Višestruko uspoređivanje

Hasić, Almeida

Master's thesis / Diplomski rad

2014

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:217:970538

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-04-20**

Repository / Repozitorij:

Repository of the Faculty of Science - University of Zagreb

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO-MATEMATIČKI FAKULTET
MATEMATIČKI ODSJEK

Almeida Hasić

VIŠESTRUKO USPOREĐIVANJE

Diplomski rad

Zagreb, 2014.

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO-MATEMATIČKI FAKULTET
MATEMATIČKI ODSJEK

Almeida Hasić

VIŠESTRUKO USPOREĐIVANJE

Diplomski rad

Voditelj rada:
Prof. dr. sc. Anamarija Jazbec

Zagreb, 2014.

Ovaj diplomski rad obranjen je dana _____ pred ispitnim povjerenstvom u sastavu:

1. _____ , predsjednik
2. _____, član
3. _____, član

Povjerenstvo je rad ocijenilo ocjenom _____.

Potpisi članova povjerenstva:

1. _____
2. _____
3. _____

Zahvaljujem svojoj mentorici Anamariji Jazbec na pristupačnosti i trudu pri izradi ovog diplomskog rada, a posebno na vrlo ugodnoj suradnji te svojoj obitelji i prijateljima na strpljenju i podršci.

Sadržaj

Sadržaj iv

Uvod 1

1. Parametrijski testovi 3

    1.1. Fisherova LSD procedura 4

    1.2. Tukey procedura 6

    1.3. Bonferroni procedura 10

    1.4. Newman-Keuls procedura 13

    1.5. Dunnett procedura 14

    1.6. Sheffé procedura 17

    1.7. Kontrasti 19

2. Neparametrijski testovi 27

    2.1. Neparametrijska višestruka usporedba 29

    2.2. Neparametrijski višestruki kontrasti 32

3. Primjeri 34

4. Prilozi 60

Bibliografija 67

Uvod

Koristeći jednofaktorsku analizu varijance eksperimentalnog dizajna, možemo testirati nultu hipotezu $H_0: \mu_1 = \mu_2 = \cdots = \mu_k$. Međutim, odbacivanje H_0 ne upućuje da su svih k aritmetičkih sredina međusobno različite, također ne znamo koliko razlika postoji kao ni gdje su razlike locirane unutar k populacijskih sredina.

Na primjer, ako je $k = 3$, a $H_0: \mu_1 = \mu_2 = \mu_3$ odbačena, tada ne znamo koja je od $H_A: \mu_1 \neq \mu_2 = \mu_3$, $H_A: \mu_1 = \mu_2 \neq \mu_3$, $H_A: \mu_1 \neq \mu_2 \neq \mu_3$ odgovarajuća alternativna hipoteza.

Generalno je krivo upotrijebiti višestruki t-test kako bi se ispitale razlike između svih mogućih parova aritmetičkih sredina. U ovom radu bit će prezentirane statističke procedure koje se mogu koristiti u takve svrhe; njih zovemo testovi višestrukog uspoređivanja ili post-hoc testovi.

Problem višestrukog uspoređivanja privlači mnogo pažnje i često se pojavljuje u statističkoj literaturi, ali još uvijek ne postoji dogovor koja je najbolja procedura za rutinsku upotrebu. Među najuže prihvaćenim i najčešće korištenim metodama su Turkey test (Turkey, 1953) i Newman-Keuls test (Newman, 1939; Keuls, 1952). Duncanov test (Duncan, 1955), često znan kao „Duncanov novi test višestrukih raspona“ zato što nasljeđuje ranije procedure (Duncan, 1951), također možemo pronaći u literaturi, ali je baziran na drugačijoj teoriji, koja nije toliko široko prihvaćena kao Tukey i Newman-Keuls procedure, te je proglašena od strane nekih (Day and Quinn, 1989) da daje slabe rezultate. Također se možemo susresti s procedurom nazvanom test „najmanje značajne razlike“ (LSD) koju je razvio Fisher (1953).[1] Druge procedure višestrukog uspoređivanja zovu se Dunn ili Bonferroni test te Sheffé-ov test. Sve te procedure bit će obrađene u prvom poglavlju. Također će biti prikazana posebna vrsta usporedbi nazvana višestruki kontrasti.

Testiranje višestrukih usporedbi se najčešće provodi samo ako analiza varijance odbaci višestruku hipotezu jednakosti sredina. Zbog toga, na te se testove odnosi kao na „a posteriori“ testove.

U trećem poglavlju dotaknut će se tema neparametrijskih ekvivalenata post-hoc testova ANOVA procedure, to jest višestruko uspoređivanje nakon Kruskal-Wallis testa.

U posljednjem poglavlju bit će provedena potpuna analiza nad stvarnim podacima pomoću programskog paketa SAS.

Poglavlje 1

Parametrijski testovi

Postupak odabira ispravnog statističkog testa može biti problematičan zadatak, no dobro poznavanje i razumijevanje odgovarajućih statističkih pojmova može pomoći u donošenju ispravne odluke. Posebno je potrebno znati s kojim tipom podataka raspolažemo, kako su ti podaci organizirani, koliko uzoraka imamo i jesu li podaci zavisni ili nezavisni; moramo si postaviti pitanje slijede li podaci iz populacije Gaussovu raspodjelu ili ne.

Ukoliko podaci dolaze iz normalne distribucije s jednakim populacijskim varijancama te su mjerenja nezavisna, možemo primjeniti parametrijske testove. Općenito, parametarijske metode zahtjevaju više pretpostavki nego neparametrijske. No ukoliko su sve pretpostavke zadovoljene, one daju točnije i preciznije rezultate. Stoga se smatra da imaju jaču statističku snagu.

Parametrijski testovi odnose se na parametar pretpostavljene razdiobe i koriste se za provjeru hipoteza o nepoznatoj vrijedosti parametra. Kada je potrebno usporediti aritmetičke sredine dviju normalno distribuiranih populacija koristi se Studentov t-test. Za testiranje značajnosti razlike među aritmetičkim sredinama triju ili više uzoraka koji predstavljaju populacije, koristi se jednosmjerna analiza varijance, to jest ANOVA[1]. Ona se temelji na F-testu[2] kojim procjeni vjerojatnost da su ustanovljene razlike slučajne (ne odbacujemo hipotezu H_0, to jest uzorci su iz iste populacije) ili statistički značajne (odbacujemo H_0, uzorci su iz različitih populacija). Ukoliko postoji statistički

[1] Za detalje pogledati [5].

[2] Testna statistika: $F = \frac{MST}{MSE}, MST = \frac{SST}{m-1}, MSE = \frac{SSE}{n-m}$, gdje je SST suma kvadrata zbog tretmana, a SSE suma kvadrata pogrešaka, m-1 i n-m broj stupnjeva slobode.

značajna razlika među aritmetičkim sredinama promatranih populacija, potrebno je odgovoriti na pitanje između kojih populacija postoji razlika. Daljnjom analizom pomoću post-hoc testova saznajemo između kojih parova uzoraka su razlike statistički značajne i na kojoj razini vjerojatnosti.

Testovi višestrukih usporedbi aritmetičkih sredina ili post-hoc testovi imaju istu temeljnu pretpostavku kao i testovi analize varijance: homogenost i normalnost varijanci unutar populacije. Iako se Turkey test pokazao kao snažan uzimajući u obzir te pretpostavke (Keselman, 1976), snaga procedura nije dobro poznata, prisutni su mogući štetni efekti na greške Tipa I[3] i Tipa II[4] ukoliko su pretpostavke uvelike prekršene.[1] Pretpostavka homogenosti varijance je mnogo ozbiljnija, parametrijsko testiranje višestrukih usporedbi ne bi se smjelo izvoditi ukoliko je proglašena heteroskedastičnost.

Nakon analize varijance, kada se pokaže postojanje statistički značajne razlike između uzoraka, javlja se problem odlučivanja koji se nivo kategorijske varijable razlikovao u okviru grupe tretmana koje promatramo. Postoje različiti pristupi koji se mogu upotrijebiti, a najvažniji su navedeni u nastavku.

1.1. Fisherova LSD procedura

Fisherov LSD test (eng. Least significant difference) ili test najmanje značajne razlike je najjednostavniji pristup za uspoređivanje aritmetičkih sredina promatranih uzoraka. Međutim, to je također i najmanje pouzdan test budući da se višestruko uspoređivanje provodi istovremeno, svaki sa jednakom razinom značajnosti α, najčešće 5%. To znači da je rizik od nastajanja greške u uspoređivanju svih sredina veći od te razine značajnosti. Dodatni problem koji se

[3] Greška tipa I predstavlja lažno pozitivan rezultat, to jest nulta hipoteza je odbačena iako je istinita, vjerojatnost pogreške 1. vrste označava se sa α i naziva se razina značajnosti.

[4] Greška tipa II odnosi se na promašaj stvarnog učinka, nulta hipoteza je prihvaćena iako je neistinita, vjerojatnost pogreške 2. vrste naziva se jakost testa i označava sa β.

ponekad javlja je pronalazak značajnog ukupnog rezultata u F-testu, a pojedinačne razlike se ne mogu prepoznati.

Kako bi pronašli koja se aritmetička sredina razlikuje od koje, koristimo formulu za računanje najmanje značajne razlike (LSD):

$$\text{LSD} = t_{v,\alpha/2} s \sqrt{\frac{1}{n_i} + \frac{1}{n_j}} \tag{1.1}$$

gdje je $s = \sqrt{\text{MSE}}$ (MSE – varijanca reziduala) iz tablice ANOVA, t vrijednost t-distiribucije (Tablica 4.1) sa stupnjevima slobode reziduala $v = k(n_i - 1)$ i razinom značajnosti testa $\alpha/2$, tj. riječ je o dvosmjernom testu. $n_i, i = 1, \dots, k$ je broj opservacija u i-tom uzorku, a $n_j, j = 1, \dots, k$ broj opservacija u j-tom uzorku, k je broj promatranih uzoraka, to jest broj razina kategorijske varijable. Svake dvije aritmetičke sredine izdvojene LSD testom su identificirane kao različite. Obično, n_i i n_j su planirani da budu jednake, iako to nije nužno.

Pokažimo postupak na slijedećem primjeru (obrada istih podataka u programskom paketu SAS prikazana je u Poglavlju 3).

Primjer 1

Provedeno je istraživanje o koncentraciji stroncija (mg/ml) u pet različitih voda Sjeverne Amerike (1 – Grayson's Pond, 2 – Beaver Lake, 3 – Angler's Cove, 4 – Appletree Lake, 5 – Rock River) tako da je u svakoj od njih uzeto po 6 uzoraka vode (baza je preuzeta iz [1], str. 210) te su podaci prikazani slijedećom tablicom:

Grayson's Pond	*Beaver Lake*	*Angler's Cove*	*Appletree Lake*	*Rock River*
28,2	39,6	46,3	41,0	56,3
33,2	40,8	42,1	44,1	54,1
36,4	37,9	43,5	46,4	59,4
34,6	37,1	48,8	40,2	62,7
29,1	43,6	43,7	38,6	60,0
31,0	42,4	40,1	36,3	57,3

Potrebno je utvrditi je li jednaka koncentracija stroncija unutar svakog oblika vode i među kojima postoje značajnije razlike. Razina značajnosti je 5%.

Za $k = 5$ i $n_i = 6$ broj stupnjeva slobode iznosi $v = 5(6-1) = 25$. Iz ANOVA tablice pročitamo dobivenu srednjekvadratnu pogrešku $\mathrm{MSE} = 9{,}7652$ te računamo $s = \sqrt{\mathrm{MSE}} = \sqrt{9{,}7652} = 3{,}1249$. U Tablici 4.1 pronađemo vrijednost za razinu značajnosti $\alpha/2 = 0{,}025$ koja iznosi $t_{25,0.025} = 2{,}06$. Sada računamo LSD vrijednost prema Formuli (1.1)

$$\mathrm{LSD} = 2{,}06 \times 3{,}1249 \times \sqrt{\frac{1}{6} + \frac{1}{6}} = 3{,}7166.$$

Prema dobivenim aritmetičkim sredinama uzoraka računamo razlike i uspoređujemo sa LSD vrijednosću (navedene su samo neke, detaljnije pogledati u Poglavlju 3):

$$|\bar{x}_1 - \bar{x}_2| = 8{,}15 \geq 3{,}7166,$$
$$|\bar{x}_1 - \bar{x}_3| = 12 \geq 3{,}7166,$$
$$|\bar{x}_2 - \bar{x}_4| = 0{,}87 < 3{,}7166,$$
$$|\bar{x}_3 - \bar{x}_4| = 2{,}98 < 3{,}7166.$$

Na temelju izračuna možemo zaključiti da se grupe 1 i 2 te 1 i 3 međusobno statistički značajno razlikuju, dok na primjer grupe 2 i 4 te 3 i 4 se značajno ne razlikuju.

1.2. Tukey procedura

Turkey test je često korištena procedura višestrukog uspoređivanja. Uzima u obzir nultu hipotezu $H_0 : \mu_B = \mu_A$ u odnosu na alternativnu hipotezu H_A: $\mu_B \neq \mu_A$, gdje oznake A i B označavaju bilo koji mogući par grupa. Za k grupa, može biti učinjeno $\frac{k(k-1)}{2}$ različitih usporedbi po parovima. To dobijemo pomoću

$$ {}_kC_2 = \frac{k!}{2!\,(k-2)!} = \frac{k(k-1)(k-2)!}{2!\,(k-2)!} = \frac{k(k-1)}{2} \qquad (1.2) $$

gdje je ${}_kC_2$ broj kombinacija k grupa učinjen s dvije u isto vrijeme.

