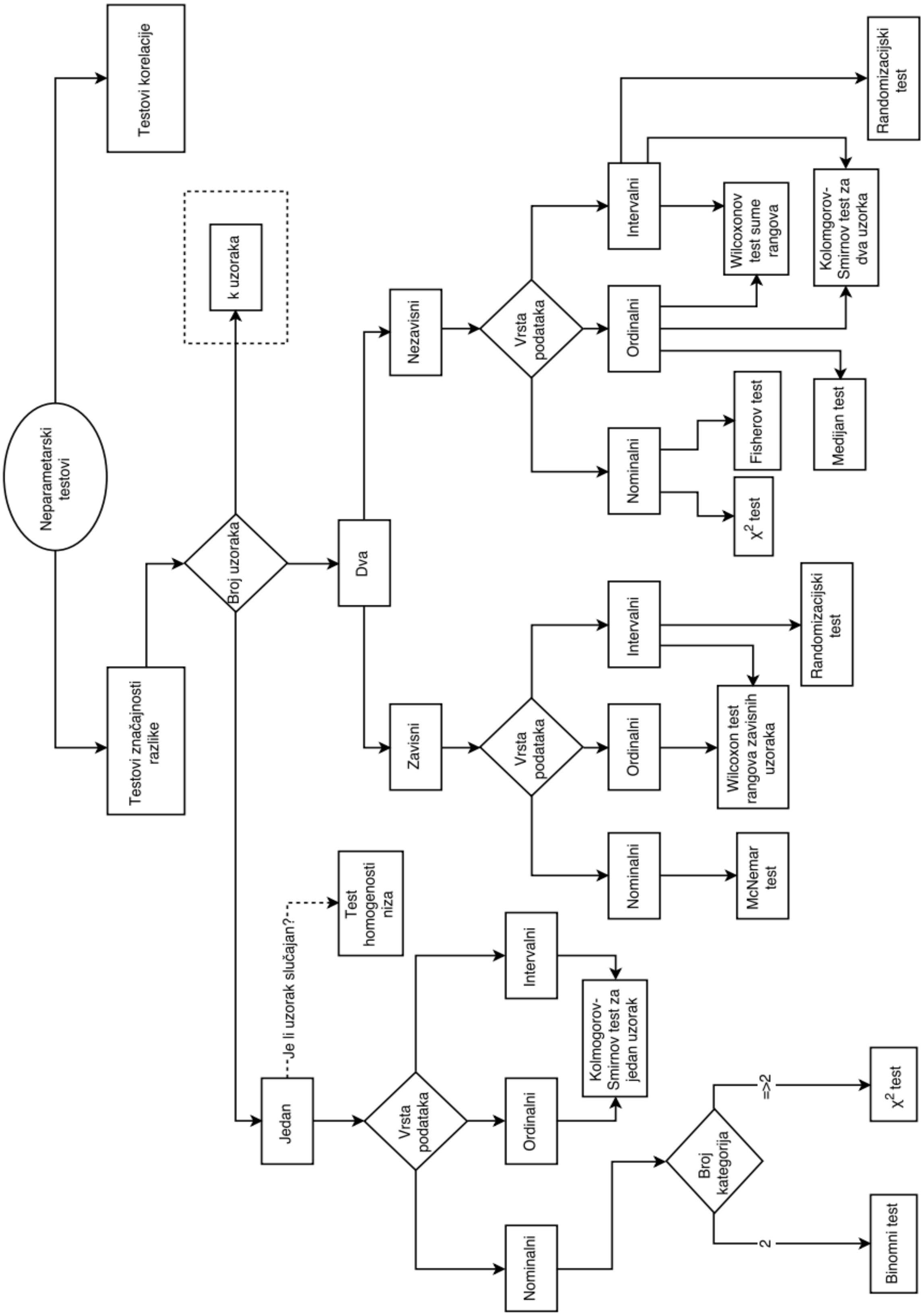
n_j	Rangovi uzoraka				$\sum R_j$
1	13	11	6	7	34,5
2	15	16	9	14	38
3	5	8	10	12	26,5
4	1,5	3	1,5	4	37

Nakon toga se izračuna vrijednost testa H prema jednadžbi (2.20), $H = 0,89$. Kako je $k > 3$ kritičnu vrijednost iščitavamo iz tablice B za $k - 1 = 3$ stupnja slobode i granicu značajnosti od 0,05 dakle $H_{krit} = 7,82$. Pošto je $H < H_{krit}$ ne možemo odbacit nultu hipotezu i zaključujemo da tehnike miješanja ne utječu na tlačnu čvrstoću betona.

