



UNIVERZITET U NOVOM SADU
**Univerzitetski centar za primenjenu
statistiku**



Neke primene statistike u psihijatriji

Master rad

Mentor:

Prof. dr Dragić Banković

Kandidat:

Snežana Todorović

Novi Sad, 2013.

Univerzitet u Novom Sadu
Univerzitetski centar za primenjenu statistiku - UCPS
Ključna dokumentacijska informacija

Redni broj: RBR	
Identifikacioni broj: IBR	
Tip dokumentacije: TD	Monografska dokumentacija
Tip zapisa: TZ	Tekstualni štampani materijal
Vrsta rada (dipl., mag., dokt.): VR	Master rad
Ime i prezime autora: AU	Snežana Todorović
Mentor (titula, ime, prezime, zvanje): MN	prof. dr Dragić Banković
Naslov rada: NR	Neke primene statistike u psihijatriji
Jezik publikacije: JP	Srpski (latinica)
Jezik izvoda: JI	srp. / eng.
Zemlja publikovanja: ZP	Republika Srbija
Uže geografsko područje: UGP	Vojvodina
Godina: GO	2013
Izdavač: IZ	Autorski reprint
Mesto i adresa: MA	Novi Sad, Trg D. Obradovića 5

Fizički opis rada: FO	(8 / 61 / 33 / 2 / 6 / 23 / 0) (broj poglavlja / stranica / tabela / slika / grafikona / referenci / priloga)
Naučna oblast: NO	Matematika
Naučna disciplina: ND	Primenjena statistika (biomedicina)
Predmetna odrednica, ključne reči: PO	shizofrenija, citokini, neparametarski testovi, ROC krive, binarna-logistička regresija
UDK	
Čuva se: ČU	Biblioteka
Važna napomena: VN	
Izvod: IZ	Svrha ovog rada je da prikaže istraživanje o pojavi shizofrenije. Statističke analize su urađene pomoću nekoliko neparametarskih testova, t – testa za nezavisne uzorke, ANOVA analize, t-testa za ponovljena merenja, analize ROC krivih, binarne logističke regresije. Pokazalo se da je citokin 5 najznačajniji parametar.
Datum prihvatanja teme od strane NN veća: DP	
Datum odbrane: DO	
Članovi komisije: (ime i prezime / titula / zvanje / naziv organizacije / status) KO	predsednik: prof. dr Zorana Lužanin, redovni profesor, Prirodno-matematički fakultet, Novi Sad član: prof. dr Andreja Tepavčević, redovni profesor, Prirodno-matematički fakultet, Novi Sad mentor: prof. dr Dragić Banković, redovni profesor, Državni univerzitet u Novom Pazaru

University of Novi Sad
University Center for Applied Statistics
Key word documentation

Accession number: ANO	
Identification number: INO	
Document type: DT	Monograph documentation
Type of record: TR	Textual printed material
Contents code: CC	Master's thesis
Author: AU	Snežana Todorović
Mentor: MN	Phd Dragić Banković
Title: TI	Some applications of statistics in psychiatry
Language of text: LT	Serbian (Latin)
Language of abstract: LA	eng. / srp.
Country of publication: CP	Republic of Serbia
Locality of publication: LP	Vojvodina
Publication year: PY	2013
Publisher: PU	Author's reprint
Publication place: PP	Novi Sad, Trg D. Obradovića 5

Physical description: PD	(8 / 61 / 33 / 2 / 6 / 23 / 0) (chapters / page of theses / tables / pictures / graphics / references / enclosures)
Scientific field SF	Mathematics
Scientific discipline SD	Applied statistics (biomedicine)
Subject, Key words SKW	Schizophrenia, cytokines, non-parametric tests, ROC curves, binary logistic regression
UC	
Holding data: HD	Library
Note: N	
Abstract: AB	The purpose of this thesis is to illustrate research about schizophrenia. Statistical analysis were made using several non-parametric tests, t – test for independent samples, t – test for repeated measures, ANOVA analysis, ROC curve, binary logistic regression. In concluded, the most important parameter is cytokine 5.
Accepted on Scientific Board on: AS	
Defended: DE	
Thesis Defend Board: DB	<p>president: Phd Zorana Lužanin, full professor at Faculty of Science in Novi Sad</p> <p>member: Phd Andreja Tepavčević, full professor at Faculty of Science in Novi Sad</p> <p>mentor: Phd Dragić Banković, full professor at State University in Novi Pazar</p>

Predgovor

Istraživanje za ovaj master rad je sprovedeno na Klinici za psihijatriju u Kragujevcu. U master radu su se ispitale razlike serumskih koncentracija citokina između ispitanika kojima je postavljena dijagnoza akutnog psihotičnog poremećaja, ispitanika sa shizofrenijom i zdravih ispitanika. Posebno se ispitivao uticaj citokina na pojavu bolesti, kao i povezanost citokina sa drugim parametrima. Ispitana je veza pola i starosti sa dijagnozom i pojavom bolesti. Pored toga ispitan je uticaj biohemijskih parametara i parametara PANSS testa na bolest, kao i promena između dva merenja PANSS testom. Za ta ispitivanja korišćeni su sledeće statističke metode: MANOVA, Kruskal – Wallis-ov test, tabele kontigencije, ANOVA, Mann – Whitney-ev U test, T – test za nezavisne uzorke, T – test uparenih uzoraka, Pirsonov koeficijent korelacije. Pomoću binarne logističke regresije utvrđeno je od čega zavisi pojava bolesti, a markeri za pojavu bolesti ocenjeni su pomoću ROC krivih. Ispitana je promena između dva merenja citokina pomoću Wilcoxon-ovog testa.

Sve statističke metode koje se koriste u radu su teorijski prikazane. Statističke analize rađene su u Softveru za statističku obradu podataka IBM SPSS Statistics, verzija 20.0.

Sadržaj:

Predgovor	1
1. Uvod	1
2. Ciljevi i značaj studije	2
3. Statističke metode	2
3.1. Pronalaženje i uklanjanje pogrešnih podataka.....	3
3.2. Nedostajući podaci	4
3.3. Procena normalnosti raspodele	4
3.3.1. Univarijantna normalnost.....	4
3.3.1.1. Koeficijent asimetrije i spljoštenosti	4
3.3.1.2. Kolmogorov - Smirnov test	5
3.4. Multivarijantna analiza varijanse (MANOVA).....	6
3.5. Kruskal-Wallis-ov test	9
3.6. χ^2 test za nezavisnost $k \times 2$ tabele (tabele kontigencije)	10
3.7. ANOVA različitih grupa	11
3.8. Mann – Whitney U test	12
3.9. T – test za nezavisne uzorke	14
3.10. T – test uparenih uzoraka	15
3.11. Koeficijent korelacije	16
3.12. Binarna logistička regresija.....	18
3.13. ROC krive	20
3.14. Wilcoxon test ranga.....	23
4. Psihijatrijske metode	24
4.1. Uzorak.....	24
4.2. Instrumenti kliničko-psihijatrijske procene	24
5. Analiza podataka	28
6. Rezultati	28
6.1. Citokini i dijagnoza.....	29
6.2. Pol i dijagnoza	34
6.3. Starost i dijagnoza.....	35
6.4. Citokini i pojava bolesti.....	35
6.5. Pol i pojava bolesti.....	37
6.6. Starost i pojava bolesti	38
6.7. Biohemijski parametri i bolesni (akutni i hronični).....	39
6.8. PANSS i bolesni	43
6.9. Promene vrednosti parametara PANSS skale između dva merenja.....	44
6.10. Povezanost citokina i ostalih parametara	46
6.11. Od čega zavisi pojava bolesti?	48
6.12. Potencijalni markeri za pojavu bolesti.....	49
6.13. Promene citokina između dva merenja.....	52
6.14. Šta utiče na promene citokina između dva merenja?	53
7. Diskusija	57
8. Zaključak	59
Literatura	60

1. Uvod

Statistika je prvobitno proučavala masovne pojave u ljudskom društvu kroz prikupljanje, upoređivanje i tumačenje podataka o stanovništvu, imovini i slično. Još u IV veku p.n.e. Aristotel je uradio „popis” 158 gradova – država na teritoriji gde su živeli Grci, što bi mogao da bude začetak nauke o društvu – demografiji, na teritoriji današnje Evrope. Nakon toga, u XVII veku, u Nemačkoj je urađen komparativni opis država. U istom veku u Engleskoj, u vreme epidemije kuge, napravljen je popis umrlih i vrsta bolesti od kojih su preminuli, po parohijama. Sa današnje tačke gledišta, to je primena statistike u medicini. Oblast primene statistike se mnogo proširila, danas skoro da nema domena ljudske delatnosti bez manje ili veće primene statistike.

Psihijatrija je grana medicine usmerena na izučavanje, prepoznavanje i zbrinjavanje psihičkih poremećaja. Glavni zadaci ove discipline su da definiše i dijagnostikuje mentalne bolesti, da pronade načine za njihovo lečenje i razvije metode za otklanjanje njihovih uzroka i, konačno, koncipira mere prevencije. U ovom radu više reči će biti o shizofreniji.

Shizofrenija je hronični psihotični duševni poremećaj koji predstavlja klinički sindrom sastavljen od psihičkih simptoma i simptoma ponašanja. Heterogena klinička slika bolesti može uključivati: abnormalnosti percepcije u obliku halucinatornih doživljavanja, poremećaje procene određene sumanutim mislima i verovanjima, poremećaje toka misli koji se manifestuju kroz poremećaj govora, redukciju emocionalne i voljne sfere i osećaja zadovoljstva, različite kognitivne poteškoće – posebno one vezane za memoriju i izvršne funkcije, uočljivo ponašanje bazirano na psihotičnim iskustvima i oštećenom sastavu samokontrole i motoričke promene funkcionisanja [1].

Iako se danas shizofreniji pristupa kao jedinstvenom entitetu, ona se vrlo verovatno sastoji od niza heterogenih, još nedovoljno poznatih poremećaja. U prilog navedenom govore nepogrešivi podaci vezani za individualnu kliničku prezentaciju bolesti, reakcije na terapiju, toka bolesti, te kapaciteta obolelih osoba za socijalno funkcionisanje.

Etiologija¹ shizofrenije još nije dovoljno razjašnjena. Iako raspolažemo podacima iz niza istraživanja usmerenih na pronalaženje potencijalnih bioloških markera, pato-

¹ Uzrok

loških neuralnih mehanizama i genetskih aberacija koje bi bile odgovorne za javljanje psihotičnih poremećaja, danas još nemamo jasna i nepogrešiva saznanja o uzroku poremećaja.

Epidemiološka ispitivanja su ukazala na uticaj genetskih i sredinskih faktora u etiologiji shizofrenije. Genetski faktori mogu da dovedu do grešaka u razvoju moždanih struktura i u sinaptičkim vezama. S druge strane, tu je čitav niz sredinskih faktora koji mogu da dovedu do daljih oštećenja. Sredinski faktori dele se na primarno biološke i primarno nebiološke faktore. Primarni biološki faktori uključuju komplikacije trudnoće i porođaja. Primarni nebiološki faktori uključuju psihosocijalne stresore, kao što je život u visoko urbanim sredinama i disfunkcionalnom porodičnom okruženju [3].

2. Ciljevi i značaj studije

Glavni cilj ove studije je ispitati razlike serumskih koncentracija citokina između ispitanika kojima je postavljena dijagnoza akutnog psihotičnog poremećaja, ispitanika sa shizofrenijom i zdravih ispitanika, kao i moguća povezanost citokina sa rezultatima krvne slike i PANSS² testa. Cilj je ispitati i uticaj citokina na pojavu bolesti, kao i vezu pola i starosti sa dijagnozom i pojavom bolesti. Utvrdiće se od čega zavisi pojava bolesti i otkriti potencijalni markeri za pojavu bolesti. Ispitaće se promena između dva merenja citokina, kao i šta utiče na tu promenu i mogući odgovor na terapiju.

Značaj studije je u otkrivanju eventualne veze između psihopatologije i određenog imunskog statusa osobe (citokina), odnosno povezanosti sptomatologije, toka i odgovora na primenjenu terapiju sa serumskim koncentracijama citokina.

3. Statističke metode

U ovom radu, pre svega, rešen je problem pogrešnih podataka, nedostajućih podataka i normalnosti. Nakon toga primenjen je sledeće statističke metode: MANOVA, Kruskall – Wallis-ov test, tabele kontigencije, ANOVA, Mann – Whitney-ev U test, t – test za nezavisne uzorke, t – test uparenih uzoraka, Pirsonov koeficijent korelacije, binarna logistička regresija, ROC krive, Wilcoxon-ov test.

² Skala pozitivnih i negativnih simptoma

MANOVA i binarna logistička regresija su metode multivarijacione analize, dok su Kruskal – Wallis-ov test, χ^2 test za nezavisnost k x 2 tabele (tabele kontigencije), Mann – Whitney-ev U test, χ^2 test za nezavisnot 2 x 2 tabele (tabele kontigencije), ROC krive, Wilcoxon-ov test neparametarske statističke metode. ANOVA, t- test za nezavisne uzorke i t – test uparenih uzoraka su parametarske statističke metode.

Multivarijacione analize se zasnivaju na nizu pretpostavki, koje će detaljnije biti obrađene u okviru analize za koju su potrebne.

Neparametarske statističke metode nemaju pretpostavku o obliku raspodele, pa se nazivaju i metode slobodne raspodele. Primenuju se kada promenljiva nije normalno raspoređena. Otporne su na prisustvo netipičnih tačaka.

Za parametarske analize karakteristično je da pretpostavka o normalnosti treba da je zadovoljena. Osetljive su i na netipične tačke, ali SPSS u okviru ANOVA analize nudi mogućnost izbacivanja netipičnih tačaka.

Sve statističke analize radiće se u Softveru za statističku obradu podataka IBM SPSS Statistics, verzija 20.0.

3.1. Pronalaženje i uklanjanje pogrešnih podataka

Postupak čišćenja podataka obuhvata:

1. Traženje grešaka – proveravanje da li je vrednost neke promenljive izvan raspona mogućih vrednosti.
2. Pronalaženje i ispravljanje grešaka – pronalaženje grešaka tj. analiziranje slučaja u kome se greška javlja i ispravljanje ili brisanje te vrednosti.

Traženje grešaka kod kategorijalnih promenljivih se vrši preko provere minimuma i maksimuma frekvencija. Preko frekvencija se može proveriti i broj nedostajućih podataka u kategorijalnoj varijabli. Kod neprekidnih podataka koristi se deskriptivna statistika u kojoj se posmatra minimum, maksimum, aritmetička sredina i standardna devijacija. Pomoću deskriptivne statistike može se proveriti i broj nedostajućih podataka.