Dakle, u primjeru sa 5 grupa podataka imamo ${}_5C_2 = 10$ kombinacija. Kada na temelju jednofaktorske analize varijance odbacimo hipotezu H_0, koristimo Tukey-ev post-hoc test kako bi odredili između kojih aritmetičkih sredina populacija postoji značajna razlika.

Testna statistika dana je formulom

$$ Q = q_{\alpha,v,k}\sqrt{\frac{\mathrm{MSE}}{n_i}} \qquad (1.3) $$

U Primjeru 1 p-vrijednost dobivena u ANOVA tablici je manja od 0,0001 stoga je hipoteza $H_0\colon \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$ odbačena. Prvi korak u daljnjoj analizi je numeriranje i uređivanje svih aritmetičkih sredina uzoraka po veličini u rastućem poretku. U danom primjeru aritmetičke sredine poredane u rastućem poretku su

$\bar{x}_5 = 58{,}30,$
$\bar{x}_3 = 44{,}08,$
$\bar{x}_4 = 41{,}10,$
$\bar{x}_2 = 40{,}23,$
$\bar{x}_1 = 32{,}08.$

Zatim se tabeliraju parovi razlika aritmetičkih sredina uzoraka $|\bar{x}_B - \bar{x}_A|$.

Usporedbe	Razlike
5 – 1	26,22
5 – 2	18,07
5 – 4	17,20
5 – 3	14,22
3 – 1	12,00
3 – 2	3,85
3 – 4	2,98
4 – 1	9,02
4 – 2	0,87
2 – 1	8,15

Tablica 1.1: Razlike aritmetičkih sredina po parovima

Kao što razlika između aritmetičkih sredina dijeljena s odgovarajućom standardnom pogreškom daje t-vrijednost, tako q-vrijednost u Tukey-evom testu računa se dijeljenjem te razlike s

$$\mathrm{SE} = \sqrt{\frac{\mathrm{MSE}}{n_i}} \tag{1.4}$$

gdje MSE predstavlja srednjekvadratnu pogrešku dobivenu u analizi varijance, a n je broj podataka u svakoj od grupa A i B.

$$\mathrm{SE} = \sqrt{\frac{9{,}7652}{6}} = 1{,}2757$$

Ako je izračunata q-vrijednost

$$q = \frac{|\bar{x}_B - \bar{x}_A|}{\mathrm{SE}} \tag{1.5}$$

veća ili jednaka od kritične vrijednosti, $q_{\alpha,v,k}$ iz Tablice 4.2, tada je $H_0: \mu_A = \mu_B$ odbačena.

Za dani primjer kritična vrijednost iznosi $q_{0.05,24,5} = 4{,}166$ (uzimamo vrijednost za 24 stupnja slobode budući da u tablici nije navedena potrebna vrijednost). Sada izračunavamo q-vrijednost za svaki par grupa:

Usporedbe	q	Usporedbe	q
5 – 1	$\frac{26{,}22}{1{,}2757} = 20{,}553$	3 – 2	$\frac{3{,}85}{1{,}2757} = 3{,}018$
5 – 2	$\frac{18{,}07}{1{,}2757} = 14{,}165$	3 – 4	$\frac{2{,}98}{1{,}2757} = 2{,}336$
5 – 4	$\frac{17{,}20}{1{,}2757} = 13{,}483$	4 – 1	$\frac{9{,}02}{1{,}2757} = 7{,}071$
5 – 3	$\frac{14{,}22}{1{,}2757} = 11{,}147$	4 – 2	$\frac{0{,}87}{1{,}2757} = 0{,}682$
3 – 1	$\frac{12{,}00}{1{,}2757} = 9{,}407$	2 – 1	$\frac{8{,}15}{1{,}2757} = 6{,}389$

Tablica 1.2: q-vrijednosti po parovima uzoraka

Kritična vrijednost u ovom testu poznata je pod nazivom Studentizirani raspon[5], u oznaci q, te je ovisan o α (razina značajnosti), v (pogreška stupnja slobode za analizu varijance) i k (ukupan broj testiranih aritmetičkih sredina).

Razina značajnosti, α, je vjerojatnost nailaženja na barem jednu pogrešku Tipa I (to jest vjerojatnost lažnog odbacivanja barem jedne nulte hipoteze) tijekom uspoređivanja svih parova aritmetičkih sredina; to nije vjerojatnost da se učini greška Tipa I za jednu usporedbu.

Zaključci testiranja višestrukog uspoređivanja ovise o redoslijedu u kojem se usporedbe po parovima razmatraju. Pravilno je usporediti prvu po veličini najveću srednju vrijednost s najmanjom, zatim najveću sa slijedećom najmanjom, i tako nastaviti sve dok se najveća sredina ne usporedi s drugom po

[5] Termin eng. „Studentized range“ prvi su upotrijebili E. S. Pearson i H. O. Hartley 1953. godine (David, 1995).[1]

veličini najvećom. Nakon toga se uspoređuje druga najveća s najmanjom, zatim s drugom po redu najmanjom itd.

Tako iz Tablice 1.2 možemo zaključiti da se grupa 5 statistički značajno razlikuje od grupe 1 zbog $q = 20{,}553 \geq 4{,}166 = q_{0.05,24,5}$. Analogno, grupa 5 značajno se razlikuje od svih ostalih grupa kao i grupa 1, dok za grupe 2 i 3 možemo zaključiti da se statistički značajno ne razlikuju kao ni grupe 3 i 4 te 4 i 2.

Drugo važno pravilo unutar procedure je to da ako nisu pronađene razlike između dviju aritmetičkih sredina, tada je zaključak da ne postoje razlike između sredina, uključujući i te dvije, i takve razlike se ne testiraju.

Ako veličine k grupa nisu jednake, potrebna je mala modfikacija unutar Tukey procedure. Za svaku usporedbu koja sadrži nejednak n, standardna pogreška se računa koristeći slijedeću aproksimaciju (Kramer, 1956; Tukey, 1953) [1]:

$$\mathrm{SE} = \sqrt{\frac{\mathrm{MSE}}{2}\left(\frac{1}{n_A} + \frac{1}{n_B}\right)} \tag{1.6}$$

Također je moguće da $H_0: \mu_1 = \mu_2 = \cdots = \mu_k$ bude odbačena u analizi varijance i da test višestruke usporedbe ne uspije otkriti razlike između bilo kojih parova aritmetičkih sredina. Ovakav slučaj se ne pojavljuje često, ali to se odražava na činjenicu da je analiza varijance jači test nego test višestrukog uspoređivanja (to jest greška Tipa II će se prije pojaviti u višestrukom uspoređivanju nego u analizi varijance). Ponavljanje eksperimenta sa većom veličinom uzorka teži boljem otkrivanju razlika među aritmetičkim sredinama u analizi višestruke usporedbe.

1.3. Bonferroni procedura

Bonferoni test je tip testa višestruke usporedbe koji se koristi u statističkim analizama. Ovom se procedurom pokušava spriječiti netočno pojavljivanje statističke značajnosti smanjivanjem vrijednosti α. No, procedura je kritizirana jer

je prekonzervativna i može ne zapaziti neke značajne razlike. Za sve parove usporedbi, Bonferroni procedura daje duže pouzdane intervale nego Tukey test.
Prema [3] Bonferroni test provodi se u nekoliko koraka. Najprije se odredi ukupna razina značajnosti α ili eventualno ukupni pouzdani interval $100(1-\alpha)\%$. U danom primjeru je $\alpha = 0{,}05$, odnosno 95% pouzdani interval. Zatim se odredi razina značajnosti za svaku usporedbu sa $\alpha/\,{}_kC_2$, gdje se ${}_kC_2$ računa se iz Formule (1.2). Sada za $k = 5$ uzoraka imamo ${}_5C_2 = 10$ usporedbi te je

$$\frac{\alpha}{{}_5C_2} = \frac{0{,}05}{10} = 0{,}005.$$

Za dani α, pouzdani interval $100(1-\alpha/\,{}_kC_2)\%$ za $\mu_i - \mu_j,\ i,j = 1,\dots,k$ računa se pomoću formule

$$\left(\bar{x}_i - \bar{x}_j\right) \pm t_{\alpha',v}\, s \sqrt{\frac{1}{n_i} + \frac{1}{n_j}} \qquad (1.7)$$

gdje je $s = \sqrt{\text{MSE}}$, $\alpha' = \frac{\alpha}{2\,{}_kC_2}$ te $t_{\alpha',v}$ t-vrijednost iz Tablice 4.1 sa $v = k(n_i - 1)$ stupnjeva slobode.
Dakle, u primjeru imamo $v = 25$ stupnjeva slobode, stoga je

$s = \sqrt{9{,}7652} = 3{,}1249,$

$\alpha' = \frac{0{,}05}{2\times 10} = 0{,}0025$ i

$t_{0.0025,25} = 3{,}078,$

a aritmetičke sredine pojedinih uzoraka su

$\bar{x}_1 = 32{,}08,$
$\bar{x}_2 = 40{,}23,$
$\bar{x}_3 = 44{,}08,$
$\bar{x}_4 = 41{,}10,$

$\bar{x}_5 = 58{,}30.$

Sada možemo procjeniti $99{,}5\%$ pouzdane intervale za svaku pojedinu usporedbu aritmetičkih sredina populacija.

$$\left(\bar{x}_i - \bar{x}_j\right) \pm t_{0.0025,25} \times 3{,}1249 \times \sqrt{\frac{1}{6} + \frac{1}{6}}, i, j = 1, \dots, k$$

$$\left(\bar{x}_i - \bar{x}_j\right) \pm 5{,}55$$

$$(\bar{x}_1 - \bar{x}_2) \pm 5{,}55$$
$$(32{,}08 - 40{,}23) \pm 5{,}55$$
$$-8{,}15 \pm 5{,}55$$

Iz ovog računa slijedi da je za uspoređivanje aritmetičkih sredina prvog i drugog uzorka pouzdani interval (-13.7, -2.6). To znači da aritmetička sredina drugog uzorka prekoračuje aritmetičku sredinu prvog uzorka za minimalno 2,6 i maksimalno 13,7 iz čega zaključujemo kako se ta dva uzorka statistički značajno razlikuju. Analogno računamo pouzdani interval za treći i četvrti uzorak:

$$(\bar{x}_3 - \bar{x}_4) \pm 5{,}55$$
$$(44{,}08 - 41{,}1) \pm 5{,}55$$
$$2{,}98 \pm 5{,}55$$

Možemo zaključiti kako se uzorci 3 i 4 statistički značajno ne razlikuju jer pouzdani interval (-2.57, 8.53) za razlike aritmetičkih sredina tih uzoraka sadrži vrijednost 0. Ostale pouzdane intervale je jednostavno izračunati te se analogno donose zaključci.

1.4. Newman-Keuls procedura

Newman- Keuls test često se pojavljuje i pod imenom „Student-Newman-Keuls test“ ili „SNK test“. Provodi se poput Tukey testa sa jednom iznimkom. Aritmetičke sredine uzoraka se rangiraju, odrede se razlike među parovima aritmetičkih sredina te se izračunavanju standardne pogreške kao u Sekciji 1.2. Zatim se izračunava testna statistika q jednako kao i q za Tukey test (to jest pomoću Formule (1.5)). Razlika naspram Tukey testa je u određivanju kritične vrijednosti $q_{\alpha,\nu,p}$, gdje je p broj aritmetičkih sredina u rasponu sredina koje se u tom trenutku testiraju. Za usporedbu rangiranih sredina $\bar{x}_5$ i $\bar{x}_1$ iz Primjera 1, uzimamo u obzir 5 aritmetičkih sredina, stoga je $p = 5$, dok za uspoređivanje $\bar{x}_3$ i $\bar{x}_2$ vrijedi $p = 3$.

Sada u Tablici 4.2 pronalazimo kritične vrijednosti $q_{0.05,24,p}$ za svaki dobiveni p:

Usporedbe	p	$q_{0.05,24,p}$	Usporedbe	p	$q_{0.05,24,p}$
5 – 1	5	4,166	3 – 2	3	3,532
5 – 2	4	3,901	3 – 4	2	2,919
5 – 4	3	3,532	4 – 1	3	3,532
5 – 3	2	2,919	4 – 2	2	2,919
3 – 1	4	3,901	2 – 1	2	2,919

Tablica 1.3: Broj aritmetičkih sredina i kritične vrijednosti po parovima uzoraka

Na temelju Tablice 1.2 i 1.3 sada možemo donijeti zaključak o statistički značajnim razlikama među grupama. U prvoj usporedbi je $q = 20{,}553 \geq 4{,}166 = q_{0.05,24,5}$ stoga odbacujemo hipotezu o jednakosti prve i pete populacije.

Analogno vrijedi i za parove aritmetičkih sredina grupa 5 i 2, 5 i 4, 5 i 3, 3 i 1, 4 i 1 te 2 i 1, dok za preostale tri usporedbe ne možemo odbaciti takvu hipotezu.

Tip testa višestruke usporedbe koji koristi različite kritične vrijednosti za različito rangirane aritmetičke sredine, poput Newman-Kelts testa, naziva se „Test višestrukih raspona“.[1]

Ukoliko veličine svih k uzoraka nisu jednake, koristi se Formula (1.4) za izračun standardne pogreške SE, kao i kod Tukey testa.