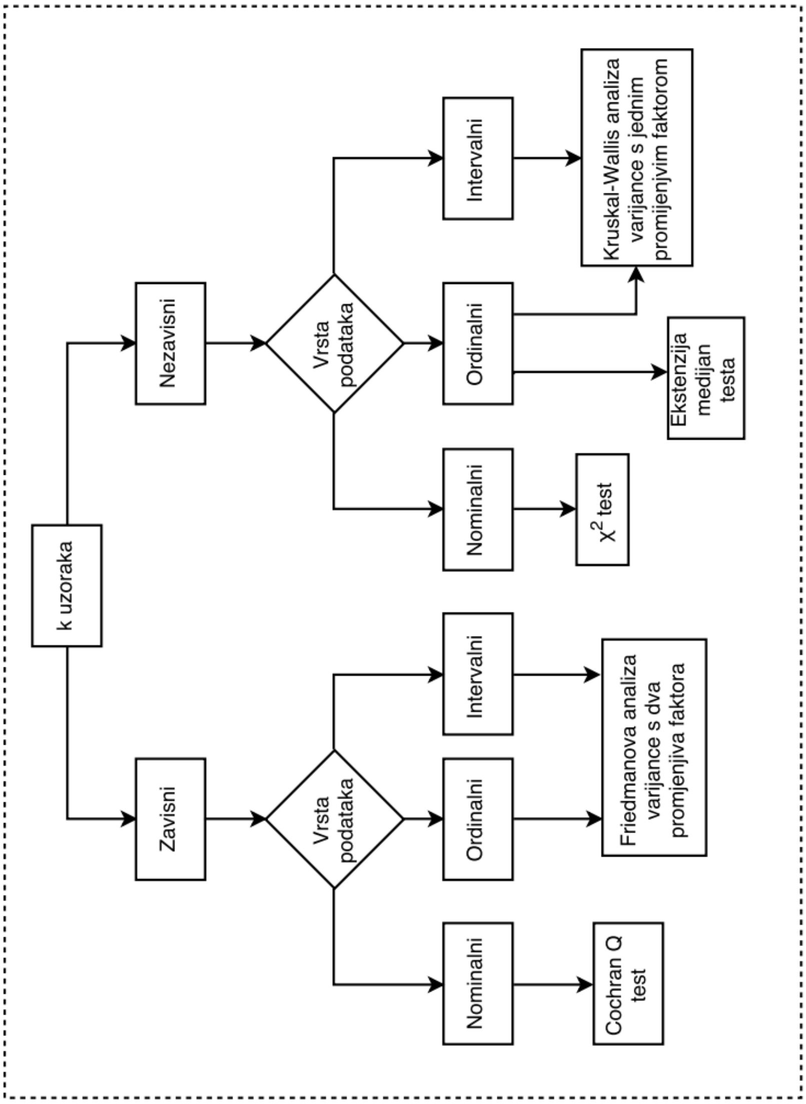
3. HODOGRAM ODABIRA METODE NEPARAMETARSKE STATISTIKE

Od svih navedenih neparametarskih statističkih testova opisanim u prethodnim poglavljima potrebno je izabrati onaj koji odgovara svrsi testa i vrsti podataka. Kako bi se jednostavno moglo utvrditi koji je test prigodan i primjenjiv izrađen je hodogram njihovog odabira koji je prikazan na slici 4. Neparametarski testovi se dijele na testove značajnosti razlike, koji testiraju jesu li uzorci iz iste populacije, i testove korelacije, koji ispituju postoji li povezanost između varijabli uzoraka.

Ako se želi testirati postoji li značajna razlika između uzoraka postupak ide kako slijedi. Prvo treba odrediti broj uzoraka. Ako imamo dva ili više uzoraka potrebno je uočiti jesu li zavisni ili nezavisni. Nakon toga se određuje vrsta podatka, jesu li u nominalnoj, ordinalnoj ili intervalnoj skali. Posljednji korak jest odabrati test, međutim ako dva ili više testova zadovoljavaju, može se pogledati njihov opis u prethodnim poglavljima za detaljnije upute o tome koji je adekvatniji, s obzirom na ono što se želi ispitati i veličinu uzorka.



Slika 4. Hodogram



Slika 4. Hodogram *(nastavak)*

4. RAZVIJENI OBRASCI NEPARAMETARSKIH TESTOVA U PROGRAMSKOM RJEŠENJU EXCEL

Kako bi se testovi mogli brzo provesti bez korištenja skupih i kompliciranih statističkih alata napravljeni su obrasci po kojemu se neki od navedenih testova rješavaju u Microsoft Excelu, koji je relativno pristupačan programski paket.

Prikaz obrasca za izračun Wilcoxonovog testa rangova zavisnih uzoraka je na slici 5.

1												
2	N	Uzorak 1	Uzorak 2	Razlika	Apsolutna razlika	Rangovi	Pozitivni rangovi	Negativni rangovi				
3	1	17	18	-1	1	2		2	H_0: $\mu_1 = \mu_2$			
4	2	16	14	2	2	6	6		H_1: $\mu_1 \neq \mu_2$	$\mu_1 < \mu_2$	$\mu_1 > \mu_2$	
5	3	21	19	2	2	6	6			N<=25	N>25	
6	4	14	11	3	3	9,5	9,5		alfa	0,05	vrsta testa:	
7	5	18	23	-5	5	11		11	n	12	jednostrani	
8	6	24	21	3	3	9,5	9,5		T-krit	14	z	-0,94136
9	7	16	10	6	6	12	12		T	27	p	0,17326
10	8	14	13	1	1	2	2		ne odbacujemo H0			
11	9	21	19	2	2	6	6					
12	10	23	24	-1	1	2		2				
13	11	13	15	-2	2	6		6				
14	12	18	20	-2	2	6		6				
15												
16												

Slika 5. Wilcoxon test rangova zavisnih uzoraka

Narančastom bojom su označena polja tablice u kojoj je dopušten unos podataka. Prvo je potrebno unijeti podatke iz uzorka, a zatim ovisno o veličini uzorka je potrebno upisati granicu značajnosti i kritičnu vrijednost iz tablice ili ako je N > 25 samo vrstu testa (da li je jednostrani ili dvostrani). Nakon toga se dobiva rezultat, koji ako se nulta hipoteza pokaže netočnom, prikazuje kao: „odbacujemo H0“. U suprotnom slučaju „ne odbacujemo H0“.

Jednako kao i za prethodni test napravljen je obrazac za izračun Wilocoxonovog testa sume rangova za primjer 2.7. U tablicu, prikazanu na slici 6, se mogu unijeti uzorci veličine do N = 50.

	N	Uzorak 1	Uzorak 2	Rang 1	Rang 2			
1								
2	*N*	Uzorak 1	Uzorak 2	*Rang 1*	*Rang 2*			
3	*1*	3238	3261	*13*	*20*	H_0: $\mu_1 = \mu_2$		
4	*2*	3195	3187	*3*	*1*	H_1: $\mu_1 \neq \mu_2$	$\mu_1 < \mu_2$	$\mu_1 > \mu_2$
5	*3*	3246	3209	*16*	*5*			
6	*4*	3190	3212	*2*	*6*			
7	*5*	3204	3258	*4*	*19*		alfa	0,05
8	*6*	3254	3248	*18*	*17*		Wkrit	78
9	*7*	3229	3215	*11*	*7*		W	99
10	*8*	3225	3226	*9*	*10*		ne odbacujemo H0	
11	*9*	3217	3240	*8*	*14*			
12	*10*	3241	3234	*15*	*12*			
13								
14								
15								

Slika 6. Wilcoxon test sume rangova

Za izračun rangova je korištena funkcija RANK.AVG koja uzima u obzir izjednačene vrijednosti te im dodijeljuje srednju vrijednost ranga. Upisom kritične vrijednosti W_{krit} očitane iz adekvatne tablice dobiva se rezultat, koji u ovom slučaju potvrđuje da nema značajnih razlika između uzoraka.