3.2. Nedostajući podaci

Metoda *pairwise* znači da će slučaj (osoba) biti isključen samo iz onih analiza za koje mu nedostaje neki neophodan podatak. Dakle, i takvi slučajevi će biti analizirani kad god je to moguće tj. kad god postoje podaci potrebni za datu analizu.

Metoda *EM* (*Expectation maximization*) je dostupna za slučajno raspoređene nedostajuće podatke. EM formira nedostajuće podatke matrica korelacije (ili kovarijanse) pod pretpostavkom oblika raspodele za delimično nedostajuće podatke, a zasnivajući zaključke o nedostajućim vrednostima na osnovu raspodele verovatnoća. To je iterativni postupak u dva koraka: očekivanje i maksimiziranje za svaku iteraciju. Prvi, E korak pronalazi uslovno očekivanje od „nedostajućih“ podataka, posmatrajući vrednosti i procenjujući parametare, kao što su korelacije. Drugo, M korak vrši procenu maksimalne verodostojnosti tako da se nedostajući podaci popunjavaju. Konačno, nakon što se postigne konvergencija, EM matrica varijanse – kovarijanse se može obezbediti i/ili popuniti podacima sačuvanim u skupu podataka [22].

3.3. Procena normalnosti raspodele

3.3.1. Univarijantna normalnost

3.3.1.1. Koeficijent asimetrije i spljoštenosti

Normalnost promenljivih se procenjuje statističkim metodama ili grafički. Dve komponente normalnosti su *koeficijent asimetrije* i *koeficijent spljoštenosti* [22]. Asimetrična promenljiva je promenljiva čija aritmetička sredina nije u centru raspodele. Spljoštenost predstavlja izduženost raspodele, može biti raspodela sa kratkim i debelim repom ili raspodela sa dugim, tankim repom. Promenljiva može imati značajan koeficijent asimetrije, koeficijent spljoštenosti, ili oba.

Postoje značajni testovi i za koeficijent asimetrije i za koeficijent spljoštenosti koji testiraju dobijene vrednosti protiv nulte hipoteze, koja tvrdi da su koeficijenti jednaki nu-

li. Standardna greška za koeficijent asimetrije je približno $SE(S) = \sqrt{\frac{6}{N}}$, gde je N obim uzorka. Dobijena vrednost koeficijenta asimetrije se poredi sa nulom koristeći $z = \frac{S-0}{SE(S)}$ i

S je vrednost koeficijenta asimetrije. Standardna greška za koeficijent spljoštenosti je približno $SE(K) = \sqrt{\frac{24}{N}}$ i dobijena vrednost koeficijenta spljoštenosti se poredi sa nulom koristeći z raspodelu, gde je $z = \frac{K-0}{SE(K)}$, gde je K koeficijent spljoštenosti.

Ako se koeficijent asimetrije i koeficijent spljoštenosti nalaze u segmentu $[-2,2]$, obeležje ima normalnu raspodelu.

3.3.1.2. Kolmogorov - Smirnov test

Test Kolmogorov-Smirnova je jedan od neparametarskih testova i primenjuje se za obeležja sa neprekidnom raspodelom. On ispituje da li je funkcija raspodele posmatranog obeležja $F(x)$ jednaka nekoj unapred zadatoj raspodeli (funkciji normalne raspodele) $F_0(x)$, što i predstavlja nultu hipotezu H_0 , nasuprot alternativne H_1 koja tvrdi da se funkcije raspodela razlikuju.

Neka $F_N^*(x)$ označava uzoračku funkciju raspodele za posmatrano obeležje. Test statistika je:

$$D_N = \sup_{x \in \mathbb{R}} |F_N^*(x) - F_0(x)|.$$

Kolmogorov je pokazao da za neprekidne funkcije raspodela važi, za svako $\lambda > 0$,

$$\lim_{N \rightarrow \infty} P\{\sqrt{N} D_N < \lambda\} = \lim_{N \rightarrow \infty} P\left\{D_N < \frac{\lambda}{\sqrt{N}}\right\} = K(\lambda) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 \lambda^2}.$$

Konvergencija je brza i aproksimacija je zadovoljavajuća za $N \geq 20$. Test statistika se poredi sa tabličnom vrednošću $D_N(\alpha)$ koja se dobija iz tablica Kolmogorov Smirnovove raspodele. Ako je $D_N > D_N(\alpha)$, nulta hipoteza se odbacuje. U suprotnom se prihvata.

3.4. Multivarijantna analiza varijanse (MANOVA)

Multivarijantna analiza varijanse je proširenje analize varijanse, koje se upotrebljava kada ima više od jedne zavisne promenljive. Trebalo bi da su te zavisne promenljive na neki način povezane ili da postoji neki konceptualni razlog zašto se razmatraju zajedno. MANOVA poredi te grupe i kazuje da li je verovatno da su srednje razlike u uticajima grupa na tu kombinaciju zavisnih promenljivih nastale slučajno. Da bi to uradila, MANOVA pravi novu zbirnu zavisnu promenljivu, dobijenu linearnom kombinacijom svih prvobitnih zavisnih promenljivih. Zatim analizira varijansu te nove, kombinovane zavisne promenljive. MANOVA kazuje da li je razlika srednjih vrednosti zavisne promenljive između populacija statistički značajna i daje univarijacionu analizu svake od zavisnih promenljivih zasebno.

Za jednofaktorsku MANOVA analizu potrebna je:

- jedna kategorijska, nezavisna promenljiva;
- dve ili više neprekidnih, zavisnih promenljivih.

MANOVA poredi dve ili više grupa po srednjim vrednostima određene kombinacije obeležja. Ispituje se nulta hipoteza da se srednje vrednosti određenog skupa zavisnih promenljivih ne menjaju u zavisnosti od nivoa grupišuće promenljive.

Pre sprovođenja MANOVA analize, potrebno je proveriti da li podaci zadovoljavaju sledeće pretpostavke:

Velicina uzorka

U svakoj grupi mora biti više slučajeva nego što ima zavisnih promenljivih. To je apsolutni minimum, a bilo bi dobro da ih ima još više. Kada je uzorak veliki, manje su i posledice kršenja nekih drugih pretpostavki (npr. normalnosti).

Normalnost raspodele

Ispitivanje značajnosti rezultata MANOVA analize zaspovano je na multivarijacionoj normalnoj raspodeli, ali je u praksi ova analiza prilično otporna na umereno narušavanje normalnosti (osim kada je ono posledica postojanja netipičnih tačaka). Treba proveriti

i univarijacionu normalnost i multivarijacionu normalnost. U slučaju da **univarijaciona normalnost** nije zadovoljena, potrebno je sve promenljive transformisati.

Postoji više metoda transformacije promenljivih. Jedna od metoda je **Tukey metoda za normalizaciju promenljivih**. Tukey za transformaciju koristi sledeći izraz $\frac{r - 1/3}{w + 1/3}$, gde je r rang, a w težina promenljive.

Multivarijaciona normalnost se proverava na osnovu Mardijinog testa [22]. Kada je vrednost Mardijine test statistike veća od 3 postoji blago odstupanje, a kada je veća od 10 postoji strašno odstupanje od multivarijacione normalne raspodele.

Netipične tačke

Netipične tačke (engl. outliers) su rezultati koji su mnogo iznad ili ispod ostalih rezultata. MANOVA je veoma osetljiva na netipične tačke. Nakon provere normalnosti, eventualne normalizacije promenljivih, potrebno je proveriti postojanje univarijacionih netipičnih tačaka, a i multivarijacionih netipičnih tačaka.

Tehnike za otkrivanje **univarijacionih netipičnih tačaka** su:

- Podaci se poređaju u varijacioni niz, a zatim se izračunaju aritmetičke sredine promenljivih bez 5% gornjih i 5% donjih slučajeva. Upoređuje se ta aritmetička sredina sa pravom aritmetičkom sredinom određenog obeležja. Ako se te dve aritmetičke sredine mnogo razlikuju, potrebno je proveriti da li su 5% gornjih i 5% donjih slučajeva netipične tačke.
- Dalja provera otkrivanja netipičnih tačaka vrši se pomoću pravougaonog dijagrama (engl. boxplot). Ako postoje netipične vrednosti, biće prikazane kao kružići označeni brojevima. SPSS smatra netipičnim vrednostima sve koje su od ivica pravougaonika udaljene više od 1.5 standardnu devijaciju. Ekstremne vrednosti, označene sa * su one koje su od ivica pravougaonika udaljene više od 3 standardne devijacije.

U slučaju postojanja netipičnih i ekstremnih vrednosti, moguće ih je ukloniti ili zameniti.

Multivarijacione netipične tačke su subjekti s čudnom kombinacijom dobijenih vrednosti raznih zavisnih promenljivih. Provera postojanja multivarijacionih netipičnih tačaka se vrši pomoću Mahalanobisovih distanci. Poredi se najveća vrednost Mahalanobi-

sovih distanci sa vrednošću χ^2 verovatnoće za $p = 0.001$ i broj stepeni slobode jednak broju promenljivih korišćenih u MANOVA analizi. Svi ispitanici za koje je vrednost Mahalanobisovih distanci veća od vrednosti χ^2 verovatnoće smatraju se multivarijacionim netipičnim tačkama [22].

Linearnost

Linearnost znači da između svih parova zavisnih promenljivih treba da postoji pravolinijski odnos. To se može oceniti na više načina, od kojih je najjednostavniji nacrtati matricu dijagrama raspršenja za sve zavisne promenljive, zasebno za grupe nezavisne promenljive.

Homogenost regresije

Ova pretpostavka je važna samo ako nameravamo da radimo postepenu analizu (engl. stepdown analysis). To se radi kada iz nekog teorijskog ili konceptualnog razloga rangirate zavisne promenljive [22].

Multikolinearnost i singularnost

MANOVA najbolje radi kada su zavisne promenljive umereno korelirane. U slučaju manjih korelacija, trebalo bi razmisliti o sprovođenju zasebne jednofaktorske analize varijanse za svaku zavisnu promenljivu. Kada su zavisne promenljive jako korelirane, to se naziva **multikolinearnost**. Do toga može doći kada je jedna od promenljivih kombinacija drugih. To bi bila **singularnost**, koju možete izbeći tako što ćete znati kakve su vam promenljive i kako se dobijaju rezultati. Najjednostavnija provera multikolinearnosti je provera korelacija zavisnih promenljivih. Razlog za brigu su korelacije oko 0.8 ili 0.9. Kada se pronade neka tolika korelacija, trebalo bi da ukloniti jedan od jako koreliranih parova zavisnih promenljivih ili ih objedinite u zajedničko merilo.

Homogenost matrica varijanse i kovarijanse

Test kojim se proverava homogenost matrica varijanse ili kovarijanse naziva se Box-ov test za jednakost kovarijansi matrice [22].

Multivarijacioni testovi značajnosti kojima se testira značajnost razlika u populacijama po linearnoj kombinaciji zavisnih promenljivih su: Wilks' Lambda, Hotelling's Trace, Pillai's Trace. Kada su podaci loši (mali uzorak, nejednak broj slučajeva po ćelijama, narušene pretpostavke), obično se navodi Pillai's Trace [22].

3.5. Kruskal-Wallis-ov test

Kruskal-Wallis-ov test naziva se i Kruskal-Wallis-ov H test. To je uopštenje Mann-Whitney-evog testa za k uzoraka. Rezultati se pretvaraju u rangove, pa se porede srednji rangovi svake grupe. To je analiza različitih grupa, pa u svakoj grupi moraju biti drugi ispitanici. Testira se nulta hipoteza

H_0 : Medijane su jednake u svih k populacija,

protiv alternativne H_1 : Medijane nisu jednake u svih k populacija.

Uslovi:

- Obim uzorka u svakoj grupi mora biti najmanje 5 da bi mogli primeniti χ^2 , iako obimi uzorka ne moraju biti jednaki.
- Obeležje nema normalnu raspodelu u svih k populacija.

Svi uzorci se uređuju po rastućem poretku i dodeljuje se rang svakom elementu svakog uzorka. Gde se vrednosti ponavljaju, koristi se srednja vrednost ranga tih vrednosti. Izračunava se rang sume za svaki od k uzoraka.

Neka imamo k uzoraka obima N_1, N_2, \dots, N_k , sa ukupnim obimom uzorka $N = N_1 + N_2 + \dots + N_k$. Neka su podaci iz svih uzoraka uzetih zajedno rangirani i neka su sume rangova iz k uzoraka R_1, R_2, \dots, R_k , respektivno. Ako se definiše test statistika sa

$$H = \left\{ \frac{12}{N(N+1)} \sum_{j=1}^k \frac{R_j^2}{N_j} \right\} - 3(N+1)$$

onda može da se pokaže da je raspodela funkcije H približno χ^2 - raspodela sa $k - 1$ stepeni slobode. Nulta hipoteza o jednakosti medijana se odbacuje kada je H vrednost veća od tablične vrednosti.

Kruskall-Wallis-ov H test je neparametarska alternativa jednofaktorskoj analizi varijanse različitih grupa.

3.6. χ^2 test za nezavisnost $k \times 2$ tabele (tabele kontigencije)

Neka su X i Y dve kategorijske promenljive, cilj ovog testa je da ispita da li su te dve promenljive nezavisne. Testira se nulta hipoteza

$$H_0 : X \text{ i } Y \text{ su nezavisni}$$

protiv alternativne $H_1 : X \text{ i } Y \text{ nisu nezavisni.}$

Neophodno je da svaki od k uzoraka bude dovoljno veliki. To podrazumeva da očekivana frekvencija svake ćelije bude bar jednaka 5 ili da 80% ćelija ima očekivane frekvencije 5 ili više.

Hipoteza se proverava na uzorku obima N tako što se evidentira koliko elemenata uzorka pripada svakoj od $k \times 2$ mogućih kombinacija obeležja X i Y. Te frekvencije se obeležavaju sa $f_{ij}, i = 1, \dots, k, j = 1, 2$ i prikazane su u tabeli 3.1. Poslednju vrstu i poslednju kolonu tabele 3.1 čine marginalne frekvencije za X i za Y. Marginalne frekvencije i ukupan obim uzorka se računaju koristeći sledeće formule:

$$f_{i.} = \sum_{j=1}^2 f_{ij}, \quad f_{.j} = \sum_{i=1}^k f_{ij}, \quad N = \sum_{i=1}^k f_{i.} = \sum_{j=1}^2 f_{.j}.$$

Tabela 3.1. $k \times 2$ tabela kontigencije

X/Y	y_1	y_2	Marginalne frekvencije za X
x_1	f_{11}	f_{12}	$f_{1.}$
\vdots	\vdots	\vdots	\vdots
x_i	f_{i1}	f_{i2}	$f_{i.}$
\vdots	\vdots	\vdots	\vdots
x_k	f_{k1}	f_{k2}	$f_{k.}$
Marginalne frekvencije za Y	$f_{.1}$	$f_{.2}$	N

Ukoliko su obeležja X i Y nezavisna, važi bi sledeće

$$p_{ij} = p(i)p(j) \Leftrightarrow \frac{f_{ij}}{N} = \frac{f_i}{N} \cdot \frac{f_j}{N},$$

odatle se dobijaju teorijske frekvencije f'_{ij} za svaku ćeliju u tabeli 3.1: $f'_{ij} = \frac{f_i \cdot f_j}{N}$.