U danom primjeru Tukey test i SNK test doveli su do jednakog zaključka što nije uvijek slučaj. Newman-Keuls procedura donosi zaključke u kojima su razlike među aritmetičkim sredinama statistički značajnije nego kod Tukey procedure, to jest prethodno testiranje je snažnije. Neki autori, na primjer Miller (1981), kritiziraju Tukeyev test s argumentom da je konzervativan, to jest zaključuje premalo statistički značajnih razlika, dok su drugi, npr. Einot i Gabriel (1975) i Ramsey (1978), protiv tog testa jer može lažno proglasiti značajne razlike s većom vjerojatnošću nego α.[1]

Kao kompromis između ova dva postupka, Tukey je prodložio korištenje kritične vrijednosti koja se dobije kao aritmetička sredina kritičnih vrijednosti oba testa (to jest aritmetička sredina od $q_{\alpha,v,k}$ i $q_{\alpha,v,p}$). Taj se postupak naziva „potpuno značajan test razlika“, a naziv se ponekad koristi i kao sinonim za Tukey proceduru.

1.5. Dunnett procedura

Dunnett (1995) specijalno je dizajnirao ovu procedru za uspoređivanje jednog individualnog uzorka, koji se naziva kontrolnim, sa svim ostalima. Koristi se upravo u eksperimentima u kojima je cilj odrediti je li se jedna aritmetička

sredina uzorka statistički značajno razlikuje od ostalih $k-1$ sredina uzoraka ukoliko je broj uzoraka jednak k. [1]
Dvije sredine su statistički značajno različite ukoliko je njihova razlika veća ili jednaka od vrijednosti dane formulom

$$D = d_{\alpha,v,k}\sqrt{\frac{2\mathrm{MSE}}{n_i}} \tag{1.8}$$

Pretpostavimo da nas zanima kako se drugi uzorak, odnosno njegova aritmetička sredina, razlikuje od ostalih u Primjeru 1. Dakle kontrolna aritmetička sredina iznosi $\bar{x}_{control} = 40{,}23$ te ju uspoređujemo sa $k-1=4$ rangirane sredine uzoraka. Broj stupnjeva slobode je $v = k(n_i - 1) = 25$.

Znajući ukupan broj uzoraka u eksperimentu i broj stupnjeva slobode za analizu varijance, gdje je nulta hipoteza H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$, u Tablici 4.3 možemo pronaći vrijednost Dunnettove kritične vrijednosti $d_{\alpha,v,k}$ koja iznosi $d_{0.05,24,5} = 2{,}61$ (budući da u Tablici 4.3 nije dana kritična vrijednost za 25 stupnjeva slobode, uzimamo prvu vrijednost prije toga). Stopa pogreške, α, označava vjerojatnost da ćemo učiniti pogrešku Tipa I negdje između usporedbi parova aritmetičkih sredina.

Standardna pogreška za Dunnettov test dana je formulom

$$\mathrm{SE} = \sqrt{\frac{2\mathrm{MSE}}{n_i}} \tag{1.9}$$

kada su svi uzorci jednake duljine $n_i, i = 1, \dots, k$, odnosno formulom

$$\mathrm{SE} = \sqrt{\mathrm{MSE}\left(\frac{1}{n_A} + \frac{1}{n_{control}}\right)} \tag{1.10}$$

kada uzorci nisu jednake duljine, n_A predstavlja duljinu uzorka iz neke grupe A, a $n_{control}$ veličinu kontrolnog uzorka.

U Primjeru 1 je za $\mathrm{MSE} = 9{,}7652$ standardna pogreška jednaka

$$\mathrm{SE} = \sqrt{\frac{2 \times 9{,}7652}{6}} = 1{,}8042.$$

Testna statistika je

$$d = \frac{|\bar{x}_{control} - \bar{x}_A|}{\mathrm{SE}}. \qquad (1.11)$$

Za usporedbu sa aritmetičkom sredinom drugog uzorka, u Tablici 1.4 prikazane su testne statistike:

Usporedbe	d
2 – 5	10,02
2 – 3	2,13
2 – 4	0,48
2 – 1	4,52

Tablica 1.4: Testne statistike po parovima aritmetičkih sredina

Na temelju prethodne tablice zaključujemo da se aritmetička sredina drugog uzorka statistički značajno razlikuje od prve i pete sredine (to vidimo iz $d = 10{,}02 \geq 2{,}61 = d_{0.05,24,5}$ i $d = 4{,}52 \geq 2{,}61$) dok se za grupe 3 i 4 ne možemo donijeti isti zaključak.

Kada je želja istraživača usporedba skupine aritmetičkih sredina s kontolnom, tada bi kontrolna grupa trebala sadržavati više zapažanja od ostalih. Dunnett (1955) je pokazao kako bi optimalna veličina kontrolne grupe trebala imati malo manje od $\sqrt{k-1}$ pomnoženo s veličinom svake ostale grupe. [1]

1.6. Sheffé procedura

Sheffé (1953) je osmislio proceduru višestrukog uspoređivanja (ponekad nazvanu i S test) koja testira nultu hipotezu $H_0: \mu_A - \mu_B = 0$ kao i Tukey i Newman-Keuls metode (ali s manjom snagom); također se može koristiti i za posebna razmatranja koja će biti prikazana u Sekciji 1.7.
Za navedenu hipotezu vrijedi formula za testnu statistiku

$$S = \frac{|\bar{x}_B - \bar{x}_A|}{\mathrm{SE}}, \tag{1.12}$$

gdje je standardna pogreška SE dana formulom

$$\mathrm{SE} = \sqrt{\mathrm{MSE}\left(\frac{1}{n_A} + \frac{1}{n_B}\right)} \tag{1.13}$$

te je kritična vrijednost

$$S_\alpha = \sqrt{(k-1)F_{\alpha,k-1,N-k}}. \tag{1.14}$$

U navedenim formulama MSE predstavlja srednje kvadratnu pogrešku dobivenu u ANOVA tablici, k broj promatranih uzoraka, N ukupan broj podataka, a $F_{\alpha,k-1,N-k}$ vrijednost Fisherove F-razdiobe (vidi Tablicu 4.4) za razinu značajnosti α i $k-1$ i $N-k$ stupnjeva slobode.

Pokažimo na primjeru postupak Shefféovog višestrukog uspoređivanja.
Za $\alpha = 0{,}05$ kritična vrijednost za svaku usporedbu je

$$S_{0{,}05} = \sqrt{(5-1) \times F_{0.05,4,25}} = \sqrt{4 \times 2{,}76} = 3{,}32.$$

Budući da su sve veličine uzoraka jednake, SE za svaku usporedbu iznosi

$$\mathrm{SE} = \sqrt{9{,}7652 \times \left(\frac{1}{6} + \frac{1}{6}\right)} = 1{,}8.$$

Sada za rangirane aritmetičke sredine računamo testnu statistiku za svaku usporedbu posebno.

Usporedbe	S	Usporedbe	S
5 – 1	$\frac{26{,}22}{1{,}8} = 14{,}57$	3 – 2	$\frac{3{,}85}{1{,}8} = 2{,}14$
5 – 2	$\frac{18{,}07}{1{,}8} = 10{,}04$	3 – 4	$\frac{2{,}98}{1{,}8} = 1{,}66$
5 – 4	$\frac{17{,}20}{1{,}8} = 9{,}56$	4 – 1	$\frac{9{,}02}{1{,}8} = 5{,}01$
5 – 3	$\frac{14{,}22}{1{,}8} = 7{,}9$	4 – 2	$\frac{0{,}87}{1{,}8} = 0{,}48$
3 – 1	$\frac{12{,}00}{1{,}8} = 6{,}67$	2 – 1	$\frac{8{,}15}{1{,}8} = 4{,}53$

Tablica 1.5: Testne statistike po parovima usporedbi

Uspoređujući vrijednosti iz Tablice 1.5 s $S_{0{,}05} = 3{,}32$ zaključujemo kako se prvi i peti uzorak statistički značajno razlikuju od drugih dok se drugi, treći i četvrti uzorak statistički značajno ne razlikuju. Možemo primjetiti da je zaključak jednak kao i kod Tukey testa, što nije uvijek slučaj. Zapravo, Tukey test otkriva manje razlika nego Shefféov, to jest Shefféov test je skloniji tome da napravi grešku Tipa II. Zbog toga se uglavnom ne preporučuje za ovakav tip višestrukih usporedbi.

1.7. Kontrasti

Kontrasti su sredstva za testiranje teorije vezane za individualne tretmane ili grupe tretmana. Ako imamo 6 tretmana A, B, C, D, E, F možemo na primjer istražiti jesu li A i B jednaki ili možda želimo testirati jesu li A i B jednaki uzorcima D i F.

Kontrast je definiran kao

$$C = \sum_{i=1}^{k} c_i \left(\sum_{j=1}^{n_i} x_j \right) \tag{1.15}$$

gdje su $c_i, i = 1, \dots, k$ koeficijenti uz aritmetičke sredine (težine kontrasta) i k broj uzoraka koji promatramo. Kako bi to bio važeći kontrast, ova linearna kombinacija mora imati koeficijente za koje vrijedi

$$\sum_{i=1}^{k} c_i = 0. \tag{1.16}$$

U primjerima iznad možemo uzeti slijedeće koeficijente:

Kontrast	A	B	C	D	E	F
$A - B$	1	-1	0	0	0	0
$A, B - D, F$	1	1	0	-1	0	-1

Tablica 1.6: Koeficijenti za kontraste $A - B$ i $A, B - D, F$

Suma kvadrata za svaki kontrast od interesa računa se pomoću formule

$$\mathrm{SS}_C = \frac{C^2}{n \sum_{i=1}^{k} {c_i}^2}, \tag{1.17}$$

ima 1 stupanj slobode te se testira nasuprot srednje kvadratnoj pogrešci kao inače. U praksi, suma kvadrata se razdjeljuje na dijelove zbog pojedinih kontasta.

Ortogonalni kontrasti

Ukoliko želimo istražiti više od jednog kontrasta istovremeno, tada bi testovi kontrasta trebali biti nezavisni. To se postiže odabirom kontrasta koji su ortogonalni jedan na drugog.

Ako su koeficijenti jednog kontrasta c_i, a drugog d_i, tada su dva kontrasta ortogonalna ako vrijedi

$$\sum_{i=1}^{k} c_i d_i = 0. \qquad (1.18)$$

Broj ortogonalnih kontrasta koji se mogu istovremeno uspoređivati jednak je broju stupnjeva slobode za tretmane. Za 6 tretmana navedenih ranije postoji 5 kontrasta koje možemo promatrati.

Otkrivanje koji skup kontrasta su ortogonalni može biti vrlo teško i zahtjeva dosta pokušaja u otkrivanju. Prethodno spomenuta dva kontrasta su ortogonalni, a mogući skup kontrasta bi bio

Kontrast	A	B	C	D	E	F
$A-B$	1	−1	0	0	0	0
$A,B-D,F$	1	1	0	−1	0	−1
$A,B-C,E$	1	1	−1	0	−1	0
$D-F$	0	0	0	1	0	−1
$C-E$	0	0	1	0	−1	0

Tablica 1.7: Svi koeficijenti za 6 tretmana

Provedimo analizu dva ortogonalna kontrasta izabrana u Primjeru 1, odnosno testirajmo hipoteze

$$H_0: \mu_1 = \mu_2$$

$$H_0: \mu_1 + \mu_2 = \mu_4 + \mu_5$$

Budući da je broj stupnjeva slobode $df = 25$, a svaki kontrast ima samo $df_C = 1$ stupanj slobode, za $\alpha = 0{,}05$ u Tablici 4.5 se nalazi kritična vrijednost $f_{0.05,1,25} = 4{,}24$. Potrebne su nam još sume danih vrijednosti i srednje kvadratna pogreška $\mathrm{MSE} = 9{,}7652$ za izračun testnih statistika F_{1-2} i F_{12-45}.

Grupa	$\sum_{j=1}^{n_i} x_j$
1	192,5
2	241,4
3	264,5
4	246,6
5	349,8

Tablica 1.8: Sume vrijednosti po uzorcima

Koeficijenti kontrasta dani su slijedećom tablicom

Kontrasti	Grupa 1	Grupa 2	Grupa 3	Grupa 4	Grupa 5
$1-2$	1	-1	0	0	0
$1{,}2-4{,}5$	1	1	0	-1	-1

Tablica 1.9: Koeficijeti za kontraste $1-2$ i $1{,}2-4{,}5$

Kontrasti iznose

$$C_{1-2} = 1 \times 192{,}5 - 1 \times 241{,}4 = -48{,}9$$

$$C_{12-45} = 1 \times 192{,}5 + 1 \times 241{,}4 - 1 \times 246{,}6 - 1 \times 349{,}8 = -162{,}5$$

Sume kvadrata računaju se po Formuli (1.17), dakle

$$SS_{1-2} = \frac{2391{,}21}{6 \times 2} = 199{,}27$$

$$SS_{12-45} = \frac{26406{,}25}{6 \times 4} = 1100{,}26.$$

Testne statistike za svaki kontrast za dane hipoteze glase

$$F_{1-2} = \frac{SS_{1-2}/df_C}{\text{MSE}} = \frac{199{,}27}{9{,}7652} = 20{,}41 \geq 4{,}24 = f_{0.05,1,25}$$

$$F_{12-45} = \frac{SS_{12-45}/df_C}{\text{MSE}} = \frac{1100{,}26}{9{,}7652} = 112{,}67 \geq 4{,}24.$$

Iz ovoga zaključujemo kako se hipoteze H_0 navedene ranije trebaju odbaciti na razini značajnosti $5\ \%$.