U primjeru 2.8 za Kolmogorov-Smirnov test za dva uzorka u tablici 15 odmah su prikazani grupirani podatci po intervalima, tj. frekvencija rezultata po intervalima. Međutim u obrascu, dostupnom u prilogu, je potrebno samo upisati podatke i definirati intervale, a funkcijom FREQUENCY dolazi do automatskog izračuna kumulativne frekvencije. Postoje tri ćelije za prikaz rezultata. Ako je $N > 40$ vrši se provjera za jednostrani i dvostrani test te se izbacuju rezultati, pri čemu će rezultat koji vrijedi za $N \leq 40$ ostati sakriven kako ne bi došlo do zabune jer za takav slučaj se ne koristi testni parametar χ^2. Prikaz radnog lista na kojemu se nalazi ovaj test je na slici 7.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Uzorak 1	Uzorak 2	Intervali	*Kumul.f.1*	*Kumul.f.2*	*Sn1(x)*	*Sn2(x)*	*KD*	*Razlika*	H_0: $\mu_1 = \mu_2$					
2	1	1	2	*11*	*1*	*0,25*	*0,018519*	*10*	*0,231481*	H_1: $\mu_1 \neq \mu_2$	$\mu_1 < \mu_2$	$\mu_1 > \mu_2$			
3	1	3	5	*18*	*4*	*0,409091*	*0,074074*	*14*	*0,335017*		N>40		N>40	N<=40	
4	1	4	8	*26*	*10*	*0,590909*	*0,185185*	*16*	*0,405724*		jednostra		dvostrani		
5	1	5	11	*29*	*22*	*0,659091*	*0,407407*	*7*	*0,251684*	alfa	0,05	alfa	0,05	alfa	0,05
6	2	6	14	*34*	*34*	*0,772727*	*0,62963*	*0*	*0,143098*	χ2 krit	5,991465	koef	1,36	Kdkrit	
7	2	7	17	*39*	*48*	*0,886364*	*0,888889*	*9*	*0,002525*	χ2	15,96399	Dkrit	0,276203	Kd	16
8	2	8	20	*44*	*54*	*1*	*1*	*10*	*0*	odbacujemo H0		D	0,405724		
9	1	6		*0*	*0*	*0*	*0*	*0*	*0*			odbacujemo H0			
10	1	7		*0*	*0*	*0*	*0*	*0*	*0*						
11	2	8		*0*	*0*	*0*	*0*	*0*	*0*			alfa	koef		
12	1	9	*N*	*44*	*54*				*0,405724*			0,1	1,22		
13	3	10	*54*									0,05	1,36		
14	4	11										0,025	1,48		
15	5	9										0,01	1,63		
16	3	10										0,005	1,73		
17	4	11										0,001	1,95		
18	5	9													

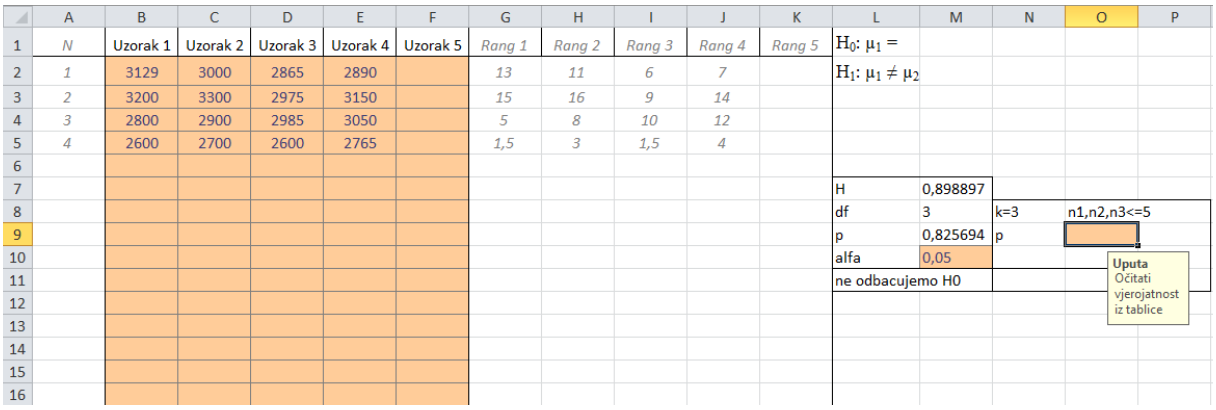
Slika 7. Kolmogorov-Smirnov test za dva uzorka

Radni list na kojemu se nalazi obrazac [Slika 8.] za rješavanje Friedmanovog testa omogućava testiranje do 6 uzoraka iako je to lako proširivo umetanjem stupaca i modificiranjem funkcije RANK.AVG. Potrebno je samo upisati podatke iz uzorka i kritičnu vrijednost testnog parametra za granicu značajnosti.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1															
2	*N*	Uzorak 1	Uzorak 2	Uzorak 3	Uzorak 4	Uzorak 5	Uzorak 6	*Rangovi uzorka 1*	*Rangovi uzorka 2*	*Rangovi uzorka 3*	*Rangovi uzorka 4*	*Rangovi uzorka 5*	*Rangovi uzorka 6*	H_0: $\mu_1 = \mu_2$	
3	*1*	9	4	1	7			*4*	*2*	*1*	*3*			H_1: $\mu_1 \neq \mu_2$	
4	*2*	6	5	2	8			*3*	*2*	*1*	*4*				
5	*3*	9	1	2	6			*4*	*1*	*2*	*3*				
6															
7															
8														N	3
9														k	4
10														χ2	7,400
11														α	0,05
12														χ2krit	7,4
13														odbacujemo H0	
14															