Test statistika je

$$\chi^2 = \sum_{i=1}^k \sum_{j=1}^2 \frac{(f'_{ij} - f_{ij})^2}{f'_{ij}}.$$

Ovo se poredi sa vrednošću dobijenom iz tablica χ^2 sa $k - 1$ stepen slobode. Nulta hipoteza nezavisnosti obeležja X i Y se odbacuje ako je realizovana vrednost test statistike veća od tablične vrednosti.

Za veličinu uticaja (tj. jačinu veze između promenljivih) $k \times 2$ tabele koristi se Cramer's V pokazatelj koji u obzir uzima i broj stepeni slobode.

U specijalnom slučaju, $k = 2$, koristi se Yates-ova korekcija. Ona nadoknađuje precenjenu vrednost χ^2 koja je posledica malog broja dimenzija tabele 2×2 . Yates-ova korekcija oduzima 0.5 od vrednosti f_{11} i f_{22} , a dodaje 0.5 ćelijama f_{12} i f_{21} , tako da ukupne marginalne raspodele ostaju iste. Jačina veze između promenljivih određuje se Fi koeficijentom. To je koeficijent korelacije, pri čemu veći broj pokazuje jaču vezu između dve promenljive.

3.7. ANOVA različitih grupa

Jednofaktorska analiza varijanse znači da postoji jedna kategorijska promenljiva sa više od dve kategorije i jedna neprekidna promenljiva. Analiza varijanse je naziv dobila po tome što poredi varijansu između raznih grupa s tom promenljivošću unutar svake grupe. Izračuna se količnik F , koji predstavlja varijansu između grupa podeljenu varijansom unutar grupa. Količnik F ima Fišerovu raspodelu sa $k - 1, N - k$ stepeni slobode. Velika vrednost količnika F pokazuje da je veća promenljivost između grupa nego unutar svake grupe.

Statistički značajan pokazatelj F (tabela 3.2) govori da treba odbaciti nultu hipotezu, tj. tvrdnju da su srednje vrednosti obeležja u populaciji jednake. Međutim, to ne govori koje se grupe razlikuju, to se utvrđuje naknadnim analizama.

Tabela 3.2. ANOVA tabela

Izvor odstupanja	SS	df	MS	F	p
Između grupa	SS_{reg}	$k - 1$	$SS_{reg}/(k - 1)$	$\frac{SS_{reg}}{k - 1}$	p vrednost za F
Unutar grupa	SS_{rez}	$N - k$	$SS_{rez}/(N - k)$	$\frac{SS_{rez}}{N - k}$	
Ukupno	$SStot$	$N - 1$			ODLUKA Hipoteza H_0 se ...

U tabeli korišćene su sledeće oznake:

- SS_{reg} - suma kvadrata odstupanja među grupama,
- SS_{rez} - rezidualna suma kvadrata odstupanja,
- $SStot$ - ukupna suma kvadrata,
- k - broj grupa,
- N - obim uzorka.

To se može zapisati formulama:

$$\left. \begin{aligned} SS_{reg} &= \sum_{i=1}^k N_i (\bar{X}_i - \bar{X})^2 \\ SS_{rez} &= \sum_{i=1}^k \sum_{j=1}^{N_j} (X_{ij} - \bar{X}_i)^2 \\ SS_{tot} &= \sum_{i=1}^k \sum_{j=1}^{N_j} (X_{ij} - \bar{X})^2 \end{aligned} \right\} \Rightarrow SStot = SS_{reg} + SS_{rez}.$$

3.8. Mann – Whitney U test

Mann – Whitney U test se koristi za ispitivanje razlika između dve nezavisne populacije. Ovaj test je neparametarska alternativa t-testu nezavisnih uzoraka. Umesto da

poredi srednje vrednosti dveju populacija, kao što to radi t – test, Mann – Whitney U test poredi njihove medijane. Testira se nulta hipoteza

H_0 : Medijane obeležja dveju populacija su jednake,

protiv alternativne H_1 : Medijane obeležja dveju populacija nisu jednake.

Od svih elemenata uzorka potrebno je formirati varijacioni niz i dodeliti vrednost ranga svakom članu varijacionog niza, počevši od najmanje ka najvećim vrednostima. Ako dva ili više elemenata uzorka ima identične vrednosti, dodeljuje im se srednja vrednost ranga istih vrednosti. Nakon toga, računa se suma rangova svakog uzorka, gde oznake R_1 i R_2 označavaju sume rangova prvog i drugog uzorka, redom. Odgovarajući obimi uzoraka su označeni sa N_1 i N_2 . Neka je, po konvenciji, N_1 obim manjeg uzorka, ako obimi nisu jednaki, tako da $N_1 \leq N_2$. Značajna razlika između suma rangova podrazumeva značajnu razliku između uzoraka.

Test razlika između sume rangova, koristi sledeću statistiku

$$U = N_1 N_2 + \frac{N_1(N_1 + 1)}{2} - R_1$$

koja odgovara prvom uzorku. Uzoračka raspodela od U je simetrična i ima matematičko očekivanje i varijansu datu, redom, formulama

$$\mu_U = \frac{N_1 N_2}{2} \quad \sigma_U^2 = \frac{N_1 N_2 (N_1 + N_2 + 1)}{12}$$

Ako su N_1 i N_2 najmanje jednaki 8, ispostavlja se da U ima približno normalnu raspodelu, tako da $z = \frac{U - \mu_U}{\sigma_U}$ ima normalnu raspodelu sa matematičkim očekivanjem 0 i varijansom 1. Koristeći tablice standardizovane normalne raspodele, može se odlučiti da li se uzorci značajno razlikuju ili ne.

Približna veličina uticaja, u oznaci r , se može izračunati pomoću formule $r = \frac{z}{\sqrt{N}}$, gde je N ukupan obim uzorka ($N = N_1 + N_2$). Kriterijum za veličinu uticaja je:

- $r = 0.1 =$ mali uticaj,

- $r = 0.3$ = srednji uticaj,
- $r = 0.5$ = veliki uticaj.

3.9. T – test za nezavisne uzorke

Kada postoje dve nezavisne slučajne promenljive X_1 i X_2 sa normalnom raspodelom, onda promenljiva $\bar{X}_1 - \bar{X}_2$ ima normalnu raspodelu sa matematičkim očekivanjem $\mu_1 - \mu_2$ i varijansom datom sa: $\sigma^2 = \frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}$, gde su N_1 i N_2 redom obimi uzoraka sa uzoračkim srednjim vrednostima \bar{x}_1 i \bar{x}_2 . Testira se nulta hipoteza

$$H_0 : \mu_1 = \mu_2$$

protiv alternativne

$$H_1 : \mu_1 \neq \mu_2.$$

Razmotrimo sledeće dve situacije:

- Pretpostavimo da su varijanse σ_1 i σ_2 jednake, test statistika je:

$$t^* = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_p^2}{N_1} + \frac{s_p^2}{N_2}}}$$

gde je $s_p^2 = \frac{(N_1-1)s_1^2 + (N_2-1)s_2^2}{N_1 + N_2 - 2}$ (s_1^2 i s_2^2 su uzoračke varijanse). Test statistika ima Studentovu t- raspodelu sa $df = N_1 + N_2 - 2$ stepeni slobode.

- Pretpostavimo da su varijanse σ_1 i σ_2 različite, test statistika:

$$t^* = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

ima Studentovu t – raspodelu sa

$$df = \frac{(s_1^2/N_1 + s_2^2/N_2)^2}{(s_1^2/N_1)^2/N_1 + (s_2^2/N_2)^2/N_2}$$

stepeni slobode.

Prvo je potrebno proveriti da li su varijanse jednake ili ne, za to se koristi se Levene-ov test [12].

Veličina uticaja se najčešće procenjuje pomoću pokazatelja eta kvadrat. Eta kvadrat uzima vrednost između 0 i 1 i predstavlja proporciju varijanse u zavisnoj promenljivoj objašnjenu nezavisnom promenljivom. Formula za izračunavanje eta kvadrata je:

$$\text{Eta kvadrat} = \frac{t^2}{t^2 + (N_1 + N_2 - 2)}.$$

Ako je eta kvadrat jednako 0.01 = mali uticaj, 0.06 = umeren uticaj, a 0.14 = veliki uticaj [19].

3.10. T – test uparenih uzoraka

T – test uparenih uzoraka (ili ponovljenih merenja) upotrebljava se kada postoji jedna grupa ispitanika, a podaci od njih se prikupljaju u dva navrata ili pod dva različita uslova. Pored osnovnih pretpostavki za parametarske testove, dodatna pretpostavka uparenog t - testa je da razlika između dva rezultata dobijena za svakog ispitanika treba da je normalno raspoređena. Testira se nulta hipoteza

$$H_0: \text{Nema razlike između merenja,}$$

protiv alternativne $H_1: \text{Postoje razlike između merenja.}$

Razlike d_i se formiraju za svaki par opservacija. Ako je N broj parova opservacija, varijansa razlika se računa koristeći formulu: $s^2 = \sum_{i=1}^N \frac{(d_i - \bar{d})^2}{N-1}$. Neka su srednje vrednosti uzoraka populacija označene redom, \bar{x}_1 i \bar{x}_2 , tada test statistika

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s^2}{N}}}$$

ima Studentovu t-raspedelu sa $N - 1$ stepen slobode. Test može biti jednostrani ili dvostrani.

Veličina uticaja se procenjuje pomoću pokazatelja eta kvadrat. Formula za izračunavanje eta kvadrata kod t – testa uparenih uzoraka je:

$$\text{Eta kvadrat} = \frac{t^2}{t^2 + (N - 1)}.$$

Ako je eta kvadrat jednako 0.01 = mali uticaj, 0.06 = umeren uticaj, a 0.14 = veliki uticaj [19].

3.11. Koeficijent korelacije

Korelaciona analiza se bavi merenjem nivoa povezanosti između dve promenljive, x i y . Na početku, pretpostavlja se da su x i y numerički.

Pretpostavlja se da postoji par vrednosti (x, y) , merenih na svakom od N pojedinaca u našem uzorku. Može se označiti odgovarajući par tačaka koji odgovara vrednostima svakog pojedinca na dvodimenzionalnom tačkastom dijagramu. Po konvenciji, promenljiva x se stavlja na horizontalnoj osi, a promenljiva y na vertikalnoj osi dijagrama. Iscrtavanjem tačaka za svih N pojedinaca, dobija se raspršenje tačaka koje mogu predložiti odnos između dve promenljive.

Pirsonov koeficijent korelacije

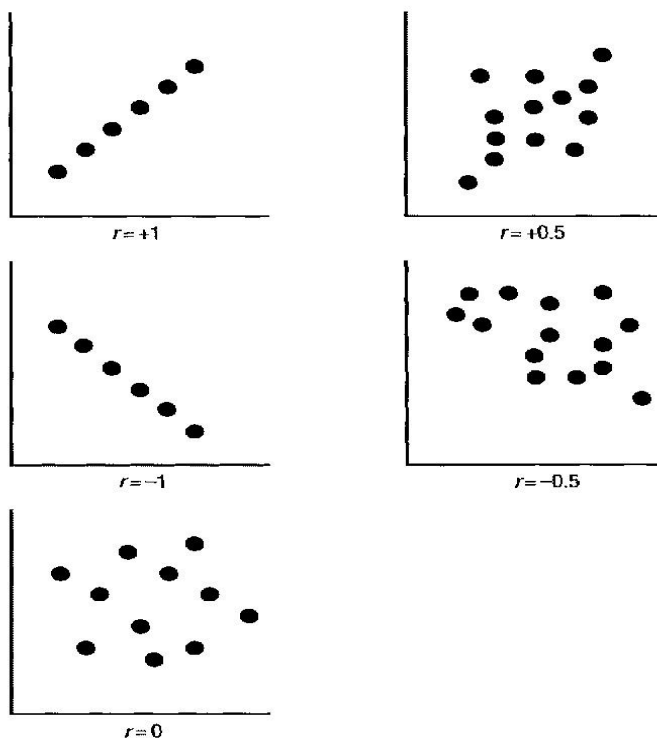
Koeficijent Pirsonove linearne korelacije prikladan je za intervalne (neprekidne) promenljive. Može se koristiti i kao mera povezanosti jedne neprekidne promenljive i jedne dihotomne promenljive. Dijagram rasturanja koji ilustruje povezanost promenljive x sa promenljivom y je prava linija. Koeficijent korelacije ocenjen iz uzorka, u oznaci r , dat je formulom

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}.$$

Svojstva koeficijenta korelacije:

1. $r \in [-1, 1]$
2. Znak od r (slika 3.1) pokazuje da li se jedna promenljiva povećava sa povećanjem druge promenljive (pozitivna korelacija) ili da li se neka promenljiva smanjuje sa povećanjem druge (negativna korelacija).
3. Veličina r pokazuje koliko su tačke blizu na pravoj. Naročito ako je $r = +1$ ili -1 , onda je savršena korelacija, jer sve tačke leže na pravoj. Ako je $r = 0$, onda ne postoji linearna korelacija. Što je r bliže granicama, veći je stepen linearne zavisnosti (slika 3.1).
4. r je bezdimenziona veličina, to nije jedinica mere.
5. Vrednost r važi samo u granicama vrednosti x i y u uzorku. Ne može se zaključiti da će imati istu vrednost kada su u pitanju vrednosti x ili y koje su ekstremnije nego vrednosti iz uzorka.
6. x i y se mogu menjati bez uticaja na vrednost r .
7. Korelacija između x i y ne znači uzrok i posledicu povezanosti.
8. r^2 predstavlja proporciju varijabilnosti od y koja se može pripisati linearnoj povezanosti sa x .

Slika 3.1. Dijagrama koji predstavljaju vrednosti koeficijenta korelacije u različitim situacijama (preuzeto iz [2])



Testiranje hipoteza za Pirsonov koeficijent korelacije sastoji se iz sledećih koraka:

1. Testira se nulta hipoteza $H_0: \rho = 0$ protiv alternativne $H_1: \rho \neq 0$, gde je ρ koeficijent korelacije populacije.
2. Prikupljaju se relevantni podaci od pojedinaca iz uzorka.
3. Izračunava se vrednost test statistike specifične za H_0
 - Ako je $n \leq 200$, r je test statistika.
 - Ako je $n > 200$, izračunava se $T = \sqrt{\frac{n-2}{1-r^2}}$ koje ima t -raspodelu sa $n - 2$ stepeni slobode.
4. Upoređuju se vrednosti test statistike sa tabličnim vrednostima [2].
 - Ako je $n \leq 150$, r se poredi sa tabličnom vrednošću Pirsonovog koeficijenta korelacije.
 - Ako je $n > 150$, T se poredi sa tabičnom vrednošću iz t - raspodele.