Polinomijalni kontrasti

Polinomijalni kontrasti su vrsta ortogonalnih kontrasta s time da ako imamo k razina kategorijske varijable, tada sumu kvadrata tretmana razdijelimo u $k-1$ ortogonalnih polinomijalnih kontrasta s $k-1$ stupnjem slobode. Na primjer, ako imamo 5 uzoraka možemo razlomiti sume kvadrata tretmana na njegove linearne, kvadratne, kubne i kvadrične komponente te razmatrati oblik reakcije (ako postoji) na povećanje vrijednosti.

Razlike koje se testiraju (to jest kontrasti) prikazani su u matrici koeficijenata za različiti broj razina kategorijske varijable.[5]

$k = 2$

Linearno	-1	1

$k = 3$

Linearno	-1	0	1
Kvadratno	1	-2	1

$k = 4$

Linearno	-3	-1	1	3
Kvadratno	1	-1	-1	1
Kubno	-1	3	-3	1

$k = 5$

Linearno	-2	-1	0	1	2
Kvadratno	2	-1	-2	-1	2
Kubno	-1	2	0	-2	1
Kvadratično	1	-4	6	-4	1

Kontrasti prikazani u matricama su ortogonalni stoga su i linearno nezavisni. Linearni, kvadratni i ostali oblici iz polinomijalnih kontrasta prilagođavaju se prema vrijednostima razina kategorijskih varijabli.

Za Primjer 1 koristimo posljednju matricu za 5 uzoraka od kojih svaki ima po $n_i = 6, i = 1, \dots, k$ zapažanja te koristimo dobivene sume vrijednosti za pojedini uzorak iz Tablice 1.8.

Pokazat ćemo postupak na linearnom i kvadratnom kontrastu koje računamo prema Formuli (1.15) te iznose

$$C_{lin} = -2 \times 192{,}5 - 1 \times 241{,}4 + 0 \times 264{,}5 + 1 \times 246{,}6 + 2 \times 349{,}8 = 319{,}8$$

$$C_{kv} = 2 \times 192{,}5 - 1 \times 241{,}4 - 2 \times 264{,}5 - 1 \times 246{,}6 + 2 \times 349{,}8 = 67{,}6$$

Kako bismo izračunali testne statistike F_{lin} i F_{kv} potrebne su nam sume kvadrata za svaki kontrast SS_{lin} i SS_{kv} (računaju se po Formuli (1.17)) te srednje kvadratna pogreška tretmana $\mathrm{MSE} = 9{,}7652$ koju dobijemo u analizi varijance (vidi Primjer 1, Poglavlje 3).

$$SS_{lin} = \frac{{C_{lin}}^2}{6 \times [(-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2]} = \frac{102272{,}04}{6 \times 10} = 1704{,}534$$

$$SS_{kv} = \frac{{C_{kv}}^2}{6 \times [2^2 + (-1)^2 + (-2)^2 + (-1)^2 + 2^2]} = \frac{4569{,}76}{6 \times 14} = 54{,}4$$

Broj stupnjeva slobode za kontraste je $df_C = 1$, a za tretman $df = 25$ stoga u Tablici 4.5 tražimo kritičnu vrijednost F-testa koja iznosi $f_{0{,}05} = 4{,}24$ za $5\,\%$ signifikantnosti, odnosno $f_{0{,}01} = 7{,}77$ za $1\,\%$. Testne statistike za pojedini kontrast tada iznose

$$F_{lin} = \frac{SS_{lin}/df_C}{\mathrm{MSE}} = \frac{1704{,}534}{9{,}7652} = 174{,}55 \geq 4{,}24$$

$$F_{lin} = 174{,}55 \geq 7{,}77$$

$$F_{kv} = \frac{SS_{kv}/df_C}{\mathrm{MSE}} = \frac{54{,}4}{9{,}7652} = 5{,}57 \geq 4{,}24$$

$$F_{kv} = 5{,}57 < 7{,}77$$

Iz dobivenih podataka jasno se vidi kako se vrijednosti približno linearno povećavaju s malom zabilješkom zakrivljenosti.

Polinomijalni kontrasti izvrsni su za testiranje, na primjer, reakcije pacijenata na određenu dozu lijeka (vidi Primjer 2 u Poglavlju 3).

Sheffé-ov test

Sheffé-ov test koristi se za uspoređivanje kontrasta kada osoba koja vrši eksperiment nije odredila ili ne može odrediti kontrast od interesa unaprijed. Ukupna razina značajnosti svake usporedbe je najviše α. Provođenje Sheffé-ovog testa je složen proces baziran na kontrastima i najbolje ga je prepustiti računalnim postupcima kao što su oni unutar računalnih paketa. U mnogim praktičnim situacijama bit ćemo zainteresirani samo za usporedbe parova tretmana.

Korištenjem podataka iz Primjera 1 može se pretpostaviti hipoteza H_0: aritmetičke sredine u drugom, četvrtom trećem uzorku ne razlikuju se od aritmetičke sredine petog uzorka. Ova hipoteza može se iskazati sažeto kao

$$H_0: \frac{\mu_2 + \mu_4 + \mu_3}{3} - \mu_5 = 0.$$

Kako bi dobili koeficijente kontrasta, $\frac{\mu_2+\mu_4+\mu_3}{3} - \mu_5$ može se zapisati i kao $\frac{\mu_2}{3} + \frac{\mu_4}{3} + \frac{\mu_3}{3} - \mu_5$, tako da očekivanjima μ_2, μ_4, μ_3 i μ_5 prethode koeficijenti $c_2 = \frac{1}{3}, c_4 = \frac{1}{3}, c_3 = \frac{1}{3}$ i $c_5 = -1$ za koje vrijedi $\sum_{i=1}^{k} c_i = 0$.

Testna statistika, S, računa se pomoću formule

$$S = \frac{|C|}{\mathrm{SE}}, \tag{1.19}$$

gdje vrijedi

$$C = \sum_{i=1}^{k} c_i \bar{x}_i, \tag{1.20}$$

$$\mathrm{SE} = \sqrt{\mathrm{MSE} \sum_{i=1}^{k} \frac{{c_i}^2}{n_i}}. \tag{1.21}$$

Pomoću prethodnih formula dalje računamo

$$\mathrm{SE} = \sqrt{9{,}7652 \times \sum_{i=1}^{5} \frac{{c_i}^2}{n_i}} = \sqrt{9{,}7652 \times \left[\frac{\left(\frac{1}{3}\right)^2}{6} + \frac{\left(\frac{1}{3}\right)^2}{6} + \frac{\left(\frac{1}{3}\right)^2}{6} + \frac{1^2}{6}\right]} = \sqrt{9{,}7652 \times \frac{2}{9}}$$

$$\mathrm{SE} = 1{,}4731$$

$$C = \sum_{i=1}^{5} c_i \bar{x}_i = 0 \times 32{,}08 + \frac{1}{3} \times 40{,}23 + \frac{1}{3} \times 44{,}08 + \frac{1}{3} \times 41{,}1 - 1 \times 58{,}3$$

$$C = -16{,}4967$$

$$S = \frac{|-16{,}4967|}{1{,}4731} = 11{,}1986$$

Kritičnu vrijednost S_α računamo pomoću Formule (1.14), za $\alpha = 0{,}05$ ona iznosi

$$S_{0{,}05} = \sqrt{(k-1)F_{0.05,k-1,N-k}} = \sqrt{(5-1)F_{0.05,4,25}} = \sqrt{4 \times 2{,}76} = 3{,}32.$$

Budući da je $S = 11{,}1986 \geq 3{,}32 = S_{0{,}05}$ odbacujemo hipotezu $H_0: \frac{\mu_2+\mu_4+\mu_3}{3} - \mu_5 = 0$ što znači da se uzorci $2, 4$ i 3 statistički značajno razlikuju od uzorka 5.

Za $k = 5$ možemo promatrati i druge hipoteze poput

$$H_0: \mu_1 - \frac{\mu_2 + \mu_3 + \mu_4}{3} = 0,$$

$$H_0: \frac{\mu_1 + \mu_5}{2} - \frac{\mu_2 + \mu_3 + \mu_4}{3} = 0,$$

$$H_0: \frac{\mu_1 + \mu_4}{2} - \frac{\mu_2 + \mu_3}{2} = 0.$$

Poglavlje 2

Neparametrijski testovi

Ukoliko nije moguće pretpostaviti normalnost ili homogenost varijanci, tada se može upotrijebiti neparametrijska analiza varijance, za kojom slijedi neparametrijsko testiranje višestrukih usporedbi. Ukoliko ne vrijede te pretpostavke, dobiveni rezultati u parametrijskom testu su varljivi i zbog toga test nije robustan. Općenito neparametrijski testovi imaju manju snagu (definiramo ju kao vjerojatnost odbacivanja nulte hipoteze ako je ona zaista pogrešna) od prethodnih. No prednost im je jednostavnost primjene i brzina te mogućnost korištenja kod malih uzoraka koji obično nemaju normalnu raspodjelu. Ukoliko su podaci dani u obliku rangiranog niza tada gotovo i ne postoju druga mogućnost nego primjeniti ovakvu vrstu testiranja.

Neparametrijski analogon F-testu u analizi varijance je Kruskal-Wallis test koji se koristi za proučavanje očekivanja triju ili više populacija. Ova procedura testira nultu hipotezu H_0: distribucije svih k populacija su jednake, nasuprot alternativne hipoteze H_1: barem jedna distribucija se razlikuje od ostalih, to jest nemaju sve populacije jednako očekivanje. Više o Kruskal-Wallis testu može se pronaći u [6].

Promatrat ćemo situaciju kada je Kruskal-Wallis test proveden i nulta hipoteza je odbačena. U takvom slučaju želimo utvrditi među kojim uzorcima postoji razlika. Za tu svrhu navodimo slijedeći primjer čija je baza preuzeta iz [3].

Primjer 3

Na slučajan način odabrano je 45 pretilih odraslih osoba kojima je dijagnosticiran stvarni dijabetes s očitim simptomima (Grupa 3), kemijski dijabetes koji se pojavljuje samo u određenim situacijama poput stresa (Grupa 2) te osobe kojim nije dijagnosticiran dijabetes, to jest s standardnim vrijednostima (Grupa 1). U svakoj grupi nalazi se jednak broj osoba. Promatra se varijabla *glutest* koja daje mjeru intolerancije na glukozu pomoću testiranja krvi.

Grupa 1	Grupa 2	Grupa 3
398	465	1468
289	558	1487
345	503	714
349	540	1470
352	469	1113
367	486	972
365	568	854
393	527	1364
379	537	832
296	466	967
353	599	920
306	429	613
335	472	857
227	456	1373
360	517	1133

Želimo provjeriti postoje li značajne razlike među tim grupama, odnosno može li se na temelju intolerancije na glukozu donositi zaključak o tome u koju grupu dijabetesa osoba pripada ili uopće nema šećernu bolest. Uzimamo razinu značajnosti 5 %.

2.1. Neparametrijska višestruka usporedba

Neparametrijska višestruka usporedba nakon Kruskal-Wallis testa koja povlači paralelu s parametrijskim Tukey testom iz Sekcije 1.2 koristi rangirane sume umjesto aritmetičkih sredina. Takva vrsta testa naziva se Nemenyi test.

Najprije se sume rangova svakog uzoraka iz Kruskal-Wallis testa postave u rastućem poretku. Rangovi $R(x_{ij}),\ i = 1, \dots, k,\ j = 1, \dots n_i$ dodjeljeni su opservacijama x_{ij}. U Primjeru 3 za $k = 3$ i $n_i = 15$, $i = 1,2,3$ izračunate su sume rangova pomoću formule

$$R_i = \sum_{j=1}^{n_i} R(x_{ij}) \tag{2.1}$$

te su one u rastućem poretku jednake $R_1 = 120$, $R_2 = 345$, $R_3 = 570$. Parovi usporedbi, odnosno razlike tih suma su dane u Tablici 2.1 počevši s razlikom između najveće i najmanje sume, a ostale se izračunaju jednako kao i u Sekciji 1.2.

Usporedba	$R_i - R_j$
$3 - 1$	$570 - 120 = 450$
$3 - 2$	$570 - 345 = 225$
$2 - 1$	$345 - 120 = 225$

Tablica 2.1: Razlike suma rangova po parovima usporedbi

Standardna pogreška računa se formulom

$$\mathrm{SE} = \sqrt{\frac{n_i(n_i k)(n_i k + 1)}{12}}, \tag{2.2}$$

q-vrijednost je dana s

$$q = \frac{R_j - R_i}{\mathrm{SE}}, \tag{2.3}$$

(Nemeny, 1963; Wilcoxon i Wilcox, 1964)[6], a Studentizirani raspon (vidi Tablicu 4.2) koji se koristi je $q_{\alpha,\infty,k}$.[1] U testiranju iz Primjera 2, za $\alpha = 0.05$ i $k = 3$ kritična vrijednost iznosi $q_{0.05,\infty,3} = 3{,}314$ te računamo standardnu pogrešku i testne statistike za svaki par usporedbi.

$$\mathrm{SE} = \sqrt{\frac{15 \times (15 \times 3)(15 \times 3 + 1)}{12}} = 50{,}867$$

Usporedba	q
$3 - 1$	$\frac{450}{50{,}867} = 8{,}847$
$3 - 2$	$\frac{225}{50{,}867} = 4{,}423$
$2 - 1$	$\frac{225}{50{,}867} = 4{,}423$

Tablica 2.2: q-vrijednosti za parove usporedbi

Sve q-vrijednosti su veće od $q_{0.05,\infty,3} = 3{,}314$ stoga odbacujemo hipoteze o međusobnoj jednakosti distribucija triju populacija.