Slika 8. Friedmanova analiza varijance s dva promjenjiva faktora

Izračun Kruskal-Wallis testa je pojednostavljen pomoću Microsoft Excela. Kao i do sada ćelije su programirane tako, da jednostavnim upisom podatka u stupce uzoraka, se dobije rezultat. Izračun uzima u obzir veličinu uzoraka pri čemu je samo potrebno upisati vjerojatnost iz tablice I, ako su uzorci dovoljno mali. Ilustracija obrasca je na slici 9.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	N	Uzorak 1	Uzorak 2	Uzorak 3	Uzorak 4	Uzorak 5	Rang 1	Rang 2	Rang 3	Rang 4	Rang 5	H_0: μ_1 =				
2	1	3129	3000	2865	2890		13	11	6	7		H_1: $\mu_1 \neq \mu_2$				
3	2	3200	3300	2975	3150		15	16	9	14						
4	3	2800	2900	2985	3050		5	8	10	12						
5	4	2600	2700	2600	2765		1,5	3	1,5	4						
6																
7												H	0,898897			
8												df	3	k=3	n1,n2,n3<=5	
9												p	0,825694	p		
10												alfa	0,05			
11												ne odbacujemo H0				
12																
13																
14																
15																
16																

Slika 9. Kruskal-Wallis analiza varijance s jednim promjenjivim faktorom

5. ZAKLJUČAK

Iz svega priloženoga može se dobiti uvid u najvažnije neparametarske statističke testove. Opisani su njihovi postupci, mogućnosti, svrha korištenja i primjeri iz inženjerske prakse. Njihova potreba se očituje u manjem broju pretpostavki i ograničenja koje se moraju zadovoljiti, kako bi rezultati testa bili smisleni, u odnosu na parametarske testove. Upotreba neparametarskih testova je potrebna kada podatci iz uzoraka nisu u intervalnoj skali, što je jedan od uvjeta parametarskih testova. Ako i jesu uzorci moraju dolaziti iz populacije s normalnom distribucijom te se u slučaju velikog odstupanja od normalnosti parametarski testovi ne mogu primijeniti. Ako se pretpostavke o normalnosti ne mogu ili ne žele provjeriti i prihvatiti mogu se primijeniti neparametarski testovi koji ne rade takve pretpostavke. Njihova prednost jest u tome što mogu raditi i s podatcima koji su u nominalnoj ili ordinalnoj skali. Takvi podatci zahtijevaju jednostavnija mjerenja i lakše se i brže prikupljaju. Testovi koji rade s ordinalnim podatcima mogu se primijeniti i na intervalnim podatcima.

U slučaju jednog uzorka, gdje se testira dolazi li uzorak iz neke točno određene populacije, mogu se primijeniti četiri testa. Za nominalne podatke, iako se neki mogu primijeniti i na ordinalne, pogodan je binomni test ako je broj kategorija dva i ako su uzorci jako mali te χ^2 ako je broj kategorija dva ili više. Ako varijable dolaze iz kontinuirane populacije najbolje je upotrijebiti Kolmogorov-Smirnov test za jedan uzorak. Test homogenosti niza se razlikuje od ostalih jer ispituje jesu li uzorci odabrani slučajno.

U slučaju dva zavisna uzorka predstavljena su tri testa. McNemar test se koristi isključivo za podatke koji su u frekvencijama diskretnih kategorija. Za ordinalna mjerenja, točnije rečeno ako se podatci mogu rangirati i ako je zadovoljena pretpostavka o kontinuiranoj distribuciji varijable, može se primijeniti Wilcoxon test rangova zavisnih uzoraka kao dobra zamjena *t*-testu. Ako su postignuta intervalna mjerenja i ako je broj uzoraka dovoljno malen koristi se randomizacijski test.

U slučaju dva nezavisna uzorka kod nominalnih podataka se preporuča χ^2, a ako je broj frekvencija manji od 5 Fisherov test. Za dihotomne podatke je pogodan medijan test, a za podatke koji se mogu rangirati ili su intervalnoj skali Wilcoxon test sume rangova te Kolmogorov-Smirnov test za dva uzorka kao dobra zamjena *t*-testu. Ako su uzorci mali i mjerenja su u intervalnoj skali preporuča se randomizacijski test za dva uzorka.

U slučaju tri ili više zavisnih uzoraka za testiranje općenite različitosti frekvencija koristi se Cochranov *Q* test. Ako su podatci u najmanje ordinalnoj skali dostupan je Friedmanov test analize varijance s dva promjenjiva faktora.

U slučaju tri ili više nezavisnih uzoraka, kod nominalnih ili ordinalnih podataka diskretnih kategorija se koristi χ^2 test. Ekstenzija medijan testa za najmanje ordinalne mjere ispituje da li *k* nezavisnih uzoraka dolazi iz populacije s istim medijanom. Međutim Kruskal-Wallis analiza varijance s jednim promjenjivim faktorom je najbolje rješnje jer koristi više informacija iz podataka.

Kako bi se korištenje ovih testova moglo vršiti bez uporabe specijalnih statističkih programa u Microsoft Excelu su napravljeni obrasci za izračun slijedećih testova: Wilcoxon test rangova zavisnih uzoraka i test sume rangova, Kolmogorov-Smirnov test za dva uzorka, Friedman i Kruskal-Wallis analize varijance.

LITERATURA

[1] Siegel S.: Nonparametric Statistics for the Behavioral Sciences, McGraw-Hill Book Company, 1956.

[2] Cajner H.: Osnove teorije uzoraka, skripte s predavanja, 2013.