Kada je vrednost realnizovane test statistike veća od tablične vrednosti, onda se odbacuje nultu hipotezu u korist alternativne.

3.12. Binarna logistička regresija

Binarna logistička regresija služi za procenu koliko dobro skup nezavisnih promenljivih predviđa ili objašnjava kategorijsku zavisnu promenljivu. Dobija se pokazatelj adekvatnosti modela tj. ocena kvaliteta predviđanja rezultata. Binarnom logističkom regresijom dobija se i pokazatelj relevantne važnosti svake nezavisne promenljive ili interakcije između njih.

Logistička regresija ne počiva na pretpostavkama o raspodeli rezultata merenja nezavisnih promenljivih. Međutim, ona je osetljiva na visoke korelacije između nezavisnih promenljivih. Na rezultate logističke regresije jako utiču i netipične tačke.

U višestrukoj linearnoj regresiji, očekivana vrednost zavisne promenljive, y , modeluje se kao linearna funkcija objašnjavajućih promenljivih:

$$E(y) = b_0 + b_1x_1 + \dots + b_qx_q.$$

Binarna promenljiva uzima vrednosti 0 i 1, očekivana vrednost je verovatnoća, p , za koju zavisna promenljiva uzima vrednost jedan.

Može se modelirati p kao linearna funkcija objašnjavajuće promenljive i procene koeficijenta regresije metodom najmanjih kvadrata, ali postoje dva problema sa ovim direktnim pristupom:

- Ocenjena vrednost p , u oznaci \hat{p} , treba da zadovolji $0 \leq \hat{p} \leq 1$.
- Posmatrane vrednosti nisu u skladu sa normalnom raspodelom sa matematičkim očekivanjem p , nego kao Bernilijeva raspodela.

Prikladniji pristup modelu p je indirektno preko **logit transformacije** od p , tj. $\ln \frac{p}{1-p}$. Ovo dovodi do modela logističke regresije datog sa

$$\ln \frac{p}{1-p} = b_0 + b_1 x_1 + \dots + b_q x_q.$$

Drugim rečima, log-šansa za preživljavanje se modelira kao linearna funkcija objašnjavajućih promenljivih.

Parametri u logističkom regresionom modelu se mogu proceniti metodom maksimalne verodostojnosti. Parametri se obično objašnjavaju datim rezultatima u terminima šansi. U terminu p , logistički regresioni model može se napisati kao

$$p = \frac{\exp(b_0 + b_1 x_1 + \dots + b_q x_q)}{1 + \exp(b_0 + b_1 x_1 + \dots + b_q x_q)}.$$

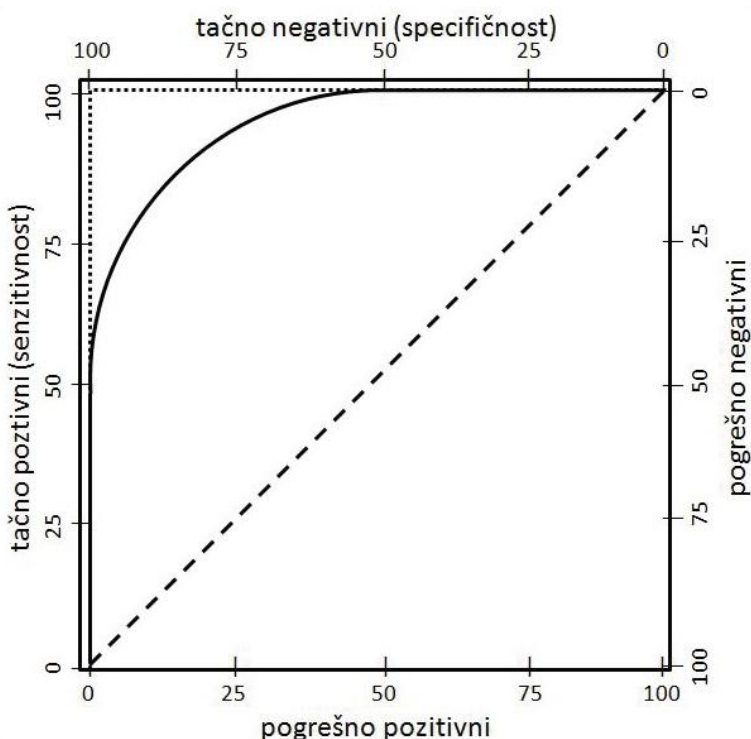
Ova funkcija linearnog prediktora je poznata kao logistička funkcija.

Konkurentni modeli u logističkoj regresiji mogu se formalno porediti sa testom verodostojnosti odnosa (engl. *likelihood ratio test*), testom skorova ili sa Wald-ovim testom. Ova tri testa su asimptotski ekvivalentni, ali se razlikuju u konačnom uzorku. Test verodostojnosti odnosa se generalno smatra najpouzdanijim, a Wald-ov test najmanje pouzdanim, iako će u mnogim primenama svi testovi dovesti do istog zaključka.

3.13. ROC krive

Cilj je da se odredi granična vrednost nekog testa, tako da se postigne najbolji odnos senzitivnosti i specifičnosti, odnosno da proporcija TP (tačno pozitivnih) i proporcija TN (tačno negativnih) bude najveća moguća. Ovo nije jednostavno, jer povećanje jedne mere povlači smanjenje druge. Kada se donosi ovakva odluka na osnovu izračunavanja senzitivnosti i specifičnosti za svaki mogući marker, dobija se dug niz paralelnih vrednosti i njihovim upoređivanjem može se ustanoviti gde je senzitivnost dovoljno visoka, a da je pri tome specifičnost najmanje oštećena. Statistička tehnika koja za cilj ima utvrđivanje granične vrednosti nekog testa koji daje najbolji odnos specifičnosti i senzitivnosti naziva se analiza **ROC krive** (engl. *Receiver Operating Characteristic Curve*).

Slika 3.2. ROC krive (preuzeto iz [10])



ROC kriva je grafički prikaz senzitivnosti i specifičnosti za svaki mogući marker za pojavu bolesti u koordinatnom sistemu gde su na ordinati (y) prikazane vrednosti senzitivnosti (proporcija tačnih pozitivnih), a na apscisi (x) vrednosti specifičnosti oduzete od 1 ($1 - \text{specifičnost}$), čime se dobijaju pogrešno pozitivni rezultati (slika 3.2). Na slici 3.2 se vidi u kakvom odnosu stoje proporcije četiri osnovne veličine iz ćelija 2×2 ma-

trice (TP – tačno pozitivni, PP – pogrešno pozitivni, TN – tačno negativni, PN – pogrešno negativni).

Tabela 3.4. Klasifikaciona matrica ROC krivih (preuzeto iz [12])

		Odluka	
		Tačno	Pogrešno
Realnost	Pozitivni	a	b
	Negativni	c	d

Iz klasifikacione matrice tabela 3.4, definišu se sledeći parametri:

- Odnos tačno pozitivnih $\equiv TP = \frac{a}{a+b}$. Poznat je kao senzitivnost, ovaj parametar govori koliko je osetljiva odluka našeg metoda u otkrivanju pogrešnog događaja. Klasifikacija metoda sa visokom osetljivošću će retko propustiti pogrešan događaj kada se dogodi.
- Odnos tačno negativnih $\equiv TN = \frac{d}{c+d}$. Poznat kao specifičnost, ovaj parametar govori koliko specifičan metod odluke je otkrio pogrešnih događaja. Klasifikacioni metod sa visokom specifičnošću će imati veoma nisku stopu netačnih.
- Odnos pogrešno pozitivnih $\equiv PP = \frac{c}{c+d} = 1 - \text{specifičnost}$.
- Odnos pogrešno negativnih $\equiv PN = \frac{b}{a+b} = 1 - \text{senzitivnost}$.

I senzitivnost i specifičnost se obično daju u procentima. Odluka metoda se smatra dobrom ako istovremeno ima visoku senzitivnost i visoku specifičnost. ROC kriva pokazuje senzitivnost u odnosu na dopunu specifičnosti (1 - specifičnost) za svaki mogući prag odluke.

Ako je pojava koja se meri takva da se oni koji su bolesni i oni koji su zdravi uopšte ne razlikuju, onda test ima šanse 50:50% da slučajno „pogodi“ ko je bolestan i koji god marker na testu da se uzme, proporcija ispravno klasifikovanih ostaje ista. Ova situacija bi na ROC grafiku bila predstavljena dijagonalnom linijom koja spaja donji levi i gornji desni ugao, odnosno dve nulte tačke sa slike 3.2. Ova dijagonala se obično zove dijagonala slučajnog ishoda i prikazana je isprekidanom linijom na slici 3.2.

Sa druge strane, može biti situacija u kojoj bi svi bolesni pacijenti bili jasno različiti od zdravih, a test dobar indikator bolesti. Ako se krene od najviših graničnih vrednosti na testu, senzitivnost (sposobnost da se prepoznaju bolesni pacijenti), će biti niska, ali će i pogrešno pozitivni biti odsutni i većina rezultata će biti pogrešno negativna. Na ROC krivoj taj rezultat će se nalaziti na samoj y-osi (jer su pogrešno pozitivni 0) i to na nižim vrednostima za senzitivnost. Svaka sledeća niža granična vrednost na ROC grafiku davaće presek senzitivnosti i (-) specifičnosti na samoj y-osi, senzitivnost će se sa spuštanjem kriterijuma povećavati (test će sve više bolesnih ljudi prepoznati kao takve), pa tako sve do savršenog kriterijuma, kada kriva stiže do gornjeg levog ugla gde se susreću tačke maksimalne senzitivnosti i specifičnosti. Sa daljim spuštanjem granične vrednosti počinju da rastu pogrešno pozitivni, a opadaju tačno negativni rezultati, odnosno specifičnost se smanjuje i ROC linija prati gornju liniju grafika. Dakle, idealni test kojim se mere jasno razdvojeni zdravi i bolesni dao bi ROC krivu koja prati levu i gornju ivicu grafika, kao što je prikazano tačkastom linijom na slici 3.2.

ROC krive, u najvećem broju slučajeva, se nalaze negde između dve opisane situacije, između dijagonale slučajnog ishoda i krive koja se poklapa sa osama senzitivnosti i specifičnosti. Što je dobijena krivulja bliže idealnoj, tačkastoj sa slike 3.2, to je test bolje razlikuje dve grupe, jer se i senzitivnost i specifičnost za svaku graničnu vrednost približavaju idealnim vrednostima i veće su šanse da se među njima pronađe ona granična vrednost koji će imati mali broj i pogrešno pozitivnih i pogrešno negativnih. Suprotno, što je krivulja bliže dijagonali slučajnog ishoda, test lošije razlikuje dve grupe i njegove klasifikatorne mogućnosti se sve manje razlikuju od slučajnog pogađanja.

Iz navedenog sledi da, što je površina koju zahvata ROC kriva veća, to test bolje razlikuje grupe. *Površine ispod ROC krive* (engl. Area Under the Curve: AUC) je pokazatelj da će verovatnoća slučajno odabrane bolesne osobe na testu za pojavu bolesti imati veću vrednost, nego zdrava osoba. Kao takva, AUC je ekvivalentna Mann-Whitney U testu. Testira se hipoteza $H_0: AUC = 0.5$. Ako je vrednost test AUC veća od 0.7, i ako se značajno razlikuje od površine ispod dijagonale, onda testirana promenljiva može biti dobar marker za pojavu bolesti. Kada se pronađu dobri markeri za pojavu bolesti, onda se za njih traži granična vrednost. Granična vrednost je vrednost promenljive koja predstavlja dobar marker, za koju je proizvod senzitivnosti i specifičnosti maksimalan.

3.14. Wilcoxon test ranga

Wilcoxon test ranga se naziva i Wilcoxon-ov test ekvivalentnih parova ili test rangova sa znakom. Namenjen je ponovljenom merenju tj. za utvrđivanje da li je razlika statistički značajna između dva zavisna uzorka. To je neparametarska alternative t – testu ponovljenih merenja, ali umesto poređenja srednjih vrednosti, Wilcoxon-ov test pretvara rezultate u rangove i njih poredi u trenutku 1 i trenutku 2.

Da bi se primenio Wilcoxon-ov test promenljive moraju biti merene bar na intervalnoj mernoj skali. Testira se nulta hipoteza

$$H_0: \text{Nema razlike između merenja,}$$

protiv alternativne $H_1: \text{Postoje razlike između merenja.}$

Pre svega računaju se razlike između rezultata koji su u paru, u oznaci d_i , $i = 1, 2, \dots, n$, a zatim se rangiraju razlike (pri tom ne uzimajući u obzir predznak razlike). Pri rangiranju najmanjoj razlici dodeljuje se rang 1, rangovi se označavaju sa r_i , $i = 1, 2, \dots, n$. Razlike jednake nuli se ispuštaju. Svakom rangui dodeljuje se predznak koji je imala njemu odgovarajuća razlika. Prave se odvojeni zbrovi pozitivnih i negativnih rangova.

Za test statistiku uzima se po modulu manji zbir rangova, u oznaci R , koji ima približno normalnu raspodelu sa matematičkim očekivanjem jednakim $\mu = \frac{\sum_{i=1}^N r_i}{2}$ i standardnom devijacijom $\sigma = \sqrt{\frac{\sum_{i=1}^N r_i^2}{4}}$ [15]. Statistika $z = \frac{R - \mu}{\sigma}$ ima normalnu raspodelu sa matematičkim očekivanjem 0 i varijansom 1. Koristeći tablice standardizovane normalne raspodele, može se odlučiti da li se razlike između merenja statistički značajno razlikuju ili ne.

Veličina uticaja u ovom testu se može izračunati na način opisan za Mann-Whitney U test, tj. deljenjem veličine z sa kvadratnim korenom iz n . Međutim, u ovoj situaciji n je broj opservacija u dva vremenska intervala, a ne broj ispitanika. Jačina uticaja se meri kao kod Mann-Whitney U testa.

4. Psihijatrijske metode

Ispitivanje je sprovedeno na Klinici za psihijatriju u Kragujevcu, uz poštovanje principa Dobre kliničke prakse [16].

4.1. Uzorak

U ispitivanju je učestvovalo 171 ispitanika oba pola, podeljena u tri grupe. Svi ispitanici obuhvaćeni ispitivanjem su imali 18 i više godina. Prvu grupu ispitanika čine pacijenti kojima je postavljena dijagnoza akutnog psihotičnog poremećaja i koji do sada nisu lečeni antipsihoticima. Drugu grupu čine pacijenti kojima je dijagnoza psihotičnog procesa (shizofrenije) postavljena pre više od godinu dana i od tada su lečeni antipsihoticima, a aktuelno su u fazi pogoršanja. Treću, kontrolnu grupu, čine zdravi dobrovoljci.