[6] Neki autori (npr. Miller, 1981) daju prednost ovom testu u ekvivalentnom obliku tako što uspoređuju razlike među aritmetičkim sredinama rangova radije nego njihove sume, u tom slučaju je standardna pogreška dana formulom $\mathrm{SE} = \sqrt{k(n_i k + 1)/12}$.

Ukoliko je poželjnija analogija sa Student-Newman-Keuls testom višestrukog uspoređivanja tada računamo standardnu pogrešku formulom

$$\mathrm{SE} = \sqrt{\frac{n_i(n_i p)(n_i p + 1)}{12}} \tag{2.4}$$

te je $q_{\alpha,\infty,p}$ kritična vrijednost iz Tablice 4.2.[7]

Obje procedure zahtjevaju jednak broj opservacija u svakom uzorku, no ukoliko to nije slučaj tada se može koristiti procedura navedena u Sekciji 2.2 ili dosta snažniji test u kojem Dunn (1964; Hollander i Wolfe, 1973) predlaže korištenje slijedećih formula za standardnu pogrešku i testnu statistiku

$$\mathrm{SE} = \sqrt{\frac{N(N+1)}{12}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}, \tag{2.5}$$

$$Q = \frac{\bar{R}_j - \bar{R}_i}{\mathrm{SE}} \tag{2.6}$$

gdje je $\bar{R}_i$ aritmetička sredina rangova iz i-tog uzorka, a N ukupan broj opservacija. Kritična vrijednost za takav test, $Q_{\alpha,k}$, dana je u Tablici 4.6.

Ukoliko se pojavljuju rangovi s jednakom vrijednošću varijable tada je slijedeća formula poboljšanje Formule (2.4).

$$\mathrm{SE} = \sqrt{\left(\frac{N(N+1)}{12} - \frac{\sum t}{12(N-1)}\right)\left(\frac{1}{n_i} + \frac{1}{n_j}\right)} \tag{2.7}$$

gdje se $\sum t$ koristi u Kruskal-Wallis testu kada su prisutni takvi rangovi te je definirana formulom

$$\sum t = \sum_{i=1}^{m} (t_i^3 - t_i) \tag{2.8}$$

[7] Ako se koriste aritmetičke sredine rangova, tada se primjenjuje formula $\mathrm{SE} = \sqrt{p(n_i p + 1)/12}$.

gdje je t_i broj rangova s jednakom vrijednošću varijable u i-toj grupi takvih rangova, a m je ukupan broj tih vezanih rangova.[8]

Nakon odbacivanja nulte hipoteze pomoću Kruskal-Wallis testa, neparametrijska analiza može tražiti jednostrane ili dvostrane značajne razlike između jednog uzorka (postavljenog kao kontrolni) i svakog drugog uzorka u bazi podataka. To se čini na jednak način kao i u Sekciji 1.5, ali koristeći sume rangova po uzorcima umjesto aritmetičkih sredina uzoraka. Standardna pogreška računa se formulom

$$\mathrm{SE} = \sqrt{\frac{n_i(n_i k)(n_i k + 1)}{6}} \tag{2.9}$$

(Wilcoxon i Wilcox, 1964), a kritičnu vrijednost $q'_{\alpha,\infty,k}$ pronalazimo u Tablici 4.3. Ovaj test zahtjeva jednake veličine uzoraka. Ako svi n_i nisu jednaki može se ponovno primjeniti procedura koju predlaže Dunn (1964; Hollander i Wolfe, 1973). Po njemu, koriste se Formule (2.5), (2.6) i (2.7) samo što se ovdje j-ti uzorak proglašava kontrolnim. Kritična vrijednost u ovom testu, koja može biti za jednostrani ili dvostrani test, je $Q'_{\alpha,k}$ (Tablica 4.7). [1]

2.2. Neparametrijski višestruki kontrasti

Višestruki kontrasti, prikazani u Sekciji 1.7, mogu biti testirani i neparametarski koristeći H-statistiku dobivenu Kruskal-Wallis testom umjesto F-statistike. Prema Formuli (1.19) računamo testnu statistiku s razlikom u definiciji kontrasta

$$C = \sum_{i=1}^{k} c_i \bar{R}_i \tag{2.10}$$

[8] Detalje pogledati u [1], str. 198.

i standardne pogreške

$$\mathrm{SE} = \sqrt{\left(\frac{N(N+1)}{12}\right)\left(\sum_{i=1}^{k}\frac{{c_i}^2}{n_i}\right)}, \qquad (2.11)$$

odnosno

$$\mathrm{SE} = \sqrt{\left(\frac{N(N+1)}{12} - \frac{\sum t}{12(N-1)}\right)\left(\sum_{i=1}^{k}\frac{{c_i}^2}{n_i}\right)} \qquad (2.12)$$

ukoliko postoje rangovi s jednakim vrijenostima zavisne varijable. Kritična vrijednost za ovaj test je $\sqrt{H_{\alpha,n_1,n_2,\dots}}$, koristeći tablicu iz [1], App104-105 za dobivanje kritične vrijednosti od H. Ukoliko je potrebna kritična vrijednost koja nije u tablici, tada se može koristiti $\chi^2_{\alpha,k-1}$.[1]

Poglavlje 3

Primjeri

Programski paket SAS pruža nam niz mogućnosti za preciznu analizu podataka, stoga ćemo na tri primjera pokazati sve što je dosada obrađeno u ovom radu.

Primjer 1

Postavljamo testne hipoteze:

H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$

H_A: Očekivanja koncentracije stroncija nisu jednaka u svim vodama

SAS kod:

```
data voda;
input  vrsta stroncij;
cards;
1 28.2
1 33.2
1 36.4
1 34.6
1 29.1
1 31.0
2 39.6
2 40.8
2 37.9
2 37.1
2 43.6
2 42.4
3 46.3
```

```
3 42.1
3 43.5
3 48.8
3 43.7
3 40.1
4 41.0
4 44.1
4 46.4
4 40.2
4 38.6
4 36.3
5 56.3
5 54.1
5 59.4
5 62.7
5 60.0
5 57.3
;
proc print;
run;
proc means data=voda;
class vrsta;
var stroncij;
run;
```

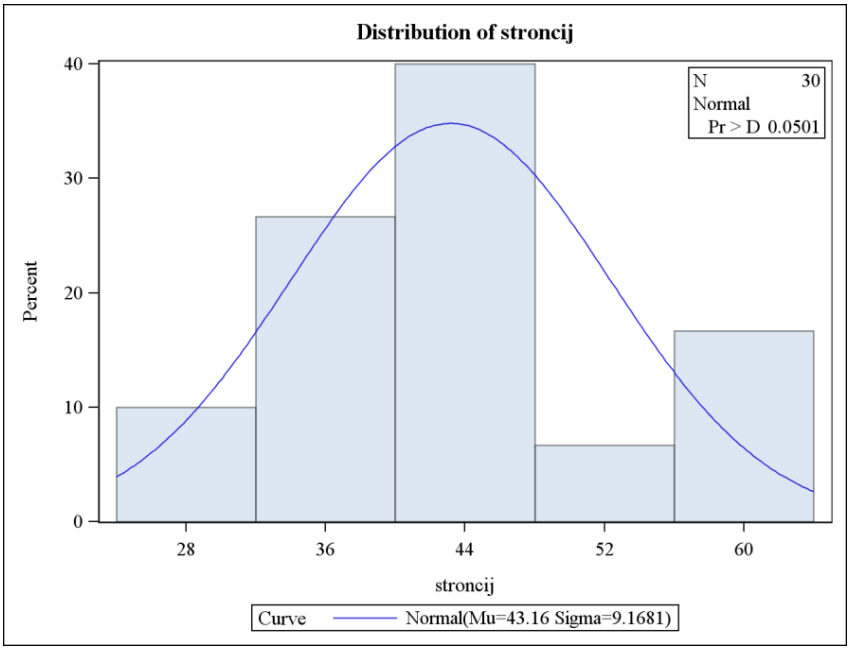
Analysis Variable : stroncij						
vrsta	**N Obs**	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**
1	6	6	32.0833333	3.2052561	28.2000000	36.4000000
2	6	6	40.2333333	2.5303491	37.1000000	43.6000000
3	6	6	44.0833333	3.0805303	40.1000000	48.8000000
4	6	6	41.1000000	3.6660606	36.3000000	46.4000000
5	6	6	58.3000000	3.0364453	54.1000000	62.7000000

Tablica 3.1: Deskriptivna statistika za varijablu *stroncij* (ispis iz SASa)

Predmet istraživanja je koncentracija stroncija u 5 vrsta vode, stoga varijabla *stroncij* predstavlja zavisnu varijablu te se promatra po varijabli *vrsta*.

Testiramo pretpostavke koje moraju biti zadovoljene za analizu varijance pomoću koda:

```
proc univariate data=voda;
   histogram stroncij / normal;
   inset n normal(ksdpval) /  pos = ne;
run;
```

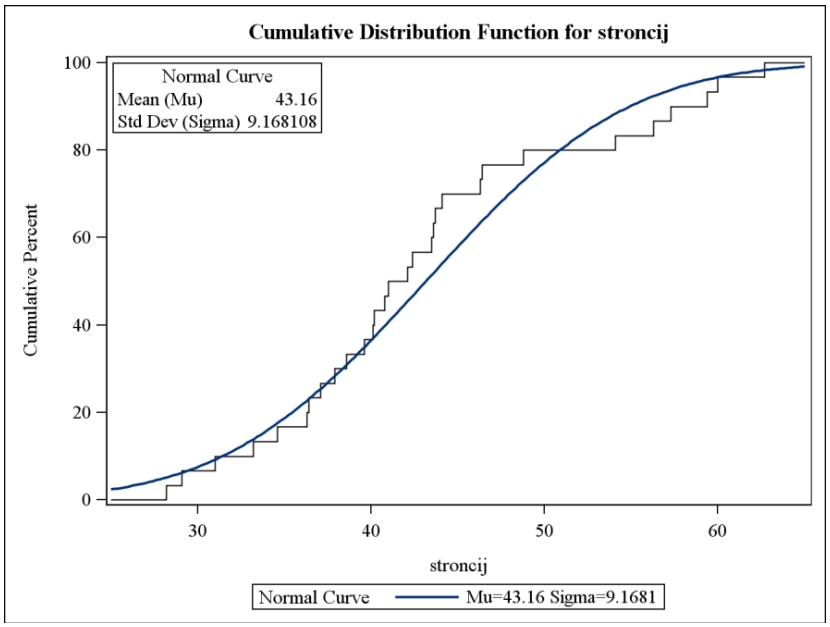


Slika 3.1: Histogram i procjenjena normalna distribucija za varijablu *stroncij* (ispis iz SASa)

Goodness-of-Fit Tests for Normal Distribution				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.15916827	**Pr > D**	0.050
Cramer-von Mises	**W-Sq**	0.11689554	**Pr > W-Sq**	0.066
Anderson-Darling	**A-Sq**	0.67928991	**Pr > A-Sq**	0.072

Tablica 3.2: Testovi normalnosti za varijablu *stroncij* (ispis iz SASa)

```
ods graphics on;
proc univariate data=voda noprint;
   cdf stroncij / normal;
   inset normal(mu sigma);
run;
```



Slika 3.2: Grafički prikaz kumulativnih frekvencija za varijablu *stroncij* (ispis iz SASa)

Prema histogramu i grafu kumulativnih frekvencija možemo zaključiti da podaci lagano slijede normalnu distribuciju. Prema p-vrijednostima procedura za testiranje normalnosti podataka ne možemo odbaciti hipotezu o normalnosti podataka.

Najvažnija pretpostavka, bez koje nema smisla provoditi analizu, je jednakost varijanci uzoraka. SAS nudi niz testova:

```
proc anova data=voda;
class vrsta;
model stroncij=vrsta;
means vrsta/HOVTEST=LEVENE;
means vrsta/HOVTEST=BARTLETT;
means vrsta/HOVTEST=OBRIEN;
means vrsta/HOVTEST=BF;
run;
```

Levene's Test for Homogeneity of stroncij Variance ANOVA of Squared Deviations from Group Means					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
vrsta	4	106.0	26.5026	0.36	0.8320
Error	25	1820.9	72.8350		

Tablica 3.3: Laveneov test homogenosti varijance za varijablu *stroncij* (ispis iz SASa)

U tablici je prikazan Levene test koji je snažniji od ostalih, stoga s velikom značajnošću možemo zaključiti kako su varijance za uzorke jednake.