[3] http://www.real-statistics.com/non-parametric-tests/one-sample-kolmogorov-smirnov-test/

[4] Montgomery D.C., Runger G.C.: Applied Statistics and Probability for Engineers, John Wiley & Sons, 2010.

[5] Bhattacharyya G.K., Johnson R.A.: Statistical Concepts and Methods, John Wiley & Sons, 2005.

[6] Cajner H.: Raspodjele podataka, skripte s predavanja, 2013.

[7] http://www.real-statistics.com/non-parametric-tests/two-sample-kolmogorov-smirnov-test/

[8] http://www.real-statistics.com/anova-repeated-measures/cochrans-q-test/

[9] Fahoome G. F.: Twenty Nonparametric Statistics And Their Large Sample Approximations. http://digitalcommons.wayne.edu/jmasm/vol1/iss2/35

[10] Cajner H.: Osnove statistike, skripte s predavanja, 2013.

[11] Koceić Bilan N., Primijenjena statistika, Prirodoslovno-matematički fakultet u Splitu, 2011.

[12] Šnajder J., Dalbelo Bašić B.: Statističko testiranje klasifikatora, bilješke s predavanja, 2015.

[13] http://www.real-statistics.com/statistics-tables/kolmogorov-smirnov-table/

[14] http://www.dzs.hr/app/rss/rjecnik_hr-en-EN.html

[15] NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/, 2012

[16] Sarapa N.: Vjerojatnost i statistika, Školska knjiga, 1993.

[17] Benšić M., Šuvak N.: Primijenjena statistika, Sveučilište J.J. Strossmayera, 2013.

[18] https://graphpad.com/guides/prism/7/statistics/index.htm?statistical_hypothesis_testing.htm

[19] Horvat J., Mijoč J.: Osnove statistike, Naklada Ljevak, 2012.

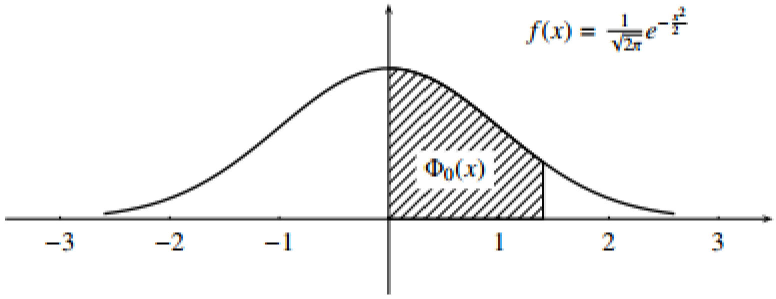
[20] https://www.uvm.edu/~dhowell/StatPages/Randomization%20Tests/RandomizationTestsOverview.html

PRILOG

I. DVD – R disk

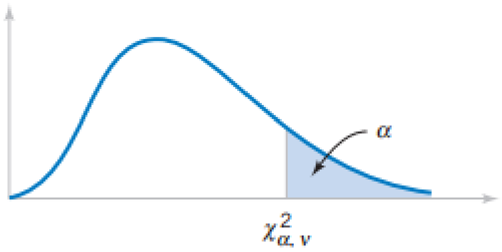
II. Statističke tablice

Tablica A Normalna jedinična razdioba (svakoj vrijednosti u tabeli prethodi decimalni zarez)



z	0	1	2	3	4	5	6	7	8	9
0.	00000	00399	00798	01197	01595	01994	02392	02790	03188	03586
0.1	03983	04380	04776	05172	05567	05962	06356	06749	07142	07535
0.2	07926	08317	08706	09095	09483	09871	10257	10642	11026	11409
0.3	11791	12172	12552	12930	13307	13683	14058	14431	14803	15173
0.4	15542	15909	16275	16640	17003	17364	17724	18082	18438	18793
0.5	19146	19497	19846	20194	20540	20884	21226	21566	21904	22240
0.6	22574	22906	23237	23565	23891	24215	24537	24857	25174	25490
0.7	25803	26114	26423	26730	27035	27337	27637	27935	28230	28523
0.8	28814	29103	29389	29673	29954	30233	30510	30785	31057	31326
0.9	31594	31858	32121	32381	32639	32894	33147	33397	33645	33891
1.	34134	34375	34613	34849	35083	35314	35542	35769	35992	36214
1.1	36433	36650	36864	37076	37285	37492	37697	37900	38100	38297
1.2	38493	38686	38876	39065	39251	39435	39616	39795	39972	40147
1.3	40320	40490	40658	40824	40987	41149	41308	41465	41620	41773
1.4	41924	42073	42219	42364	42506	42647	42785	42921	43056	43188
1.5	43319	43447	43574	43699	43822	43942	44062	44179	44294	44408
1.6	44520	44630	44738	44844	44949	45052	45154	45254	45352	45448
1.7	45543	45636	45728	45818	45907	45994	46079	46163	46246	46327
1.8	46407	46485	46562	46637	46711	46784	46855	46925	46994	47062
1.9	47128	47193	47257	47319	47381	47441	47500	47558	47614	47670
2.	47725	47778	47830	47882	47932	47981	48030	48077	48123	48169
2.1	48213	48257	48299	48341	48382	48422	48461	48499	48537	48573
2.2	48609	48644	48679	48712	48745	48777	48808	48839	48869	48898
2.3	48927	48955	48983	49009	49035	49061	49086	49110	49134	49157
2.4	49180	49202	49224	49245	49265	49285	49305	49324	49343	49361
2.5	49379	49396	49413	49429	49445	49461	49476	49491	49506	49520
2.6	49533	49547	49560	49573	49585	49597	49609	49620	49631	49642
2.7	49653	49663	49673	49683	49692	49702	49711	49719	49728	49736
2.8	49744	49752	49759	49767	49774	49781	49788	49794	49801	49807
2.9	49813	49819	49825	49830	49835	49841	49846	49851	49855	49860
3.	49865	49869	49873	49877	49881	49885	49889	49893	49896	49899
3.1	49903	49906	49909	49912	49915	49918	49921	49923	49926	49928
3.2	49931	49933	49935	49938	49940	49942	49944	49946	49948	49949
3.3	49951	49953	49955	49956	49958	49959	49961	49962	49963	49965
3.4	49966	49967	49968	49969	49970	49972	49973	49974	49974	49975
3.5	49976	49977	49978	49979	49980	49980	49981	49982	49982	49983
4.	49996	49997	49997	49997	49997	49997	49997	49997	49997	49997
4.5	49999	49999	49999	49999	49999	49999	49999	49999	49999	49999