4.2. Instrumenti kliničko-psihijatrijske procene

Psihijatrijske skale procene i imunski testovi su sprovedeni na početku i nakon 30 dana tretmana na pacijentima prve i druge grupe, a u međuvremenu su na sedmičnom nivou podvrgnuti kliničko-psihijatrijskoj eksploataciji. Za psihotične pacijente predviđena je primena specijalizovanog testa za ispitivanje pojedinih psihičkih funkcija i određenih kliničkih sindroma:

- Skala pozitivnih i negativnih simptoma – PANSS (engl. *Positive and Negative Syndrom Scale*) [7], kako bi se utvrdila težina mentalnog poremećaja, pratio tok i ishod (poboljšanje/pogoršanje) nakon 30 dana lečenja.

Pored toga pacijentima prve i druge grupe rađena je analiza kompletne krvne slike na početku i 30 dana nakon tretmana. Od svih pacijenata uzimani su podaci o polu i starosti.

Kratak pregled obeležja:

Pol je kategorijalna dihotomna (binarna) promenljiva, gde je ženski pol označen sa 0, a muški sa 1.

Starost je kvantitativna promenljiva koja uzima vrednosti veće ili jednake 18.

P1 i *P2* su kvantitativne promenljive koje predstavljaju pozitivne skale merene PANSS testom. *P1* predstavlja merenje na početku terapije, a *P2* merenje nakon 30 dana od početka terapije. Pozitivne skale čini 7 stavki od kojih se svaka meri Likertovom skalom od 1 do 7, gde 1 označava da je odsutan poremećaj, a 7 označava da je ekstreman poremećaj. Rezultati svih stavki se sabijaru i čine obeležje *P1* odnosno *P2*. S tim u vezi *P1* i *P2* uzimaju vrednosti od 7 do 49.

N1 i *N2* imaju iste karakteristike kao *P1* i *P2*. Jedina razlika je što one predstavljaju negativne skale merene PANSS testom.

Opšta psihopatološka skala PANSS testa označena je sa *G1* i *G2*, gde *G1* takođe označava merenje na početku tretmana, a *G2* merenje nakon 30 dana. Ta skala se sastoji od 16 stavki, a svaka stavka je merena Likertovom skalom. *G1* i *G2* predstavljaju zbir merenja svih stavki i uzimaju vrednosti od 16 do 112.

Kako se skala pozitivnih i nedativnih simptoma sastoji od pozitivnih, negativnih i opštih psihopatoloških stavki, onda *PANSS1* i *PANSS2* predstavljaju zbir merenja svih stavki na početku tretmana i 30 dana kasnije. Vrednosti koje uzimaju *PANSS1* i *PANSS2* su od 30 do 210.

Obeležja od *C1* do *C5* predstavljaju citokine. To su kvantitativne neprekidne promenljive. Citokini su polipeptidi ili glikopeptidi, koji prenose informacije među ćelijama i važni su medijatori upalnih bolesti [5].

WBC je skraćenica za leukocite (bela krvna zrnca) koja služe za odbranu organizma od bakterija, gljivica i virusa. Normalne vrednosti leukocita u krvi su od $6 - 10 \times 10^9$ po litru krvi [4].

RBC predstavlja eritrocite (crvena krvna zrnca) koja prenose kiseonik iz pluća u ostale delove tela. Normalne vrednosti eritrocita u krvi su od $3.5 - 5.2 \times 10^{12}$ po litru krvi [4].

HGB je skraćenica za hemoglobin. To je složen protein sa molekulskom težinom 64 458 Da. To je jedini protein krvi koji ne pripada proteinima plazme, nego se nalazi unutar eritrocita. Hemoglobina u krvi ima 120 – 160g/l [4].

HCT predstavlja hematokrit. Hematokrit je odnos uobličjenih elemenata prema celokupnoj krvi ili odnos uobličjenih elemenata i tečne faze, tj. odnos eritrocita i tečne faze. Iznosi od 40 – 45 % [4].

Trombociti, u oznaci *PLT*, su krvne pločice. Ima ih $150 - 300 \times 10^9$ po litru krvi [4].

PCT je prokalcitonin. To je zapreminski udeo trombocita u jedinici pune krvi.

LYM predstavlja limfocite. Oni poseduju jedno nesegmentisano jedro koje zauzima najveći deo ćelije, tako da obično citoplazmu čini tanak polumesečni sloj oko jedra. Normalne vrednosti limfocita su $1.5 - 4 \times 10^9$ po litru krvi [4].

AcidUric je mokraćna kiselina. Normalne vrednosti za muškarce su od 200 – 420 *mmol/l*, a za žene su od 140 – 340 *mmol/l* [8].

GLU je glukoza (šećer). Normalne vrednosti su od 3.5 – 6.5 *mmol/l* [21].

UREA objašnjava stanje bubrega. Normalne vrednosti pripadaju intervalu od 1.7 – 8.3 [8].

KREATININ takođe objašnjava stanje bubrega. Normalne vrednosti za muškarce su od 62 – 106 *mmol/l*, a za žene su od 44 – 80 *mmol/l*. Povišene vrednosti javljaju se kod bubrežnih oboljenja. Povećana urea, a normalan kreatinin može da bude i rezultat ishrane bogate proteinima [8].

HOLESTEROL se takođe nalazi u krvnoj plazmi, a predstavlja masnoću u krvi. Normalne vrednosti su od 3.6 – 6.5 *mmol/l* [21].

TRIGLICERIDI predstavljaju masnoće u krvi. Preporučeni su manji od 1.7, dok su visoki veći od 2.3 *mmol/l* [8].

HDL predstavlja holesterol (dobar). Preporučene vrednosti su veće od 1.54, a niske vrednosti su manje od 1.0 *mmol/l* [8].

LDL je holesterol (loš). Preporučene vrednosti su manje od 2.6, a visoke su veće od 4.1 *mmol/l* [8].

BILIRUBIN objašnjava stanje jetre. Normalna vrednost je od 1.1 – 18.8 *mmol/l* [8].

PROTEINI u krvnoj plazmi uzimaju normalne vrednosti od 60 do 80 *g/l* [21].

Fe je gvožđe. Ono objašnjava koštanu srž. Za muškarce normalne vrednosti su od 10.6 – 28.3 *mmol/l*, a za žene su od 6.6 – 26.0 *mmol/l* [8].

AST je oznaka za aspartat aminotransperaze, iznosi od 7 – 38 *LJ/l* [8].

ALT je skraćenica za alanin aminotransperaze. Normalne vrednosti su od 8 – 41 *LJ/l* [8].

ALP je alkalna fosfataza. Normalne vrednosti za odrasle su od 0 – 270 *LJ/l* [8].

GamaGT predstavlja glutamil. Normalna vrednost za muškarce je od 8 – 40, a za žene 5 – 35 *LJ/l*. [8].

CK je keratin kineza.

CRP je skraćenica koja označava C reaktivni protein. Reč je o proteinu koji se proizvodi u jetri. Njegova koncentracija u venskoj krvi se koristi kao značajan biohemijski parametar za mnoga stanja. Normalna koncentracija *CRP*-a se smatra do 10*mg/l* [9].

K je količina kalijuma u krvnoj plazmi. Normalna vrednost kalijuma je od 3.5 – 5 *mmol/l* [21].

Na predstavlja količinu natrijuma u krvnoj plazmi. Normalna vrednost natrijuma je od 136 – 146 *mmol/l* [21].

SE predstavlja brzinu sedimentacije eritrocita po Vestergernu, iznosi za muškarce od 0 – 15 *mm* za 1 sat, a za žene od 0 – 20 *mm* za 1 sat [21].

DIJAGNOZA je kategorijalna varijabla sa tri kategorije. Predstavlja tri grupe ispitanika: $-1 = „zdravi”$, $0 = „akutni”$ i $1 = „shizofrenija (hronični)”$.

Terapija je kategorijalna varijabla sa tri kategorije. Predstavlja tri različita leka.

Pojava_bolesti je dihotomna promenljiva, gde su $0 = „zdravi”$ i $1 = „bolesni”$. Bolesne čine akutni i shizofrenični.

5. Analiza podataka

Primenom deskriptivne statistike i frekvencija, pronađeno je da promenljiva N1 ima 1 pogrešnu vrednost. Kako promenljiva N1 uzima vrednosti u segmentu [7, 49], pogrešna vrednost je 59. Promenljiva G1 uzima vrednosti u segmentu [16, 112], pa u datim podacima ima jednu pogrešnu vrednost, a to je 120. Promenljiva P2 može uzimati vrednosti iz segmenta [7, 49]. U datim podacima u promenljivoj P2 se pojavljuje vrednost 152, to implicira da je 152 pogrešna. Sve pogrešne vrednosti je potrebno izbrisati, a zatim tretirati kao nedostajuće podatke.

Problem nedostajućih podataka, koji se javio u istraživanju, rešavan je u dva navrata. Prvo su registrovani i eliminisani svi nedostajući podaci u obeležjima pol, starost, dijagnoza i terapija. Oni su eliminisani metodom *pairwise*. Zatim su preostali nedostajući podaci, nadomešteni postupkom zasnovanim na maksimizaciji očekivanih vrednosti (*Expectation – Maximization - EM*).

6. Rezultati

Kako je u radu ispitivano dosta parametara i primenjen veliki broj statističkih metoda, rezultati će biti prikazani u nekoliko odeljaka. Demografski podaci prikazani su u tabeli 6.1.

Tabela 6.1. Demografski podaci

	Godine	34.82 ± 9.73
Pol	Žene	79 (46.2%)
	Muškarci	89 (52.0%)
	Nedostajući podaci	3 (1.8%)
Dijagnoza	Zdravi	35 (20.5%)
	Akutni	89 (52.0%)
	Shizofrenija	47 (27.5%)
	Nedostajući podaci	0 (0%)

6.1. Citokini i dijagnoza

Kako bi se izabrala odgovarajuća statistička metoda za ispitivanje statističke značajnosti razlika srednjih vrednosti citokina (prvo merenje) između dijagnoza, potrebno je proveriti da li citokini imaju normalnu raspodelu u svakoj grupi dijagnoza.

Tabela 6.2. Kolmogorov – Smirnov test normalnosti

DIJAGNOZA	Kolmogorov-Smirnov ^a			
	Statistika	df	p	
Citokin 1	Zdravi	0.361	35	< 0.0005
	Akutni	0.392	89	< 0.0005
	Shizofrenija	0.310	47	< 0.0005
Citokin 2	Zdravi	0.454	35	< 0.0005
	Akutni	0.336	89	< 0.0005
	Shizofrenija	0.401	47	< 0.0005
Citokin 3	Zdravi	0.390	35	< 0.0005
	Akutni	0.345	89	< 0.0005
	Shizofrenija	0.366	47	< 0.0005
Citokin 4	Zdravi	0.186	35	0.003
	Akutni	0.250	89	< 0.0005
	Shizofrenija	0.394	47	< 0.0005
Citokin 5	Zdravi	0.290	35	< 0.0005
	Akutni	0.165	89	< 0.0005
	Shizofrenija	0.156	47	0.006

a. Lilliefors korekcija

Kako je u svakoj grupi obim uzorka veći od 30, koristi se Kolmogorov-Smirnov test normalnosti. Iz tabele 6.2 se vidi da je *p* vrednost Kolmogorov-Smirnovog testa za svaku grupu svakog citokina manje od 0.05, što implicira da ni jedna grupa ni jednog ci-

tokina nema normalnu raspodelu. U tabeli 6.3 prikazane su vrednosti citokina prema dijagnozi.

Tabela 6.3. Vrednosti citokina prema dijagnozi

	Zdravi Medijana (25. perc – 75. perc)	Akutni Medijana (25. perc – 75. perc)	Shizofrenija Medijana (25. perc – 75. perc)	p
Citokin 1	0.00 (0.00 – 20.06)	0.00 (0.00 – 14.49)	1.27 (0.00 – 97.39)	0.198
Citokin 2	0.00 (0.00 – 0.00)	0.00 (0.00 – 24.62)	0.00 (0.00 – 4.28)	0.177
Citokin 3	0.00 (0.00 – 14.38)	0.00 (0.00 – 15.39)	4.54 (0.00 – 14.37)	0.636
Citokin 4	1940.32 (474.57 – 3895.24)	895.53 (0.00 – 4372.83)	849.65 (45.75 – 2881.95)	0.535
Citokin 5	17.24 (12.86 – 21.88)	35.95 (18.39 – 73.97)	55.95 (18.72 – 105.18)	< 0.0005

S obzirom da citokini nemaju normalnu raspodelu u svakoj grupi, za testiranje razlika srednjih vrednosti citokina (prvo merenje) između dijagnoza korišćen je Kruskal-Wallis-ov test.

Kruskal-Wallis-ov test je otkrio da razlika srednjih vrednosti citokina 1 između zdravih, akutnih i shizofreničnih nije statistički značajna (Grupa zdravi $N = 35$, Grupa akutni $N = 89$, Grupa shizofrenija $N = 47$), $\chi^2(2, N = 171) = 3.24$, $p = 0.198$.

Isti test pokazao je da razlika srednjih vrednosti citokina 2 između zdravih, akutnih i shizofreničnih nije statistički značajna (Grupa zdravi $N = 35$, Grupa akutni $N = 89$, Grupa shizofrenija $N = 47$), $\chi^2(2, N = 171) = 3.46$, $p = 0.177$.