Provodimo ANOVA-u pomoću

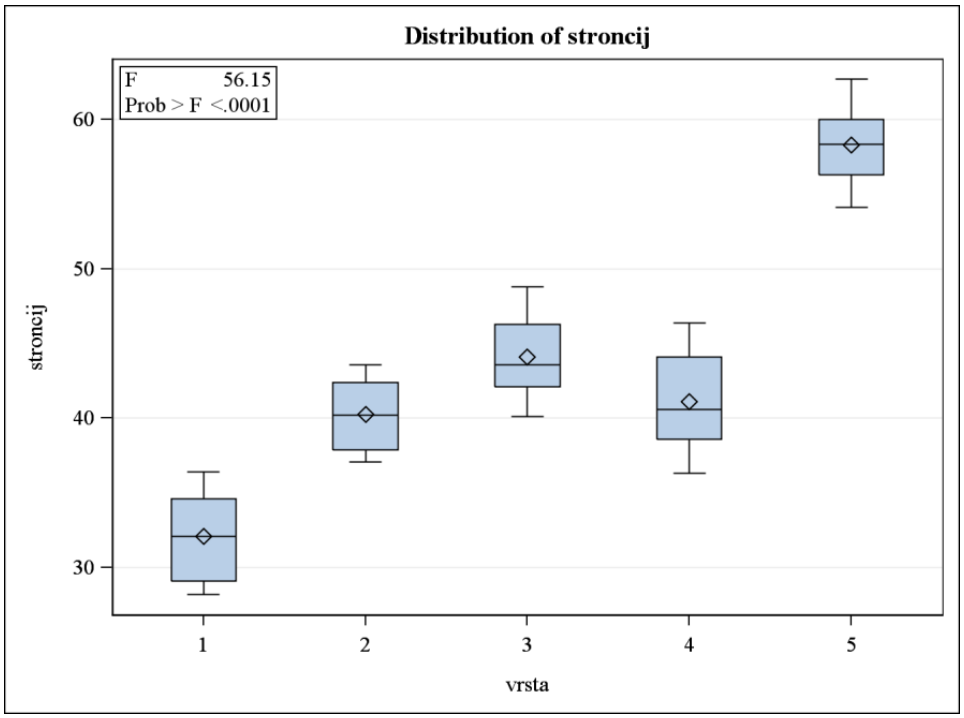
```
proc anova data=voda;
class vrsta;
model stroncij=vrsta;
run;
```

Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	4	2193.442000	548.360500	56.15	<.0001
Error	25	244.130000	9.765200		
Corrected Total	29	2437.572000			

Tablica 3.4: Rezultati ANOVA-e za varijablu *stroncij* (ispis iz SASa)

Iz tablice vidimo kako je p-vrijednost vrlo mala što znači da s velikom značajnošću odbacujemo danu hipotezu H_o, što se jasno vidi i iz Box-Plot dijagrama za varijablu *stroncij*. Također možemo vidjeti kako Rock River ima

najveću koncentraciju stroncija te se znatno razlikuje od svih drugih, dok Beaver Lake, Anger's Cove i Appletree Lake imaju slične aritmetičke sredine varijable *stroncij.*



Slika 3.3: Box-Whisker grafički prikaz varijable *stroncij* po analiziranim vrstama jezera (ispis iz SASa)

Budući da je hipoteza o jednakosti očekivanja odbačena ima smisla promatrati među kojim vodama je nastala značajnija razlika, stoga koristimo post-hoc testove.

Unutar gore navedene `proc anova` procedure dodajemo kodove za pojedine testove višestruke usporedbe.

LSD procedura dobijemo pomoću naredbe `means vrsta/LSD;`

Alpha	0.05
Error Degrees of Freedom	25
Error Mean Square	9.7652
Critical Value of t	2.05954
Least Significant Difference	3.7158

Means with the same letter are not significantly different.				
t Grouping		**Mean**	**N**	**vrsta**
	A	58.300	6	5
	B	44.083	6	3
	B			
C	B	41.100	6	4
C				
C		40.233	6	2
	D	32.083	6	1

Tablica 3.5: Rezultati Fisherovog LSD testa (ispis iz SASa)

Ovdje je zapisana t-vrijednost, to jest kritična vrijednost za 25 stupnjeva slobode i razinom značajnosti $\alpha = 0.05$ te LSD vrijednost (testna statistika za usporedbu značajnih razlika među vrstama voda).

LSD procedura postavila je Rock River i Grayson's Pound kao zasebne grupe koje se u potpunosti razlikuju od ostalih. Vidimo kako koncentracija stroncija nije značajno različita između Beaver Lake i Appletree Lake, to može dovesti do zaključka kako su u jezerima podjednake razine stroncija. Također su u zasebnu grupu izdvojeni Appletree Lake i Angler's Cove.

Ako želimo vidjeti pozdane intervale na razini značajnosti 5% za pojedine razlike, koristimo naredbu `means vrsta/LSD cldiff;.`

Comparisons significant at the 0.05 level are indicated by ***.				
vrsta Comparison	**Difference Between Means**	**95% Confidence Limits**		
5 - 3	14.217	10.501	17.932	***
5 - 4	17.200	13.484	20.916	***
5 - 2	18.067	14.351	21.782	***
5 - 1	26.217	22.501	29.932	***
3 - 5	-14.217	-17.932	-10.501	***
3 - 4	2.983	-0.732	6.699	
3 - 2	3.850	0.134	7.566	***
3 - 1	12.000	8.284	15.716	***
4 - 5	-17.200	-20.916	-13.484	***
4 - 3	-2.983	-6.699	0.732	
4 - 2	0.867	-2.849	4.582	
4 - 1	9.017	5.301	12.732	***
2 - 5	-18.067	-21.782	-14.351	***
2 - 3	-3.850	-7.566	-0.134	***
2 - 4	-0.867	-4.582	2.849	
2 - 1	8.150	4.434	11.866	***
1 - 5	-26.217	-29.932	-22.501	***
1 - 3	-12.000	-15.716	-8.284	***
1 - 4	-9.017	-12.732	-5.301	***
1 - 2	-8.150	-11.866	-4.434	***

Tablica 3.6: Rezultati Fisherovog LSD testa s 95% pouzdanim intervalima (ispis iz SASa)

Slijedeća vrlo važna procedura je Tukey procedura koja se u SAS-u izvršava naredbom `means vrsta/TUKEY;`, odnosno, ako želimo pouzdane interval kao i u prethodnom testu, koristimo naredbu `means vrsta/TUKEY cldiff;`.

Alpha	0.05
Error Degrees of Freedom	25
Error Mean Square	9.7652
Critical Value of Studentized Range	4.15336
Minimum Significant Difference	5.2986

Means with the same letter are not significantly different.			
Tukey Grouping	**Mean**	**N**	**vrsta**
A	58.300	6	5
B	44.083	6	3
B			
B	41.100	6	4
B			
B	40.233	6	2
C	32.083	6	1

Tablica 3.7: Rezultati Tukey testa (ispis iz SASa)

Comparisons significant at the 0.05 level are indicated by *.**				
vrsta Comparison	**Difference Between Means**	**Simultaneous 95% Confidence Limits**		
5 - 3	14.217	8.918	19.515	***
5 - 4	17.200	11.901	22.499	***
5 - 2	18.067	12.768	23.365	***
5 - 1	26.217	20.918	31.515	***
3 - 5	-14.217	-19.515	-8.918	***
3 - 4	2.983	-2.315	8.282	
3 - 2	3.850	-1.449	9.149	

Comparisons significant at the 0.05 level are indicated by ***.				
vrsta Comparison	**Difference Between Means**	**Simultaneous 95% Confidence Limits**		
3 - 1	12.000	6.701	17.299	***
4 - 5	-17.200	-22.499	-11.901	***
4 - 3	-2.983	-8.282	2.315	
4 - 2	0.867	-4.432	6.165	
4 - 1	9.017	3.718	14.315	***
2 - 5	-18.067	-23.365	-12.768	***
2 - 3	-3.850	-9.149	1.449	
2 - 4	-0.867	-6.165	4.432	
2 - 1	8.150	2.851	13.449	***
1 - 5	-26.217	-31.515	-20.918	***
1 - 3	-12.000	-17.299	-6.701	***
1 - 4	-9.017	-14.315	-3.718	***
1 - 2	-8.150	-13.449	-2.851	***

Tablica 3.8: Rezultati Tukey testa s 95% pouzdanim intervalima (ispis iz SASa)

Tukey test, za razliku od LSD testa, govori nam kako se testirana jezera i dana uvala ne razlikuju značajno po koncentraciji stroncija, dok Grayson's Pond i Rock River ponovno odudaraju od ostalih.

Slijedi nam Bonferroni procedura s kojom dolazimo do jednakog zaključka kao i kod Tukey testa. To nije uvijek slučaj jer je ovaj test nešto blaži što se vidi iz dobivenih pouzdanih intervala koji su veći nego u prethodnom testu. Bonferoni procedura je obično pogodnija za manje uzorke.

Koristimo SAS kodove `means vrsta/BON;` i `means vrsta/BON cldiff;` kojima dobijemo slijedeće rezultate

Alpha	0.05
Error Degrees of Freedom	25
Error Mean Square	9.7652
Critical Value of t	3.07820
Minimum Significant Difference	5.5536

Means with the same letter are not significantly different.

Bon Grouping	Mean	N	vrsta
A	58.300	6	5
B	44.083	6	3
B			
B	41.100	6	4
B			
B	40.233	6	2
C	32.083	6	1

Tablica 3.9: Rezultati Bonferroni testa (ispis iz SASa)

Comparisons significant at the 0.05 level are indicated by *.**

vrsta Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
5 - 3	14.217	8.663	19.770	***
5 - 4	17.200	11.646	22.754	***
5 - 2	18.067	12.513	23.620	***

Comparisons significant at the 0.05 level are indicated by ***.				
vrsta Comparison	**Difference Between Means**	**Simultaneous 95% Confidence Limits**		
5 - 1	26.217	20.663	31.770	***
3 - 5	-14.217	-19.770	-8.663	***
3 - 4	2.983	-2.570	8.537	
3 - 2	3.850	-1.704	9.404	
3 - 1	12.000	6.446	17.554	***
4 - 5	-17.200	-22.754	-11.646	***
4 - 3	-2.983	-8.537	2.570	
4 - 2	0.867	-4.687	6.420	
4 - 1	9.017	3.463	14.570	***
2 - 5	-18.067	-23.620	-12.513	***
2 - 3	-3.850	-9.404	1.704	
2 - 4	-0.867	-6.420	4.687	
2 - 1	8.150	2.596	13.704	***
1 - 5	-26.217	-31.770	-20.663	***
1 - 3	-12.000	-17.554	-6.446	***
1 - 4	-9.017	-14.570	-3.463	***
1 - 2	-8.150	-13.704	-2.596	***

Tablica 3.10: Rezultati Bonferroni testa s 95% pouzdanim intervalima (ispis iz SASa)

Drugačijim pristupom Student-Newman-Keuls procedura daje nam jednak zaključak kao i prethodne dvije procedure. Razlika je u tome što ovdje imamo 4 različite kritične vrijednosti (objašnjenje je dano u Sekciji 1.4) koje su prikazane tablicom (koristimo kod `means vrsta/SNK;`)

Number of Means	**2**	**3**	**4**	**5**
Critical Range	3.7157708	4.4939068	4.9626569	5.2986447

Alpha	0.05
Error Degrees of Freedom	25
Error Mean Square	9.7652

Means with the same letter are not significantly different.

SNK Grouping	**Mean**	**N**	**vrsta**
A	58.300	6	5
B	44.083	6	3
B			
B	41.100	6	4
B			
B	40.233	6	2
C	32.083	6	1

Tablica 3.11: Rezultati Student-Newman-Keuls testa
(ispis iz SASa)

Procedura objašnjena u Sekciji 1.5, Dunnett procedura, koristi kontrolni uzorak te se sve ostale aritmetičke sredine uspoređuju s njegovom aritmetičkom sredinom. Po defaultu, kod `means vrsta/DUNNETT;` daje usporedbe s aritmetičkom sredinom prve grupe.

Alpha	0.05
Error Degrees of Freedom	25
Error Mean Square	9.7652
Critical Value of Dunnett's t	2.60688
Minimum Significant Difference	4.7033

Comparisons significant at the 0.05 level are indicated by *.**

vrsta Comparison	**Difference Between Means**	**Simultaneous 95% Confidence Limits**		
5 - 1	26.217	21.513	30.920	***
3 - 1	12.000	7.297	16.703	***
4 - 1	9.017	4.313	13.720	***
2 - 1	8.150	3.447	12.853	***

Tablica 3.12: Rezultati Dunnett testa za višestruku usporedbu s uzorkom Grayson's Pond (ispis iz SASa)

Sve vrste testiranih voda imaju značajno različitu koncentraciju stroncija od jezerca Grayson's Pond.

Ovakva vrsta testa pogodna je ukoliko nam je neka od ovih voda po nečemu zanimljivija za istraživanje ili se po nečemu ističe te ju želimo usporediti s ostalima. Na primjer, ovdje imamo samo jednu rijeku – Rock River te želimo vidjeti je li se koncentracija stroncija u rijeci razlikuje od ostalih oblika voda. Tada postavljamo peti uzorak kao kontrolni, odnosno njegovu aritmetičku sredinu, i to s naredbom `means vrsta/DUNNETT ('5');`.

Comparisons significant at the 0.05 level are indicated by ***.				
vrsta Comparison	**Difference Between Means**	**Simultaneous 95% Confidence Limits**		
3 - 5	-14.217	-18.920	-9.513	***
4 - 5	-17.200	-21.903	-12.497	***
2 - 5	-18.067	-22.770	-13.363	***
1 - 5	-26.217	-30.920	-21.513	***

Tablica 3.13: Rezultati Dunnett testa za višestruku usporedbu s uzorkom Rock River (ispis iz SASa)

Iz dobivenih podataka zaključujemo kako se koncentracija u ovoj rijeci svakako razlikuje od ostalih.