Tablica B Hi-kvadrat razdioba [4]



ν \ α	.995	.990	.975	.950	.900	.500	.100	.050	.025	.010	.005
1	.00+	.00+	.00+	.00+	.02	.45	2.71	3.84	5.02	6.63	7.88
2	.01	.02	.05	.10	.21	1.39	4.61	5.99	7.38	9.21	10.60
3	.07	.11	.22	.35	.58	2.37	6.25	7.81	9.35	11.34	12.84
4	.21	.30	.48	.71	1.06	3.36	7.78	9.49	11.14	13.28	14.86
5	.41	.55	.83	1.15	1.61	4.35	9.24	11.07	12.83	15.09	16.75
6	.68	.87	1.24	1.64	2.20	5.35	10.65	12.59	14.45	16.81	18.55
7	.99	1.24	1.69	2.17	2.83	6.35	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	7.34	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	8.34	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	9.34	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	10.34	17.28	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	6.30	11.34	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	12.34	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	13.34	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.27	7.26	8.55	14.34	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	15.34	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	16.34	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.87	17.34	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	18.34	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	19.34	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	20.34	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	21.34	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	22.34	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	23.34	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	24.34	34.28	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	25.34	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	26.34	36.74	40.11	43.19	46.96	49.65
28	12.46	13.57	15.31	16.93	18.94	27.34	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	28.34	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	29.34	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	39.34	51.81	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	49.33	63.17	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	46.46	59.33	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	69.33	85.53	90.53	95.02	100.42	104.22
80	51.17	53.54	57.15	60.39	64.28	79.33	96.58	101.88	106.63	112.33	116.32
90	59.20	61.75	65.65	69.13	73.29	89.33	107.57	113.14	118.14	124.12	128.30
100	67.33	70.06	74.22	77.93	82.36	99.33	118.50	124.34	129.56	135.81	140.17

Tablica C Kolmogorov-Smirnov za jedan uzorak (kritične vrijednosti statistike D) [13]

n \ α	0.01	0.05	0.1	0.15	0.2
1	0.995	0.975	0.950	0.925	0.900
2	0.929	0.842	0.776	0.726	0.684
3	0.828	0.708	0.642	0.597	0.565
4	0.733	0.624	0.564	0.525	0.494
5	0.669	0.565	0.510	0.474	0.446
6	0.618	0.521	0.470	0.436	0.410
7	0.577	0.486	0.438	0.405	0.381
8	0.543	0.457	0.411	0.381	0.358
9	0.514	0.432	0.388	0.360	0.339
10	0.490	0.410	0.368	0.342	0.322
11	0.468	0.391	0.352	0.326	0.307
12	0.450	0.375	0.338	0.313	0.295
13	0.433	0.361	0.325	0.302	0.284
14	0.418	0.349	0.314	0.292	0.274
15	0.404	0.338	0.304	0.283	0.266
16	0.392	0.328	0.295	0.274	0.258
17	0.381	0.318	0.286	0.266	0.250
18	0.371	0.309	0.278	0.259	0.244
19	0.363	0.301	0.272	0.252	0.237
20	0.356	0.294	0.264	0.246	0.231
25	0.320	0.270	0.240	0.220	0.210
30	0.290	0.240	0.220	0.200	0.190
35	0.270	0.230	0.210	0.190	0.180
40	0.250	0.210	0.190	0.180	0.170
45	0.240	0.200	0.180	0.170	0.160
50	0.230	0.190	0.170	0.160	0.150
OVER 50	$\frac{1.63}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.14}{\sqrt{n}}$	$\frac{1.07}{\sqrt{n}}$

Tablica D Kritične vrijednosti za *r*

Critical values of *r* in the runs test*

Given in the tables are various critical values of *r* for values of *m* and *n* less than or equal to 20. For the one-sample runs test, any observed value of *r* which is less than or equal to the smaller value, or is greater than or equal to the larger value in a pair is significant at the $\alpha = .05$ level.