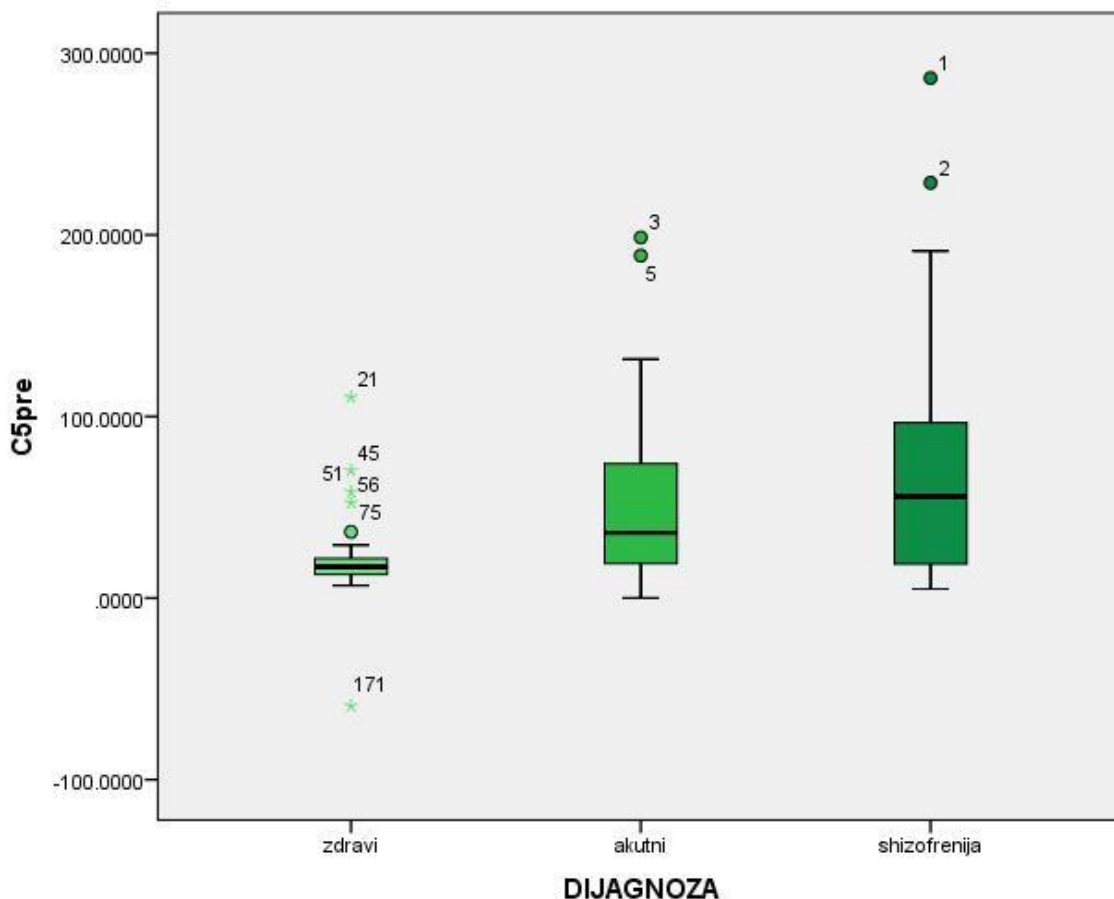
Kruskal-Wallis-ov test je potvrdio da razlika srednjih vrednosti citokina 3 između zdravih, akutnih i shizofreničnih nije statistički značajna (Grupa zdravi $N = 35$, Grupa akutni $N = 89$, Grupa shizofrenija $N = 47$), $\chi^2(2, N = 171) = 0.906$, $p = 0.636$.

Takođe, Kruskal-Wallis-ov test je otkrio da razlika srednjih vrednosti citokina 4 između dijagnoza nije statistički značajna (Grupa zdravi $N = 35$, Grupa akutni $N = 89$, Grupa shizofrenija $N = 47$), $\chi^2(2, N = 171) = 1.072$, $p = 0.585$.

Kruskal-Wallis-ov test je otkrio da je razlika srednjih vrednosti citokina 5 između zdravih 17.24 (12.86 – 21.88), akutnih 35.95 (18.39 – 73.97) i shizofreničnih 55.95 (18.72 – 105.18) statistički značajna (Grupa zdravi $N = 35$, Grupa akutni $N = 89$, Grupa shizofrenija $N = 47$), $\chi^2(2, N = 171) = 23.241$, $p < 0.0005$. Na osnovu medijane

citokina 5 (prvo merenje) i dijagnoze nacrtan je dijagram (grafik 6.1). Za crtanje dijagrama korišćena je medijana jer se Kruskal-Wallis-ovim testom testira jednakost medijana.

Grafik 6.1. Citokin 5 (prvo merenje) i dijagnoza



Statistički značajnu razliku srednjih vrednosti citokina između dijagnoza je moguće otkriti koristeći MANOVA analizu. Kako citokini nemaju normalnu raspodelu, neophodno ih je normalizovati. Za normalizaciju citokina korišćena je *Tukey* transformacija. Nakon normalizacije urađena je deskriptivna statistika noramalizovanih vrednosti citokina (prvo merenje) tj. izračunati su koeficijenti asimetrije i spljoštenosti (tabela 6.4).

Tabela 6.4. Deskriptivna statistika (koeficijenti asimetrije i spljoštenosti)

	N	Koeficijent asimetrije		Koeficijent spljoštenosti	
		Statistika	Std. greška	Statistika	Std. greška
NC1pre	171	0.895	0.186	-0.250	0.369
NC2pre	171	1.345	0.186	0.611	0.369
NC3pre	171	0.133	0.186	0.244	0.369
NC4pre	171	0.395	0.186	-0.576	0.369
NC5pre	171	0.002	0.186	-0.154	0.369
Ukupno N	171				

Iz tabele 6.4 se vidi da koeficijenti asimetrije i spljoštenosti za normalizovane vrednosti citokina pripadaju segmentu $[-2, 2]$, što pokazuje da normalizovane vrednosti citokina imaju normalnu raspodelu. Pored pretpostavke o univarijacionoj normalnosti citokina, neophodno je proveriti i ostale pretpostavke za MANOVA analizu. Dalje, sve pretpostavke MANOVA analize se proveravaju na normalizovanim vrednostima citokina i MANOVA analiza se sprovodi korišćenjem normalizovanih vrednosti citokina.

U grupama ima više ispitanika nego što je neophodan broj (tabela 6.2).

Mardijinim testom je potvrđeno da citokini imaju multivarijacionu normalnu raspodelu ($N(b2p) = 2.7996 < 3, p = 0.051$).

Tabela 6.5. Deskriptivna statistika normalizovanih vrednosti citokina

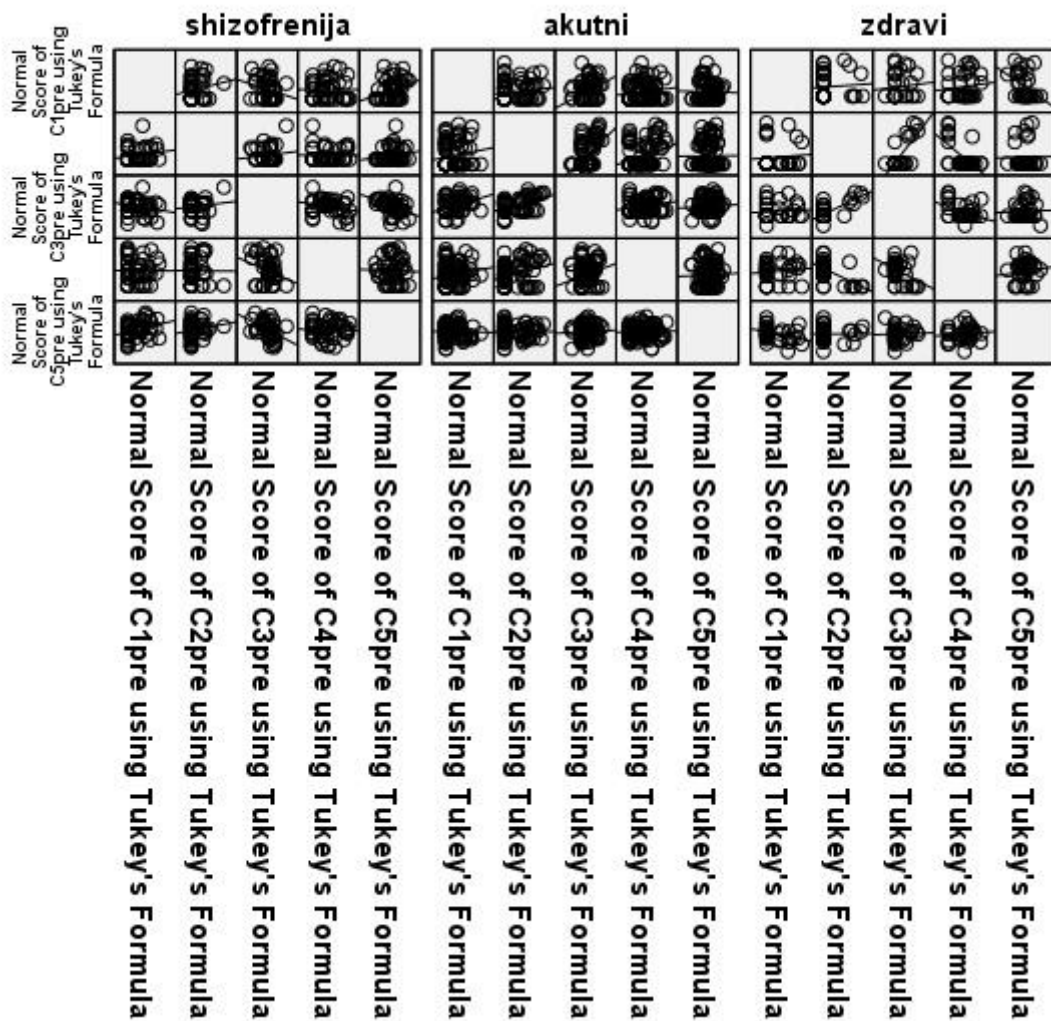
		Statistika	Std. Greška
NC1pre	Srednja vrednosti	0.064363	0.0634027
	Srednja vrednost bez 5% slučajeva	-0.007267	
	Minimum	-0.6057	
	Maksimum	2.6614	
NC2pre	Srednja vrednosti	0.075694	0.0579988
	Srednja vrednost bez 5% slučajeva	-0.006655	
	Minimum	-0.3901	
	Maksimum	2.6614	
NC3pre	Srednja vrednosti	0.021472	0.0717672
	Srednja vrednost bez 5% slučajeva	0.023858	
	Minimum	-2.6614	
	Maksimum	2.6614	
NC4pre	Srednja vrednosti	0.030791	0.0703368
	Srednja vrednost bez 5% slučajeva	-0.016989	
	Minimum	-1.1021	
	Maksimum	2.6614	
NC5pre	Srednja vrednosti	0.000112	0.0757633
	Srednja vrednost bez 5% slučajeva	0.000006	
	Minimum	-2.6614	
	Maksimum	2.6614	

Iz tabele 6.5 se vidi da srednje vrednosti citokina i srednje vrednosti bez 5% maksimalnih i minimalnih vrednosti citokina se mnogo ne razlikuju, te nema potrebe za brigu o univarijacionim netipičnim vrednostima.

Kako je $\chi^2_{6,0.001} = 22.46$ ($n = 6, p = 0.001$), gde je n broj promenljivih u analizi, a najviša vrednost Mahalanobisovih distanci u podacima je 27.18, odatle sledi da postoji jedna multivarijantna netipična tačka za koju je Mahalanobisova distanca jednaka 27.18. Ostale Mahalanobisove distance su manje od χ^2 vrednosti, te nema više multivari-

jacionih netipičnih tačaka. Pronađena multivarijaciona netipična tačka je izbačena iz dalje analize.

Grafik 6.2. Matrični tačkasti dijagram za dijagnozu i normalizovane vrednosti citokina



Na grafiku 6.2 se ne primećuju znakovi očigledne nelinearnosti, te se pretpostavka o linearnosti smatra zadovoljena.

Kako se ne radi postepena analiza, nije potrebno proveravati pretpostavku o homogenosti regresije.

Kako u tabeli 6.6 nema jako visokih korelacija između citokina, nema potrebe za brigom o multikolinearnosti.

Tabela 6.6. *Pirsonovi koeficijenti korelacije noramalizovanih vrednosti citokina*

	NC1pre	NC2pre	NC3pre	NC4pre	NC5pre
NC1pre	1	0.134	0.142	0.089	0.077
NC2pre	0.134	1	0.440**	0.104	0.076
NC3pre	0.142	0.440**	1	-0.009	-0.097
NC4pre	0.089	0.104	-0.009	1	-0.003
NC5pre	0.077	0.076	-0.097	-0.003	1

** . Korelacija je značajna na nivou 0.01 (obostrano).

Homogenost matrica varijanse i kovarijanse je proverena Boksovim testom za jednakost matrica kovarijanse. Kako je p vrednost Boksovog testa < 0.0005 , pretpostavka o homogenosti matrica varijanse je narušena [22].

MANOVA analizom utvrđena je statistički značajna razlika srednjih vrednosti citokina između dijagnoza, $F(10, 328) = 3.342$, $p < 0.0005$; Pillai's Trace lambda je 0.185; parcijalni eta kvadrat = 0.092. Kada su rezultati zavisnih promenljivih razmotreni zasebno, jedina razlika srednjih vrednosti koja je dosegla statističku značajnost (uz po Bonferoniju prilagođen nivo alfa od 0.124) bio je citokin 5, $F(2, 167) = 12.955$, $p < 0.0005$, parcijalni eta kvadrat = 0.134. Pregledom srednjih vrednosti zabeleženo da je kod zdravih neznatno manje citokina 5 (-0.67 ± 0.79) nego kod akutnih (0.09 ± 0.86) i shizofreničnih (0.36 ± 1.14).

6.2. Pol i dijagnoza

Povezanost dijagnoze sa polom ispitana je koristeći χ^2 test nezavisnosti $k \times 2$ tabele (tabela 6.7).

Tabela 6.7. *Povezanost dijagnoze sa polom*

		Pol		Ukupno
		Ženski	Muški	
DIJAGNOZA	zdravi	0.00%	20.80%	20.80%
	akutni	31.00%	22.00%	53.00%
	bolesni	16.10%	10.10%	26.20%
Ukupno		47.00%	53.00%	100.00%

Kako više od 80% ćelija zadovoljava pretpostavku o najmanjoj ćelijskoj frekvenciji, znači da nije prekršena pretpostavka (minimalna očekivana vrednost frekvencija je

16.46). χ^2 test nezavisnosti (Pirsonov χ^2) pokazao je značajnu vezu između dijagnoze i pola, $\chi^2(2, 168) = 39.345$, $p < 0.0005$, *Cramer's V* = 0.484.

Iz tabele 6.7 se vidi da grupu zdravih čine samo muškarci, njih 20.80% od celog uzorka. Akutni čine 53% celog uzorka, od toga 31% otpada na žene, a 22% na muškarce. Bolesni predstavljaju 26.20% od ukupnog uzorka, od čega su 16.10% žene, a 10.10% muškarci. Od ukupnog uzorka 47% čine žene, a 53% muškarci.

6.3. Starost i dijagnoza

Kako bi se odlučilo koji je statistički test potreban za ispitivanje statistički značajne razlike srednjih vrednosti starosti između dijagnoza, neophodno je prvo proveriti da li starost ima normalnu raspodelu u svakoj grupi dijagnoze. Kolmogorov-Smirnov test tvrdi da starost ima normalnu raspodelu u sve tri grupe (tabela 6.8).

Tabela 6.8. Kolmogorov-Smirnov test normalnosti

		Kolmogorov-Smirnov ^a		
	DIJAGNOZA	Statistika	df	p
starost	Zdravi	0.114	35	0.200*
	Akutni	0.087	89	0.089
	Shizofreni	0.086	47	0.200*

a. Lilliefors korekcija

*. Ovo je donja granica praga značajnosti.