Ostale usporedbe dobiju se analogno:

```
means vrsta/DUNNETT ('2');
```

Comparisons significant at the 0.05 level are indicated by ***.				
vrsta Comparison	**Difference Between Means**	**Simultaneous 95% Confidence Limits**		
5 - 2	18.067	13.363	22.770	***
3 - 2	3.850	-0.853	8.553	
4 - 2	0.867	-3.837	5.570	
1 - 2	-8.150	-12.853	-3.447	***

Tablica 3.14: Rezultati Dunnett testa za višestruku usporedbu s uzorkom Beaver Lake (ispis iz SASa)

```
means vrsta/DUNNETT ('3');
```

Comparisons significant at the 0.05 level are indicated by ***.				
vrsta Comparison	**Difference Between Means**	**Simultaneous 95% Confidence Limits**		
5 - 3	14.217	9.513	18.920	***
4 - 3	-2.983	-7.687	1.720	
2 - 3	-3.850	-8.553	0.853	
1 - 3	-12.000	-16.703	-7.297	***

Tablica 3.15: Rezultati Dunnett testa za višestruku usporedbu s uzorkom Angler's Cove (ispis iz SASa)

```
means vrsta/DUNNETT ('4');
```

Comparisons significant at the 0.05 level are indicated by ***.				
vrsta Comparison	**Difference Between Means**	**Simultaneous 95% Confidence Limits**		
5 - 4	17.200	12.497	21.903	***
3 - 4	2.983	-1.720	7.687	
2 - 4	-0.867	-5.570	3.837	
1 - 4	-9.017	-13.720	-4.313	***

Tablica 3.16: Rezultati Dunnett testa za višestruku usporedbu s uzorkom Appletree Lake (ispis iz SASa)

Kao što se iz prethodnih outputa vidi Dunnett test daje u ovom slučaju jednake zaključke kao i prethodna tri testa.

Procedura Scheffé uglavnom se koristi za kontraste, no može se provesti i ovdje te nam za dane podatke daje jednake zaključke iako radi veće greške Tipa II nego Tukey test.

```
means vrsta/SCHEFFE;
```

Alpha	0.05
Error Degrees of Freedom	25
Error Mean Square	9.7652
Critical Value of F	2.75871
Minimum Significant Difference	5.9933

Means with the same letter are not significantly different.

Scheffe Grouping	Mean	N	vrsta
A	58.300	6	5
B	44.083	6	3
B			
B	41.100	6	4
B			
B	40.233	6	2
C	32.083	6	1

Tablica 3.17: Rezultati Schefféovog testa (ispis iz SASa)

```
means vrsta/SCHEFFE cldiff;
```

Comparisons significant at the 0.05 level are indicated by ***.				
vrsta Comparison	**Difference Between Means**	**Simultaneous 95% Confidence Limits**		
5 - 3	14.217	8.223	20.210	***
5 - 4	17.200	11.207	23.193	***
5 - 2	18.067	12.073	24.060	***
5 - 1	26.217	20.223	32.210	***
3 - 5	-14.217	-20.210	-8.223	***
3 - 4	2.983	-3.010	8.977	
3 - 2	3.850	-2.143	9.843	
3 - 1	12.000	6.007	17.993	***
4 - 5	-17.200	-23.193	-11.207	***
4 - 3	-2.983	-8.977	3.010	
4 - 2	0.867	-5.127	6.860	
4 - 1	9.017	3.023	15.010	***
2 - 5	-18.067	-24.060	-12.073	***
2 - 3	-3.850	-9.843	2.143	
2 - 4	-0.867	-6.860	5.127	
2 - 1	8.150	2.157	14.143	***
1 - 5	-26.217	-32.210	-20.223	***
1 - 3	-12.000	-17.993	-6.007	***
1 - 4	-9.017	-15.010	-3.023	***
1 - 2	-8.150	-14.143	-2.157	***

Tablica 3.18: Rezultati Schefféovog testa s 95% pouzdanim intervalima (ispis iz SASa)

U Poglavlju 1 obrađena su dva ortogonalna kontrasta (usporedba Grayson's Pond-a s Beaver Lake te drugi koji daje usporedbu te dvije vode sa zadnje dvije,

to jest s Appletree Lake i Rock River) koja ćemo sada testirati pomoću programskog paketa SAS.

```
proc glm data = voda;
  class vrsta;
  model stroncij = vrsta;
  means vrsta /deponly;
  contrast 'Compare 1st & 2nd grp' vrsta 1 -1 0 0 0;
  contrast 'Compare 1st & 2nd with 4rd & 5th grp' vrsta 1 1 0 -1 -1;
run;
quit;
```

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
Compare 1st & 2nd grp	1	199.267500	199.267500	20.41	<.0.001
Compare 1st & 2nd with 4rd & 5th grp	1	1100.260417	1100.260417	112.67	<.0001

Tablica 3.19: Kontrasti – usporedba Grayson's Pond-a s Beaver Lake i usporedba Grayson's Pond-a i Beaver Lake s Appletree Lake i Rock River (ispis iz SASa)

Obje hipoteze se odbacuju kao što smo već vidjeli i u Sekciji 1.7.

Sada ćemo promatrati trend prema kojem se ponašaju podaci pomoću polinomijalnih kontrasta.

```
Proc glm ;
Class vrsta;
Model stroncij=vrsta;
contrast 'linear' vrsta -2 -1 0 1 2;
contrast 'quadratic' vrsta 2 -1 -2 -1 2;
contrast 'cubic' vrsta -1 2 0 -2 1;
contrast 'quartic' vrsta 1 -4 6 -4 1;
run;
```

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
linear	1	1704.534000	1704.534000	174.55	<.0001
quadratic	1	54.401905	54.401905	5.57	0.0264
cubic	1	359.660167	359.660167	36.83	<.0001
quartic	1	74.845929	74.845929	7.66	0.0105

Tablica 3.20: Polinomijalni kontrasti za bazu *voda* (ispis iz SASa)

Na temelju dobivenih rezultata možemo zaključiti kako podacima najviše odgovara linearan rast i kubični. No ovdje nije pogodno koristiti polinomijalne kontraste, više nas zanimaju prethodni post-hoc testovi jer nam objašnjavaju među kojim vodama postoji značajnija razlika u koncentraciji stroncija, nego poredak tih voda.

Primjer 2

Uzet ćemo primjer iz [2] (str. 94) kako bismo demonstrirali očiti trend. Promatrana je ječmena trava koja je izložena različitim dozama olova. Uzete su po četiri biljke u svakoj parceli koje imaju najdulji korijen te je njihov prosjek prikazan u tablici za tri mjerenja.

Doza (*ppm*[9])	Duljina (*mm*)		
15	12,5	10,3	8,3
30	11,0	4,8	9,3
45	2,5	1,3	1,0

[9] eng. parts per million upotrebljava se za izražavanje koncentracije u relativnim proporcijama, $1\ \text{ppm} = 1\ \text{mg/l}$

SAS kod:

```
data olovo;
input doza mm;
cards;
15 12.3
15 10.3
15 8.3
30 11
30 4.8
30 9.3
45 2.5
45 1.3
45 1
;
proc print;
run;

Proc glm data = olovo;
Class doza;
Model mm=doza;
contrast 'linear' br -1 0 1;
contrast 'quadratic' br 1 -2 1;
run;
```

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
linear	1	113.5350000	113.5350000	22.87	0.0031
quadratic	1	11.6805556	11.6805556	2.35	0.1760

Tablica 3.21: Polinomijalni kontrasti za bazu *olovo* (ispis iz SASa)

Iz dobivenih p-vrijednosti vidimo kako porastom doze (koncentracije olova u biljkama) imamo linearnu reakciju s malom zakrivljenošću. Pogledamo li prikupljene podatke, to znači kako većom izloženošću olovu, biljke imaju manji korijen.

Primjer 3

Za navedene podatke u Poglavlju 2 najprije ćemo provjeriti jesu li podaci normalno distribuirani i imaju li homogene varijance kako bi odlučili hoćemo li koristiti parametrijsko ili neparametrijsko testiranje.

```
data diabetes;
input  grupa glutest;
cards;
1 398
1 289
1 345
1 349
1 352
1 367
1 365
1 393
1 379
1 296
1 353
1 306
1 335
1 227
1 360
2 465
2 558
2 503
2 540
2 469
2 486
2 568
2 527
2 537
2 466
2 599
2 429
2 472
2 456
2 517
3 1468
3 1487
3 714
3 1470
3 1113
3 972
3 854
3 1364
3 832
3 967
3 920
3 613
3 857
3 1373
```

```
3 1133
;
proc print; run;
proc univariate data=diabetes;
   histogram glutest / normal;
   inset n normal(ksdpval) /  pos = ne;
run;
```

Goodness-of-Fit Tests for Normal Distribution				
Test	**Statistic**		**p Value**	
Kolmogorov-Smirnov	**D**	0.22459998	**Pr > D**	<0.010
Cramer-von Mises	**W-Sq**	0.51633410	**Pr > W-Sq**	<0.005
Anderson-Darling	**A-Sq**	2.92170737	**Pr > A-Sq**	<0.005

Tablica 3.22: Testovi za provjeru normalnosti podataka (ispis iz SASa)

```
proc anova data=diabetes;
class grupa;
model glutest=grupa;
means grupa/HOVTEST=LEVENE;
run;
```

Levene's Test for Homogeneity of glutest Variance ANOVA of Squared Deviations from Group Means					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
grupa	2	6.109E10	3.054E10	19.15	<.0001
Error	42	6.699E10	1.5949E9		

Tablica 3.23: Rezultati ANOVA-e (ispis iz SASa)

Vidi se kako niti jedna od pretpostavki za ANOVA-u nije zadovoljena stoga moramo koristiti neparametrijski Kruskal-Wallis test za testiranje jednakosti aritmetičkih sredina uzoraka.

```
PROC NPAR1WAY DATA=diabetes WILCOXON;
CLASS grupa;
EXACT WILCOXON / MC;
VAR glutest;
run;
```

Kruskal-Wallis Test	
Chi-Square	39.1304
DF	2
Pr > Chi-Square	<.0001

Tablica 3.24: Rezultat Kruskal-Wallis testa (ispis iz SASa)

Nulta hipoteza $H_0: \mu_1 = \mu_2 = \mu_3$ je odbačena s velikom razinom značajnosti, stoga nastavljamo promatrati koje su grupe međusobno različite.

Prikazat ćemo najprije rangove dodijeljene svakoj zapaženoj vrijednosti glutesta jer nam je to potrebno kad se radi analiza bez SAS-a.

```
PROC RANK DATA=diabetes OUT=diabetesrnks;
VAR glutest;
ranks glutestRank;
run;

PROC PRINT DATA=diabetesrnks;
title "Rangiranje glutesta";
run;
```

Obs	grupa	glutest	glutestRank
1	1	398	15
2	1	289	2
3	1	345	6
4	1	349	7
5	1	352	8
6	1	367	12
7	1	365	11
8	1	393	14
9	1	379	13
10	1	296	3

Obs	grupa	glutest	glutestRank
11	1	353	9
12	1	306	4
13	1	335	5
14	1	227	1
15	1	360	10
16	2	465	18
17	2	558	28
18	2	503	23
19	2	540	27
20	2	469	20
21	2	486	22
22	2	568	29
23	2	527	25
24	2	537	26
25	2	466	19
26	2	599	30
27	2	429	16
28	2	472	21
29	2	456	17
30	2	517	24
31	3	1468	43
32	3	1487	45
33	3	714	32
34	3	1470	44
35	3	1113	39
36	3	972	38
37	3	854	34
38	3	1364	41
39	3	832	33
40	3	967	37
41	3	920	36
42	3	613	31

Obs	grupa	glutest	glutestRank
43	3	857	35
44	3	1373	42
45	3	1133	40

Tablica 3.25: Rezultat rangirana varijable *glutest* (ispis iz SASa)

Kruskal-Wallis test odbacio je nultu hipotezu te se sada provode post-hoc testovi. Budući da su veličine uzoraka jednake korist ćemo Nemenyi test pomoću macro naredbe navedene u [8]. Pozivamo ju u SAS-u s naredbom

```
%LET NUMGROUPS=3;
%LET DATANAME=diabetes;
%LET OBSVAR=glutest;
%LET GROUP=grupa;
%LET ALPHA=0.05;

ODS HTML STYLE=STATISTICAL;
   %KW_MC(source=&DATANAME, groups=&NUMGROUPS, obsname=&OBSVAR, gpname=&GROUP,
sig=&alpha);
   run;
ods HTML close;
```

(U macro naredbi KW_MC uvrštene su i prethodne naredbe za Kruskal-Wallis test i rangove.)

```
Group sample sizes are equal. Performed Nemenyi test, alpha=0.05

Comparison group = grupa

Compare   Diff      SE        q         q(0.05)    Conclude
-----------------------------------------------------------
3 vs  1   450       50.87     8.85       3.315    Reject
3 vs  2   225       50.87     4.42       3.315    Reject
2 vs  1   225       50.87     4.42       3.315    Reject

Reference: Biostatistical Analysis, 4th Edition, J. Zar, 2010.
```

Tablica 3.26: Rezultat Nemenyi testa (ispis iz SASa)

Dobiven je output u kojem se vidi kako su sve tri grupe statistički značajno različite, to jest možemo testiranje glukoze uzeti kao određujući faktor u dijagnosticiranju dijabetesa.