m \ n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2											2 –	2 –	2 –	2 –	2 –	2 –	2 –	2 –	2 –
3					2 –	2 –	2 –	2 –	2 –	2 –	2 –	2 –	2 –	3 –	3 –	3 –	3 –	3 –	3 –
4				2 9	2 9	2 –	3 –	3 –	3 –	3 –	3 –	3 –	3 –	3 –	4 –	4 –	4 –	4 –	4 –
5			2 9	2 10	3 10	3 11	3 11	3 –	3 –	4 –	4 –	4 –	4 –	4 –	4 –	4 –	5 –	5 –	5 –
6		2 –	2 9	3 10	3 11	3 12	3 12	4 13	4 13	4 13	4 13	5 –	5 –	5 –	5 –	5 –	5 –	6 –	6 –
7		2 –	2 –	3 11	3 12	3 13	4 13	4 14	5 14	5 14	5 14	5 15	5 15	6 15	6 –	6 –	6 –	6 –	6 –
8		2 –	3 –	3 11	3 12	4 13	4 14	5 14	5 15	5 15	6 16	6 16	6 16	6 16	6 17	7 17	7 17	7 17	7 17
9		2 –	3 –	3 –	4 13	4 14	5 14	5 15	5 16	6 16	6 16	6 17	7 17	7 18	7 18	7 18	8 18	8 18	8 18
10		2 –	3 –	3 –	4 13	5 14	5 15	5 16	6 16	6 17	7 17	7 18	7 18	7 18	8 19	8 19	8 19	8 20	9 20
11		2 –	3 –	4 –	4 13	5 14	5 15	6 16	6 17	7 17	7 18	7 19	8 19	8 19	8 20	9 20	9 20	9 21	9 21
12	2 –	2 –	3 –	4 –	4 13	5 14	6 16	6 16	7 17	7 18	7 19	8 19	8 20	8 20	9 21	9 21	9 21	10 22	10 22
13	2 –	2 –	3 –	4 –	5 –	5 15	6 16	6 17	7 18	7 19	8 19	8 20	9 20	9 21	9 21	10 22	10 22	10 23	10 23
14	2 –	2 –	3 –	4 –	5 –	5 15	6 16	7 17	7 18	8 19	8 20	9 20	9 21	9 22	10 22	10 23	10 23	11 23	11 24
15	2 –	3 –	3 –	4 –	5 –	6 15	6 16	7 18	7 18	8 19	8 20	9 21	9 22	10 22	10 23	11 23	11 24	11 24	12 25
16	2 –	3 –	4 –	4 –	5 –	6 –	6 17	7 18	8 19	8 20	9 21	9 21	10 22	10 23	11 23	11 24	11 25	12 25	12 25
17	2 –	3 –	4 –	4 –	5 –	6 –	7 17	7 18	8 19	9 20	9 21	10 22	10 23	11 23	11 24	11 25	12 25	12 26	13 26
18	2 –	3 –	4 –	5 –	5 –	6 –	7 17	8 18	8 19	9 20	9 21	10 22	10 23	11 24	11 25	12 25	12 26	13 26	13 27
19	2 –	3 –	4 –	5 –	6 –	6 –	7 17	8 18	8 20	9 21	10 22	10 23	11 23	11 24	12 25	12 26	13 26	13 27	13 27
20	2 –	3 –	4 –	5 –	6 –	6 –	7 17	8 18	9 20	9 21	10 22	10 23	11 24	12 25	12 25	13 26	13 27	13 27	14 28

Tablica E Wilcoxon test rangova zavisnih uzoraka (kritične vrijednosti za T) [4]

α / n^*	0.10 0.05	0.05 0.025	0.02 0.01	0.01 0.005	Two-sided tests One-sided tests
4					
5	0				
6	2	0			
7	3	2	0		
8	5	3	1	0	
9	8	5	3	1	
10	10	8	5	3	
11	13	10	7	5	
12	17	13	9	7	
13	21	17	12	9	
14	25	21	15	12	
15	30	25	19	15	
16	35	29	23	19	
17	41	34	27	23	
18	47	40	32	27	
19	53	46	37	32	
20	60	52	43	37	
21	67	58	49	42	
22	75	65	55	48	
23	83	73	62	54	
24	91	81	69	61	
25	100	89	76	68	

Tablica F Wilcoxon test sume rangova (kritične vrijednosti za W pri α = 0,01, dvostrani) [4]

$w_{0.01}$

n_1 / n_2	4	5	6	7	8	9	10	11	12	13	14	15
5		15										
6	10	16	23									
7	10	17	24	32								
8	11	17	25	34	43							
9	11	18	26	35	45	56						
10	12	19	27	37	47	58	71					
11	12	20	28	38	49	61	74	87				
12	13	21	30	40	51	63	76	90	106			
13	14	22	31	41	53	65	79	93	109	125		
14	14	22	32	43	54	67	81	96	112	129	147	
15	15	23	33	44	56	70	84	99	115	133	151	171
16	15	24	34	46	58	72	86	102	119	137	155	
17	16	25	36	47	60	74	89	105	122	140		
18	16	26	37	49	62	76	92	108	125			
19	17	27	38	50	64	78	94	111				
20	18	28	39	52	66	81	97					
21	18	29	40	53	68	83						
22	19	29	42	55	70							
23	19	30	43	57								
24	20	31	44									
25	20	32										
26	21											
27												
28												

Tablica F Wilcoxon test sume rangova (kritične vrijednosti za W pri α = 0,05, dvostrani)*(nastavak)*

$w_{0.05}$

n_2 \ n_1*	4	5	6	7	8	9	10	11	12	13	14	15
4	10											
5	11	17										
6	12	18	26									
7	13	20	27	36								
8	14	21	29	38	49							
9	15	22	31	40	51	63						
10	15	23	32	42	53	65	78					
11	16	24	34	44	55	68	81	96				
12	17	26	35	46	58	71	85	99	115			
13	18	27	37	48	60	73	88	103	119	137		
14	19	28	38	50	63	76	91	106	123	141	160	
15	20	29	40	52	65	79	94	110	127	145	164	185
16	21	31	42	54	67	82	97	114	131	150	169	
17	21	32	43	56	70	84	100	117	135	154		
18	22	33	45	58	72	87	103	121	139			
19	23	34	46	60	74	90	107	124				
20	24	35	48	62	77	93	110					
21	25	37	50	64	79	95						
22	26	38	51	66	82							
23	27	39	53	68								
24	28	40	55									
25	28	42										
26	29											
27												
28												

Tablica G Kritične vrijednosti K_D za Kolomogorov-Smirnov test za dva uzorka [1]

N	One-tailed test*		Two-tailed test†	
	α = .05	α = .01	α = .05	α = .01
3	3	—	—	—
4	4	—	4	—
5	4	5	5	5
6	5	6	5	6
7	5	6	6	6
8	5	6	6	7
9	6	7	6	7
10	6	7	7	8
11	6	8	7	8
12	6	8	7	8
13	7	8	7	9
14	7	8	8	9
15	7	9	8	9
16	7	9	8	10
17	8	9	8	10
18	8	10	9	10
19	8	10	9	10
20	8	10	9	11
21	8	10	9	11
22	9	11	9	11
23	9	11	10	11
24	9	11	10	12
25	9	11	10	12
26	9	11	10	12
27	9	12	10	12
28	10	12	11	13
29	10	12	11	13
30	10	12	11	13
35	11	13	12	
40	11	14	13	

Tablica H Friedmanova analiza varijance s dva promjenjiva faktora (k = 3) [1]