Kako je pretpostavka o normalnosti ispunjena, za testiranje statističke značajnosti uticaja starosti na dijagnozu korišćena je ANOVA analiza. Jednofaktorskom dispersionom analizom nije utvrđena statistički značajna razlika na nivou $p < 0.05$ starosti između dijagnoza: $F(2, 168) = 1.205$, $p = 0.302$.

6.4. Citokini i pojava bolesti

Kako bi se izabrala odgovarajuća statistička metoda za ispitivanje statističke značajnosti razlika srednjih vrednosti citokina (prvo merenje) o odnosu na pojavu bolesti, potrebno je proveriti da li citokini imaju normalnu raspodelu u obe grupe.

Pojava bolesti predstavlja dihotomnu promenljivu koja je nastala rekodiranjem kategorijalne promenljive dijagnoza. Kategorija zdravi je zadržana i uzima vrednost 0, a kategorije akutni i shizofrenija su pretvoreni u jednu kategoriju, u nazivu bolesni, koja uzima vrednost 1.

Tabela 6.9. Kolmogorov-Smirnov test normalnosti

	pojava bolesti	Kolmogorov-Smirnov ^a		
		Statistika	df	p
Citokin 1	Zdravi	0.361	35	< 0.0005
	Bolesni	0.368	136	< 0.0005
Citokin 2	Zdravi	0.454	35	< 0.0005
	Bolesni	0.356	136	< 0.0005
Citokin 3	Zdravi	0.390	35	< 0.0005
	Bolesni	0.343	136	< 0.0005
Citokin 4	Zdravi	0.186	35	0.003
	Bolesni	0.367	136	< 0.0005
Citokin 5	Zdravi	0.290	35	< 0.0005
	Bolesni	0.160	136	< 0.0005

a. Lilliefors korekcija

Kako je u svakoj grupi obim uzorka veći od 30, koristi se Kolmogorov-Smirnov test normalnosti. Iz tabele 6.9 se vidi da je *p* vrednost Kolmogorov-Smirnovog testa u svakoj grupi svakog citokina manje od 0.05, što implicira da ni jedna populacija ni jednog citokina nema normalnu raspodelu.

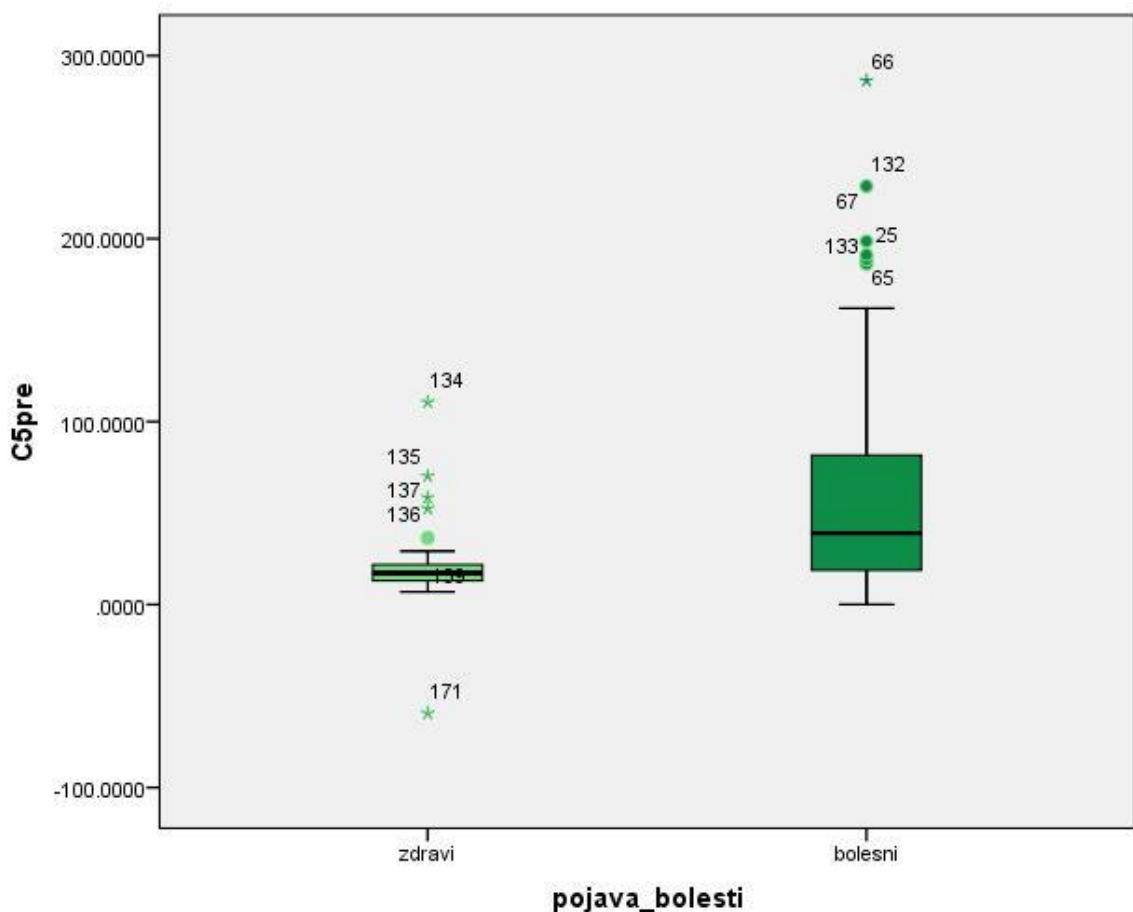
Kako citokini nemaju normalnu raspodelu u svakoj grupi, za testiranje razlika srednjih vrednosti citokina u odnosu na pojavu bolesti korišćen je Mann-Whitney U test.

Tabela 6.10. Vrednosti citokina prema pojavi bolesti

	Zdravi	Akutni	p
	Medijana	Medijana	
	(25. perc – 75. perc)	(25. perc – 75. perc)	
Citokin 1	0.00 (0.00 – 20.06)	0.00 (0.00 – 27.91)	0.198
Citokin 2	0.00 (0.00 – 0.00)	0.00 (0.00 – 11.11)	0.177
Citokin 3	0.00 (0.00 – 14.38)	0.90 (0.00 – 14.98)	0.636
Citokin 4	1940.32 (474.57 – 3895.24)	856.85 (0.00 – 4173.14)	0.535
Citokin 5	17.24 (12.86 – 21.89)	39.06 (18.75 – 81.65)	< 0.0005

Mann-Whitney U test nije otkrio statistički značajnu razliku srednjih vrednosti citokina 1, citokina 2, citokina 3 i citokina 4 između zdravih i bolesnih (tabela 6.10). Isti test tvrdi da je razlika srednjih vrednosti citokina 5 između zdravih 17.24 (12.86 – 21.89) i bolesnih 39.06 (18.75 – 81.65) statistički značajna (Grupa zdravi $N = 35$, Grupa bolesni $N = 136$), $U = 1161.5$, $z = -4.665$, $p < 0.0005$, $r = 0.36$. Na osnovu medijane citokina 5 (prvo merenje) i pojave bolesti nacrtan je dijagram (grafik 6.3). Za crtanje dijagrama korišćena je medijana, jer Mann-Whitney U test ispituje jednakost medijana.

Grafik 6.3. Medijana citokina 5 u odnosu na pojavu bolesti



6.5. Pol i pojava bolesti

Povezanost pojave bolesti sa polom ispitana je koristeći tabele kontigencije (tabele 6.11).

Tabela 6.11. Povezanost pojave bolesti sa polom

		pol		Total
		Ženski	Muški	
Pojava bolesti	Zdravi	0.00%	20.80%	20.80%
	Bolesni	47.00%	32.10%	79.20%
Ukupno		47.00%	53.00%	100.00%

Kako više od 80% ćelija zadovoljava pretpostavku o minimalnoj očekivanoj ćelijskoj frekvenciji, znači da nije prekršena pretpostavka (minimalna očekivana vrednost frekvencija je 16.46). χ^2 test nezavisnosti (uz korekciju neprekidnosti prema Yates'-u) pokazao je značajnu vezu između pola i pojave bolesti, $\chi^2(1, N = 168) = 36.895$, $p < 0.0005$, *Fi koeficijent* = - 0.483.

Iz tabele 6.11 se vidi da grupu zdravih čine samo muškarci, njih 20.80% od celog uzorka. Bolesni čine 79.20% celog uzorka, od toga 47% otpada na žene, a 32.10% na muškarce. Od ukupnog uzorka 47% čine žene, a 53% muškarci.

6.6. Starost i pojava bolesti

Kako bi se odlučilo koji je statistički test potreban za ispitivanje statistički značajne razlike između starosti i pojave bolesti, neophodno je prvo proveriti da li starost ima normalnu raspodelu i u grupi zdravih i u grupi bolesnih. Kolmogorov-Smirnov test tvrdi da starost ima normalnu raspodelu u obe grupe (tabela 6.12).

Tabela 6.12. Kolmogorov-Smirnov test normalnosti

		Kolmogorov-Smirnov ^a		
		Statistika	df	p
starost	Zdravi	0.114	35	0.200*
	Bolesni	0.074	136	0.069

a. Lilliefors korekcija

*. Ovo je donja granica praga značajnosti.

Kako je pretpostavka o normalnosti zadovoljena, za testiranje statističke značajnosti uticaja starosti na pojavu bolesti korišćen je t – test nezavisnih uzoraka. Postoji jedna netipična vrednost starosti, koja se zadržava u analizi jer rezultate značajno ne menja. T – testom nezavisnih uzoraka upoređena je starost između zdravih i bolesnih. Levene-ov test za jednakost varijansi nije značajan ($p = 0.36$), što implicira da su varijanse jednake, te

da se može primeniti t - test za nezavisne uzorke. T – test za nezavisne uzorke je otkrio da razlika srednjih vrednosti starosti između zdravih (35.46 ± 10.56) i bolesnih, (34.65 ± 9.53) nije statistički značajna, $t(169) = 0.44, p = 0.66$ (obostrano), eta kvadrat = 0.001.

6.7. Biohemijski parametri i bolesni (akutni i hronični)

Kako bi se izabrala odgovarajuća statistička metoda za ispitivanje statističke značajnosti razlika srednjih vrednosti biohemijskih parametara (prvo merenje) u odnosu na bolesne (akutne i shizofrenične - hronične), potrebno je proveriti da li biohemijski parametri imaju normalnu raspodelu u obe grupe.

Kako je u svakoj grupi obim uzorka veći od 30, koristi se Kolmogorov-Smirnov test normalnosti. Iz tabele 6.13 se vidi da je p vrednost Kolmogorov-Smirnovog testa u bar jednoj grupi kod obeležja leukociti, prokalcitonin, limfociti, mokraćna kiselina, glu-loza, urea, kreatinin, holesterol, trigliceridi, LDL-holesterol, bilirubin, proteini, gvožđe, aspartat aminotransperaze, alanin aminotransperaze, alkalna fosfataza, glutamil, keratin kineza, C reaktivni protein, kalijum i sedimentacija manja od 0.05, što implicira da navedene promenljive nemaju normalnu raspodelu po grupama bolesnika.

Tabela 6.13. Kolmogorov-Smirnov test normalnosti

	Boles- ni	Kolmogorov-Smirnov ^a				Boles- ni	Kolmogorov-Smirnov ^a		
		Statistika	df	p			Statistika	df	p
Leukociti	Akutni	0.117	89	0.004	LDL-holesterol	Akutni	0.148	89	< 0.0005
	Hronični	0.118	47	0.103		Hronični	0.137	47	0.027
Eritrociti	Akutni	0.051	89	0.200*	Bilirubin	Akutni	0.233	89	< 0.0005
	Hronični	0.097	47	0.200*		Hronični	0.218	47	< 0.0005
Hemoglobin	Akutni	0.084	89	0.169	Probeini	Akutni	0.153	89	< 0.0005
	Hronični	0.088	47	0.200*		Hronični	0.116	47	0.133
Hematokrit	Akutni	0.049	89	0.200*	Gvožđe	Akutni	0.093	89	0.057
	Hronični	0.094	47	0.200*		Hronični	0.235	47	< 0.0005
Trombociti	Akutni	0.057	89	0.200*	Aspartat aminotransperaze	Akutni	0.276	89	< 0.0005
	Hronični	0.102	47	0.200*		Hronični	0.306	47	< 0.0005
Prokalcitonin	Akutni	0.06	89	0.200*	Alanin aminotransperaze	Akutni	0.205	89	< 0.0005
	Hronični	0.156	47	0.006		Hronični	0.298	47	< 0.0005
Limfociti	Akutni	0.102	89	0.023	Alkalna fosfataza	Akutni	0.144	89	< 0.0005
	Hronični	0.078	47	0.200*		Hronični	0.081	47	0.200*
Mokraćna kiselina	Akutni	0.103	89	0.021	Glutamil	Akutni	0.377	89	< 0.0005
	Hronični	0.189	47	< 0.0005		Hronični	0.253	47	< 0.0005
Glukoza	Akutni	0.179	89	< 0.0005	Keratin kineza	Akutni	0.332	89	< 0.0005
	Hronični	0.086	47	0.200*		Hronični	0.319	47	< 0.0005
Urea	Akutni	0.169	89	< 0.0005	C reaktivni protein	Akutni	0.315	89	< 0.0005
	Hronični	0.083	47	0.200*		Hronični	0.335	47	< 0.0005
Kreatinin	Akutni	0.099	89	0.032	Kalijum	Akutni	0.115	89	0.006
	Hronični	0.145	47	0.014		Hronični	0.088	47	0.200*
Hholesterol	Akutni	0.212	89	< 0.0005	Natrijum	Akutni	0.075	89	0.200*
	Hronični	0.125	47	0.063		Hronični	0.115	47	0.143
Trigliceridi	Akutni	0.455	89	< 0.0005	Sedimentacija	Akutni	0.143	89	< 0.0005
	Hronični	0.256	47	< 0.0005		Hronični	0.143	47	0.017
HDL- holesterol	Akutni	0.092	89	0.058	a. Lilliefors-ova korekcija				
	Hronični	0.113	47	0.175	*. Ovo je donja granica praga značajnosti.				

Kako navedena obeležja nemaju normalnu raspodelu po grupama bolesnika, za testiranje razlika srednjih vrednosti navedenih biohemijskih parametara između akutnih i hroničnih korišćen je Mann-Whitney U test.