Poglavlje 4

Prilozi

v	0,10	0,05	0,025	0,01	0,005	0,0025
1	3,078	6,314	12,706	31,821	63,657	127,321
2	1,886	2,920	4,303	6,965	9,925	14,089
3	1,638	2,353	3,182	4,541	5,841	7,453
4	1,533	2,132	2,776	3,747	4,604	5,598
5	1,476	2,015	2,571	3,365	4,032	4,773
6	1,440	1,934	2,447	3,143	3,707	4,317
7	1,415	1,895	2,365	2,998	3,499	4,029
8	1,397	1,860	2,306	2,896	3,355	3,833
9	1,383	1,833	2,262	2,821	3,250	3,690
10	1,372	1,812	2,228	2,764	3,169	3,581
11	1,363	1,796	2,201	2,718	3,106	3,497
12	1,356	1,782	2,179	2,681	3,055	3,428
13	1,350	1,771	2,160	2,650	3,012	3,372
14	1,345	1,761	2,145	2,624	2,977	3,326
15	1,341	1,753	2,131	2,602	2,947	3,286
16	1,337	1,746	2,120	2,583	2,921	3,252
17	1,333	1,740	2,110	2,567	2,898	3,222
18	1,330	1,734	2,101	2,552	2,878	3,197
19	1,328	1,729	2,093	2,539	2,861	3,174
20	1,325	1,725	2,086	2,528	2,845	3,153
21	1,323	1,721	2,080	2,518	2,831	3,135
22	1,321	1,717	2,074	2,508	2,819	3,119
23	1,319	1,714	2,069	2,500	2,807	3,104
24	1,318	1,711	2,064	2,492	2,797	3,091
25	1,316	1,708	2,060	2,485	2,787	3,078
26	1,315	1,706	2,056	2,479	2,779	3,067
27	1,314	1,703	2,052	2,473	2,771	3,057
28	1,313	1,701	2,048	2,467	2,763	3,047
29	1,311	1,699	2,045	2,462	2,756	3,038
30	1,310	1,697	2,042	2,457	2,750	3,030

Tablica 4.1 Kritične vrijednosti t-distribucije

v	k (p)	2	3	4	5
1		17,970	26,980	32,820	37,080
2		6,085	8,331	9,798	10,880
3		4,501	5,910	6,825	7,502
4		3,927	5,040	5,757	6,287
5		3,653	4,602	5,218	5,673
6		3,461	4,339	4,896	5,305
7		3,344	4,165	4,681	5,060
8		3,261	4,041	4,529	4,886
9		3,199	3,949	4,415	4,756
10		3,151	3,877	4,327	4,654
11		3,113	3,820	4,256	4,576
12		3,082	3,773	4,199	4,508
13		3,055	3,735	4,151	4,453
14		3,033	3,702	4,111	4,407
15		3,014	3,674	4,076	4,367
16		2,998	3,649	4,046	4,333
17		2,984	3,628	4,020	4,303
18		2,971	3,609	3,997	4,277
19		2,960	3,593	3,977	4,253
20		2,950	3,578	3,958	4,232
24		2,919	3,532	3,901	4,166
30		2,888	3,486	3,845	4,102
∞		2,772	3,314	3,633	3,858

Tablica 4.2 Kritične vrijednosti q-distribucije za $\alpha = 0{,}05$

v	k	2	3	4	5
5		2,57	3,03	3,29	3,48
6		2,45	2,86	3,10	3,26
7		2,36	2,75	2,97	3,12
8		2,31	2,67	2,88	3,02
9		2,26	2,61	2,81	2,95
10		2,23	2,57	2,76	2,89
11		2,20	2,53	2,72	2,84
12		2,18	2,50	2,68	2,81
13		2,16	2,48	2,65	2,78
14		2,14	2,46	2,63	2,75
15		2,13	2,44	2,61	2,73
16		2,12	2,42	2,59	2,71
17		2,11	2,41	2,58	2,69
18		2,10	2,40	2,56	2,68
19		2,09	2,39	2,55	2,66
20		2,09	2,38	2,54	2,65
24		2,06	2,85	2,51	2,61
30		2,04	2,32	2,47	2,58
∞		1,96	2,21	2,35	2,44

Tablica 4.3 Kritične vrijednosti za dvostrani Dunnettov test za $\alpha = 0{,}05$

df	0,10	0,05	0,025	0,01
1	55,80	225,00	900,00	5620,00
2	9,24	19,20	39,20	99,20
3	5,34	9,12	15,10	28,70
4	4,11	6,39	9,60	16,00
5	3,52	5,19	7,39	11,40
6	3,18	4,53	6,23	9,15
7	2,96	4,12	5,52	7,85
8	2,81	3,84	5,05	7,01
9	2,69	3,63	4,72	6,42
10	2,61	3,48	4,47	5,99
11	2,54	3,36	4,28	5,67
12	2,48	3,26	4,12	5,41
13	2,43	3,18	4,00	5,21
14	2,39	3,11	3,89	5,04
15	2,36	3,06	3,80	4,89
16	2,33	3,01	3,73	4,77
17	2,31	2,96	3,66	4,67
18	2,29	2,93	3,61	4,58
19	2,27	2,90	3,56	4,50
20	2,25	2,87	3,51	4,43
21	2,23	2,84	3,48	4,37
22	2,22	2,82	3,44	4,31
23	2,21	2,80	3,41	4,26
24	2,19	2,78	3,38	4,22
25	2,18	2,76	3,35	4,18
26	2,17	2,74	3,33	4,14
27	2,17	2,73	3,31	4,11
28	2,16	2,71	3,29	4,07
29	2,15	2,70	3,27	4,04
30	2,14	2,69	3,25	4,02

Tablica 4.4 Kritične vrijednosti F-distribucije za $\mathrm{DF} = 4$

df	0,10	0,05	0,025	0,01
1	39,90	161,00	648,00	4050,00
2	8,53	18,50	38,50	98,50
3	5,54	10,10	17,40	34,10
4	4,54	7,71	12,20	21,20
5	4,06	6,61	10,00	16,30
6	3,78	5,99	8,81	13,70
7	3,59	5,59	8,07	12,20
8	3,46	5,32	7,57	11,30
9	3,36	5,12	7,21	10,60
10	3,29	4,96	6,96	10,00
11	3,23	4,84	6,72	9,65
12	3,18	4,75	6,55	9,33
13	3,14	4,67	6,41	9,07
14	3,10	4,60	6,30	8,86
15	3,07	4,54	6,20	8,68
16	3,05	4,49	6,12	8,53
17	3,03	4,45	6,04	8,40
18	3,01	4,41	5,98	8,29
19	2,99	4,38	5,92	8,18
20	2,97	4,35	5,87	8,10
21	2,96	4,32	5,83	8,02
22	2,95	4,30	5,79	7,95
23	2,94	4,28	5,75	7,88
24	2,93	4,26	5,72	7,82
25	2,92	4,24	5,69	7,77
26	2,91	4,23	5,66	7,72
27	2,90	4,21	5,63	7,68
28	2,89	4,20	5,61	7,64
29	2,89	4,18	5,59	7,60
30	2,88	4,17	5,57	7,56

Tablica 4.5 Kritične vrijednosti F-distribucije za $\mathrm{DF} = 1$

k	α	0,10	0,05	0,01
2		1,645	1,960	2,576
3		2,128	2,394	2,936
4		2,394	2,639	3,144
5		2,576	2,807	3,291
6		2,713	2,936	3,403
7		2,823	3,038	3,494
8		2,914	3,124	3,570
9		2,992	3,197	3,635
10		3,059	3,261	3,692
11		3,119	3,317	3,734
12		3,172	3,368	3,789
13		3,220	3,414	3,830
14		3,264	3,456	3,868
15		3,304	3,494	3,902
16		3,342	3,529	3,935
17		3,376	3,562	3,965
18		3,409	3,593	3,993
19		3,439	3,622	4,019
20		3,467	3,649	4,044
21		3,494	3,675	4,067
22		3,519	3,699	4,089
23		3,543	3,722	4,110
24		3,566	3,744	4,130
25		3,588	3,765	4,149

Tablica 4.6 Kritične vrijednosti Q-distribucije za neparametrijsko višestruko uspoređivanje

k	α	0,10	0,05	0,01
2		1,282	1,645	2,327
3		1,645	1,960	2,576
4		1,834	2,128	2,713
5		1,960	2,242	2,807
6		2,054	2,327	2,879
7		2,128	2,394	2,936
8		2,190	2,450	2,983
9		2,242	2,498	3,024
10		2,287	2,540	3,059
11		2,327	2,576	3,091
12		2,362	2,609	3,119
13		2,394	2,639	3,144
14		2,424	2,666	3,168
15		2,450	2,690	3,189
16		2,475	2,713	3,209
17		2,498	2,735	3,227
18		2,520	2,755	3,245
19		2,540	2,773	3,261
20		2,558	2,791	3,279
21		2,576	2,807	3,291
22		2,593	2,823	3,304
23		2,609	2,838	3,317
24		2,624	2,852	3,330
25		2,639	2,866	3,342

Tablica 4.7 Kritične vrijednosti Q'-distribucije za neparametrijsko višestruko uspoređivanje s kontrolnim uzorkom

U prethodnim tablicama navedeni su samo dijelovi potrebni u radu, za potpune tablice pogledati u [1].

Bibliografija

[1] J. H. Zar, *Biostatistical Analysis*, Department of Biological Sciences, Prentice Hall, New Jersey, 1999.

[2] T. Sparks, *Statistic in Ecotoxicology*, Institute of Terrestrial Ecology, Cambridgeshire, UK, 2000.

[3] J.J. Ashton, M. Beamish, B.L. Cphlen, P. G. Moell, J.P. Pope, *Logistic Regression, Examples, Using the SAS System*, SAS Institute Inc., 1995., str. 4-7.

[4] *Statistički testovi,* dostupno na http://web.math.pmf.unizg.hr/nastava/stat/files/StatTestovi2.pdf, slajdovi 53-60. (ožujak 2014.)

[5] D. A. Dickey, *Orthogonal polynomial Coefficients*, dostupno na http://www.stat.ncsu.edu/people/dickey/courses/st512/lab09/OrthPoly.html (travanj 2014.)

[6] W. Mendenhall, III, R. J. Beaver, B. M. Beaver, *Introduction to probability and Statistic*, 2013., str. 627-632.

[7] A. C. Elliott, L. S. Hynan, *A SAS Macro implementation of a Multiple Comparison post hoc test for a Kruskal-Wallis Analysis*, dostupno na http://www.alanelliott.com/kw/KW_MC.sas (svibanj 2014.)

Sažetak

U primjeni nailazimo na problem odlučivanja postoje li razlike između više populacija te između kojih od njih se prepoznaju značajnije razlike. Ukoliko se pokazalo da za tri ili više uzoraka ANOVA odbacuje nultu hipotezu, to jest da postoji značajna razlika među uzorcima tada koristimo post-hoc testove nakon analize varijance čije su pretpostavke o homogenosti varijance uzoraka, te normalnosti podataka zadovoljene. Takvu vrstu testiranja nazivamo parametrijsko testiranje i tu spada višestruko uspoređivanje pomoću Fisherove LSD procedure koja glasi kao najslabija, Tukey i Newman-Keuls testova, Dunnettovog testa za usporedbu s kontrolnim uzorkom te Bonferroni testa. Za posebnu vrstu višestrukog uspoređivanja, kontraste, najpogodniji je Shefféov test. Druga vrsta testiranja provedena u radu, je neparametrijski analogon ANOVA proceduri i njenim post-hoc testovima. U svim testovima višestrukog uspoređivanja, jednake veličine uzoraka su poželjne za maksimalnu snagu i robusnost.

U ovom radu prikazana je teorijska pozadina višestrukog uspoređivanja te su navedena tri primjera na kojim je jasno prikazana razlika među različitim testovima. Prikazano je kako se podaci mogu obrađivati sa i bez računalne podrške. Kako bi rezultati i zaključci bili što točniji korišten je programski paket SAS te su navedeni svi potrebni kodovi.

Summary

In the practice, there is a problem of deciding if there are differences between several populations and between which of them we can recognize significant differences. If it turns out that ANOVA rejects the null hypothesis for three or more treatments, in other words that there exists a significant difference between treatments, than we use post-hoc tests after the analysis of variance whose assumptions of homogeneity of variance of the treatments and population normality of data are met. This type of testing is called parametric tests, and it includes multiple comparisons using Fisher's LSD procedure which is known as the weakest, Tukey and Newman-Keuls tests, Dunnett's test for comparison with the control, and the Bonferroni test. For a special type of the multiple comparisons, contrasts, the best suited is the Sheffe's test. Another type of tests used in this thesis is the nonparametric analogue to the ANOVA procedure and its post-hoc tests. In all multiple comparison tests, equal sample sizes are desirable for maximum power and robustness.

This thesis shows the theoretical background of multiple comparisons and it contains three examples in which is clearly shown the difference between various types of tests. It shows how the data can be processed with and without computer support. Statistical analysis software – SAS was used to achieve the most accurate results and conclusions. All required codes are listed above.

Životopis

Rođena sam 2. listopada 1988. u Doboju, Bosna i Hercegovina. Osnovnu školu pohađala sam u Zagrebu, te sam potom upisala III. Ekonomsku školu. Nakon završetka srednje škole 2007. godine upisala sam Preddiplomski sveučilišni studij Matematika na Prirodoslovno-matematičkom fakultetu, Matematički odsjek u Zagrebu. Nakon stjecanja zvanja sveučilišne prvostupnice matematike, 2011. godine upisala sam Diplomski studij Matematička statistika na istom fakultetu.