$N = 2$		$N = 3$		$N = 4$		$N = 5$	
χ_r^2	p	χ_r^2	p	χ_r^2	p	χ_r^2	p
0	1.000	.000	1.000	.0	1.000	.0	1.000
1	.833	.667	.944	.5	.931	.4	.954
3	.500	2.000	.528	1.5	.653	1.2	.691
4	.167	2.667	.361	2.0	.431	1.6	.522
		4.667	.194	3.5	.273	2.8	.367
		6.000	.028	4.5	.125	3.6	.182
				6.0	.069	4.8	.124
				6.5	.042	5.2	.093
				8.0	.0046	6.4	.039
						7.6	.024
						8.4	.0085
						10.0	.00077

$N = 6$		$N = 7$		$N = 8$		$N = 9$	
χ_r^2	p	χ_r^2	p	χ_r^2	p	χ_r^2	p
.00	1.000	.000	1.000	.00	1.000	.000	1.000
.33	.956	.286	.964	.25	.967	.222	.971
1.00	.740	.857	.768	.75	.794	.667	.814
1.33	.570	1.143	.620	1.00	.654	.889	.865
2.33	.430	2.000	.486	1.75	.531	1.556	.569
3.00	.252	2.571	.305	2.25	.355	2.000	.398
4.00	.184	3.429	.237	3.00	.285	2.667	.328
4.33	.142	3.714	.192	3.25	.236	2.889	.278
5.33	.072	4.571	.112	4.00	.149	3.556	.187
6.33	.052	5.429	.085	4.75	.120	4.222	.154
7.00	.029	6.000	.052	5.25	.079	4.667	.107
8.33	.012	7.143	.027	6.25	.047	5.556	.069
9.00	.0081	7.714	.021	6.75	.038	6.000	.057
9.33	.0055	8.000	.016	7.00	.030	6.222	.048
10.33	.0017	8.857	.0084	7.75	.018	6.889	.031
12.00	.00013	10.286	.0036	9.00	.0099	8.000	.019
		10.571	.0027	9.25	.0080	8.222	.016
		11.143	.0012	9.75	.0048	8.667	.010
		12.286	.00032	10.75	.0024	9.556	.0060
		14.000	.000021	12.00	.0011	10.667	.0035
				12.25	.00086	10.889	.0029
				13.00	.00026	11.556	.0013
				14.25	.000061	12.667	.00066
				16.00	.0000036	13.556	.00035
						14.000	.00020
						14.222	.000097
						14.889	.000054
						16.222	.000011
						18.000	.0000006

Tablica H Friedmanova analiza varijance s dva promjenjiva faktora (k = 4)*(nastavak)*

$N = 2$		$N = 3$		$N = 4$			
χ_r^2	p	χ_r^2	p	χ_r^2	p	χ_r^2	p
.0	1.000	.2	1.000	.0	1.000	5.7	.141
.6	.958	.6	.958	.3	.992	6.0	.105
1.2	.834	1.0	.910	.6	.928	6.3	.094
1.8	.792	1.8	.727	.9	.900	6.6	.077
2.4	.625	2.2	.608	1.2	.800	6.9	.068
3.0	.542	2.6	.524	1.5	.754	7.2	.054
3.6	.458	3.4	.446	1.8	.677	7.5	.052
4.2	.375	3.8	.342	2.1	.649	7.8	.036
4.8	.208	4.2	.300	2.4	.524	8.1	.033
5.4	.167	5.0	.207	2.7	.508	8.4	.019
6.0	.042	5.4	.175	3.0	.432	8.7	.014
		5.8	.148	3.3	.389	9.3	.012
		6.6	.075	3.6	.355	9.6	.0069
		7.0	.054	3.9	.324	9.9	.0062
		7.4	.033	4.5	.242	10.2	.0027
		8.2	.017	4.8	.200	10.8	.0016
		9.0	.0017	5.1	.190	11.1	.00094
				5.4	.158	12.0	.000072

Tablica I Kritične vrijednosti H za Kruskal-Wallis analizu varijance s jednim promjenjivim faktorom [1]

Sample sizes			H	p
n_1	n_2	n_3		
5	2	2	6.5333	.008
			6.1333	.013
			5.1600	.034
			5.0400	.056
			4.3733	.090
			4.2933	.122
5	3	1	6.4000	.012
			4.9600	.048
			4.8711	.052
			4.0178	.095
			3.8400	.123
5	3	2	6.9091	.009
			6.8218	.010
			5.2509	.049
			5.1055	.052
			4.6509	.091
			4.4945	.101
5	3	3	7.0788	.009
			6.9818	.011
			5.6485	.049
			5.5152	.051
			4.5333	.097
			4.4121	.109
5	4	1	6.9545	.008
			6.8400	.011
			4.9855	.044
			4.8600	.056
			3.9873	.098
			3.9600	.102
5	4	2	7.2045	.009
			7.1182	.010
			5.2727	.049
			5.2682	.050
			4.5409	.098
			4.5182	.101
5	4	3	7.4449	.010
			7.3949	.011
			5.6564	.049
			5.6308	.050
			4.5487	.099
			4.5231	.103
5	4	4	7.7604	.009
			7.7440	.011
			5.6571	.049
			5.6176	.050
			4.6187	.100
			4.5527	.102
5	5	1	7.3091	.009
			6.8364	.011
			5.1273	.046
			4.9091	.053
			4.1091	.086
			4.0364	.105
5	5	2	7.3385	.010
			7.2692	.010
			5.3385	.047
			5.2462	.051
			4.6231	.097
			4.5077	.100
5	5	3	7.5780	.010
			7.5429	.010
			5.7055	.046
			5.6264	.051
			4.5451	.100
			4.5363	.102
5	5	4	7.8229	.010
			7.7914	.010
			5.6657	.049
			5.6429	.050
			4.5229	.099
			4.5200	.101
5	5	5	8.0000	.009
			7.9800	.010
			5.7800	.049
			5.6600	.051
			4.5600	.100
			4.5000	.102