Tabela 6.14. Test statistika Mann-Whitney U testa

	Mann-Whitney U	Z	p (dvostrano)
Leukociti	1895.500	-0.897	0.370
Prokalcitonin	1442.500	-2.970	0.003
Limfociti	1866.500	-1.030	0.303
Mokraćna kiselina	1662.500	-1.963	0.050
Glukoza	1964.500	-0.581	0.561
Urea	1759.000	-1.522	0.128
Kreatinin	1946.000	-0.666	0.505
Holesterol	1929.000	-0.744	0.457
Trigliceridi	1917.500	-0.796	0.426
LDL-holesterol	1721.500	-1.694	0.090
Bilirubin	1528.500	-2.576	0.010
Proteini	2005.000	-0.396	0.692
Gvožđe	2004.500	-0.398	0.691
Aspartat aminotransperaze	1944.500	-0.673	0.501
Alanin aminotransperaze	2010.500	-0.371	0.711
Alkalna fosfataza	1727.500	-1.666	0.096
Glutamil	2064.500	-0.124	0.902
Keratin kineza	2042.500	-0.224	0.823
C reaktivni protein	1879.500	-0.970	0.332
Kalijum	2041.500	-0.229	0.819
Sedimentacija	1649.500	-2.024	0.043

a. grupišuća promenljiva: Bolest

Mann-Whitney U test nije otkrio statistički značajnu razliku u vrednosti leukocita, limfocita, glukoze, uree, kreatinina, holesterola, triglicerida, LDL-holesterola, proteina, gvožđa, aspartat aminotransperaze, alanin aminotransperaze, alkalne fosfataze, glutamila, keratin kineza, C reaktivnog proteina, kalijuma i sedimentacije između akutnih i hroničnih (tabela 6.14). Isti test tvrdi da je razlika srednjih vrednosti prokalcitonina između akutnih 0.234 (0.195 – 0.257) i hroničnih 0.261 (0.217 – 0.282) statistički značajna (Grupa akutni $N = 89$, Grupa hronični $N = 47$), $p = 0.003$, $r = 0.25$. Mann-Whitney U test je otkrio da je razlika srednjih vrednosti mokraćne kiseline između akutnih 328 (256.5 – 414.5) i hroničnih 283 (206 – 351) statistički značajna (Grupa akutni $N = 89$, Grupa hronični $N = 47$), $p = 0.05$, $r = 0.17$. Test je takođe otkrio da je razlika srednjih vrednosti bilirubina između akutnih 10.2 (7.2 – 18.98) i hroničnih 7.4 (6.2 – 11.1) statistički značajna (Grupa akutni $N = 89$, Grupa hronični $N = 47$), $p = 0.01$, $r = 0.22$. Isti test je otkrio da je razlika srednjih vrednosti sedimentacije iz-

među akutnih 14 (7.89 – 20) i hroničnih 18 (8 – 33) statistički značajna (Grupa akutni $N = 89$, Grupa hronični $N = 47$), $p = 0.043$, $r = 0.17$. U tabeli 6.15 prikazane su srednje vrednosti biohemijskih parametara prema bolesti, za koje je razlika srednjih vrednosti između akutnih i hroničnih (shizofreničnih) statistički značajna.

Tabela 6.15. Vrednosti biohemijskih parametara prema bolesti

	Akutni	Hronični	p
	Medijana (25. perc – 75. perc)	Medijana (25. perc – 75. perc)	
Prokalcitonin	0.234 (0.195 – 0.257)	0.261 (0.217 – 0.282)	0.003
Mokraćna kiselina	328 (256.5 – 414.5)	283 (206 - 351)	0.050
Bilirubin	10.2 (7.2 – 18.98)	7.4 (6.2 – 11.1)	0.010
Sedimentacija	14 (7.89 - 20)	18 (8 - 33)	0.043

Iz tabele 6.13 se vidi da promenljive eritrociti, hemoglobin, hematokrit, trombociti, HDL-holesterol i natrijum imaju normalnu raspodelu po grupama bolesnih, što implicira da se za poređenje navedenih obeležja po grupama koristi t-test za nezavisne uzorke. Kako se srednje vrednosti celog uzorka i srednje vrednosti bez 5% gornjih i 5% donjih slučajeva ne razlikuju mnogo kod navedenih obeležja, nema razloga za brigu o postojanju netipičnih tačaka.

Tabela 6.16. T – test za nezavisne uzorke

	Levene-ov test za jednakost varijansi		t-test za jednakost srednjih vrednosti						
	F	p	t	df	p (dvostrano)	Prosečna razlika	Std. greška razlike	95% CI za razlike	
								Donja granica	Gornja granica
Eritrociti	1.093	0.298	1.011	134	0.314	0.08656	0.08565	-0.08285	0.25596
Hemoglobin	1.650	0.201	2.276	134	0.024	5.91072	2.59724	0.77382	11.04761
Hematokriti	1.419	0.236	1.793	134	0.075	0.01330	0.00742	-0.00137	0.02798
Trombociti	2.162	0.144	-1.994	134	0.048	-26.68152	13.37883	-53.14252	-0.22053
Holesterol	2.141	0.146	-1.202	134	0.232	-0.09704	0.08076	-0.25677	0.06268
Natrijum	0.153	0.696	-0.413	134	0.681	-0.40561	0.98317	-2.35014	1.53892

Levene-ov test za jednakost varijansi nije značajan ni kod jedne promenljive (tabela 6.16), što implicira da su varijanse jednake kod svake promenljive po grupama. Razlika srednjih vrednosti hemoglobina između akutnih (140.78 ± 12.96) i hroničnih (134.87 ± 16.82), je statistički značajna, $t(134) = 2.276$, $p = 0.024$ (dvostrano), eta kvadrat = 0.04. Razlika srednjih vrednosti trombocita je statistički značajna između akut-

nih (272.18 ± 65.29) i hroničnih (298.86 ± 88.78), $t(134) = -1.994$, $p = 0.048$ (dvostrano), eta kvadrat = 0.029.

U tabeli 6.17 prikazane su vrednosti biohemijskih parametara prema bolesti, za koje je razlika srednjih vrednosti između akutnih i hroničnih (shizofreničnih) statistički značajna.

Tabela 6.17. Vrednosti biohemijskih parametara prema bolesti

	Akutni	Shizofrenija	p
	Aritm. Sredina \pm SD	Aritm. Sredina \pm SD	
Hemoglobin	140.78 \pm 12.96	134.87 \pm 16.82	0.024
Trombociti	272.18 \pm 65.29	298.86 \pm 88.78	0.048

6.8. PANSS i bolesni

Kako bi se izabrala odgovarajuća statistička metoda za ispitivanje statističke značajnosti razlika srednjih vrednosti parametara PANSS skale (prvo merenje) u odnosu na bolesne (akutne i shizofrenične - hronične), potrebno je proveriti da li parametri PANSS skale imaju normalnu raspodelu u obe grupe.

Tabela 6.18. Kolmogorov-Smirnov test normalnosti

	Bolesni	Kolmogorov-Smirnov ^a		
		Statistika	df	p
P1	Akutni	0.151	89	< 0.0005
	Hronični	0.108	47	0.200*
N1	Akutni	0.085	89	0.151
	Hronični	0.092	47	0.200*
G1	Akutni	0.093	89	0.057
	Hronični	0.096	47	0.200*
PANSS1	Akutni	0.111	89	0.009
	Hronični	0.115	47	0.151

a. Lilliefors-ova korekcija

*. Ovo je donja granica praga značajnosti.

Kako je u svakoj grupi obim uzorka veći od 30, koristi se Kolmogorov-Smirnov test normalnosti. Iz tabele 6.18 se vidi da je p vrednost Kolmogorov-Smirnovog testa u bar jednoj grupi kod obeležja P1, PANSS1 manja od 0.05, što implicira da promenljive nemaju P1 i PANSS1 nemaju normalnu raspodelu po grupama bolesnika.

Kako promenljive P1 i PANSS1 nemaju normalnu raspodelu po grupama bolesnika, za testiranje razlika srednjih vrednosti navedenih parametara između akutnih i hroničnih korišćen je Mann-Whitney U test.

Tabela 6.19. Test Statistika^a

	P1	PANSS1
Mann-Whitney U	1827.000	1738.500
Z	-1.212	-1.616
p (dvostrano)	.225	.106

a. grupišuća promenljiva: bolest

Mann-Whitney U test nije otkrio statistički značajnu razliku srednjih vrednosti obeležja P1, ni srednjih vrednosti obeležja PANSS1 između akutnih i hroničnih (tabela 6.19).

Iz tabele 6.18 se vidi da promenljive N1 i G1 imaju normalnu raspodelu po grupama bolesnih, što implicira da se za poređenje navedenih obeležja po grupama koristi t-test za nezavisne uzorke. Kako se srednje vrednosti celog uzorka i srednje vrednosti bez 5% gornjih i 5% donjih slučajeva ne razlikuju mnogo ni kod promenljive N1 ni kod promenljive G1, nema razloga za brigu o postojanju netipičnih tačaka. Levane-ov test za jednakost varijansi nije značajan kod promenljive N1 ($p = 0.863$), a ni kod promenljive G1 ($p = 0.49$), što implicira da su varijanse promenljivih N1 i G1 jednake između populacija. Međutim, t-test za nezavisne uzorke pokazuje da razlika srednjih vrednosti N1 između akutnih (22.63 ± 7.06) i hroničnih (23.40 ± 7.09) nije statistički značajna, $t(134) = -0.599$, $p = 0.550$ (dvostrano), $eta\ kvadrat = 0.003$. Razlika srednjih vrednosti G1 između akutnih (53.54 ± 7.80) i hroničnih (54.46 ± 6.77) nije statistički značajna, $t(134) = -0.685$, $p = 0.495$ (dvostrano), $eta\ kvadrat = 0.003$.

6.9. Promene vrednosti parametara PANSS skale između dva merenja

Kako ispitujemo merenje istih parametara u dva navrata, potrebno je utvrditi da li razlike parova imaju normalnu raspodelu. Zbog toga se izračunavaju nove promenljive koje predstavljaju razlike parametara PANSS skale merenih pre terapije i 30 dana nakon početka terapije.

Tabela 6.20. Kolmogorov-Smirnov test normalnosti

	Kolmogorov-Smirnov ^a		
	Statistika	df	p
razlikaP	0.087	136	0.014
razlikaN	0.055	136	0.200 [*]
razlikaG	0.059	136	0.200 [*]
razlikaPANSS	0.064	136	0.200 [*]

a. Lilliefors – ova korekcija

*. This is a lower bound of the true significance.

Kolmogoriv – Smirnov test (tabela 6.20) pokazuje da razlika para P1 i P2 nema normalnu raspodelu, te se za poređenje srednjih vrednosti razlika između dva merenja koristi neparametarski Wilcoxon-ov test. Iz tabele 6.20 se vidi da razlike parova ostalih promenljivih imaju normalnu raspodelu, te se za poređenje srednjih vrednosti razlika između dva merenja koristi parametarski t-test uparenih uzoraka.

Tabela 6.21. Test Statistika^b

	P2 - P1
Z	-10.014 ^a
p (obostrano)	<0.0005

a. Zasnovan na pozitivnim rangovima

b. Wilcoxon – ov test

Kao što se vidi iz tabele 6.21, Wilcoxon-ov test ranga je otkrio statistički značajno smanjenje vrednosti parametra pozitivne skale PANSS testa nakon početka terapije, uz veliki uticaj ($r = 0.61$). Medijana vrednosti parametara pozitivne skale PANSS testa je opala od 26 (22 – 29) na početku terapije do 12 (9 - 16) nakon terapije.

T – testom uparenih uzoraka (ponovljenih merenja) utvrđeno je statistički značajno smanjenje vrednosti parametara negativne skale PANSS testa pre početka terapije (22.90 ± 7.05) i nakon terapije (14.71 ± 6.82), $t(135) = 13.722$, $p < 0.0005$ (dvostrano). Vrednost eta kvadrat (0.58) pokazuje da je veoma velika razlika rezultata između dva merenja.

Istim testom utvrđeno je statistički značajno smanjenje vrednosti parametara opšte psihopatološke skale PANSS testa pre početka terapije (53.86 ± 7.45) i nakon terapije (32.43 ± 11.41), $t(135) = 22.169$, $p < 0.0005$ (dvostrano). Vrednost eta kvadrat (0.78) pokazuje da je veoma velika razlika rezultata između dva merenja.

Takođe je t - testom uparenih uzoraka utvrđeno statistički značajno smanjenje ukupne vrednosti skale PANSS testa pre početka terapije (102.46 ± 16.29) i nakon terapije (60.13 ± 21.93), $t(135) = 21.69$, $p < 0.0005$ (dvostrano). Vrednost eta kvadrat (0.78) pokazuje da je veoma velika razlika rezultata između dva merenja.

6.10. Povezanost citokina i ostalih parametara

Veza između citokina (prvo merenje) i ostalih parametara (prvo merenje) izražena je pomoću Pirsonovog koeficijenta korelacije (tabela 6.22).

Citokin 1 je u pozitivnoj korelaciji sa starošću, parametrima negativne skale PANSS testa, opštim psihopatološkim parametrima PANSS testa, leukocitima, hemoglobinom, hematokritom, trombocitima ($r = 0.177$, $p = 0.04$), prokalcitoninom, mokraćnom kiselinom, ureom ($r = 0.284$, $p = 0.001$), kreatininom, trigliceridima, HDL-holesterolom, proteinima, gvoždem, alanin aminotransperazom, alkalnom fosfatazom, kreatin kinezom, C reaktivnim proteinom i sedimentacijom.

Citokin 1 je u negativnoj korelaciji sa parametrima pozitivne skale PANSS testa, sa ukupnom vrednošću PANSS testa, eritrocitima, limfocitima, glukozom, kolesterolom, LDL-holesterolom, bilirubinom, aspartat aminotransperazom, glutamilom, kalijumom i natrijumom.

Citokin 2 je u pozitivnoj korelaciji sa parametrima pozitivne skale PANSS testa, opštim psihopatološkim parametrima PANSS testa, ukupnom vrednošću PANSS testa, eritrocitima, hemoglobinom, hematokritom, trombocitima, prokalcitoninom, limfocitima, mokraćnom kiselinom, glukozom, ureom, kreatininom, kolesterolom, trigliceridima, HDL-holesterolom, bilirubinom, proteinima, alkalnom fosfatazom, glutamilom, kreatin kinezom, C reaktivnim proteinom i kalijum.

Citokin 2 je u negativnoj korelaciji sa starošću, parametrima negativne skale PANSS testa, leukocitima, LDL-holesterolom, gvoždem, aspartat aminotransperazom, alanin aminotransperazom, natrijumom i sedimentacijom.

Tabela 6.22. Koeficijenti Pirsonove korelacije između citokina o pstalih parametara

	Citokin 1	Citokin 2	Citokin 3	Citokin 4	Citokin 5
starost	0	-0.006	-0.051	-0.162	0.116
P1	-0.041	0.116	0.045	-0.037	0.045
N1	0.042	-0.066	-0.089	0.008	0.064
G1	0.006	0.116	0.097	-0.038	0.05
PANSS1	-0.079	0.073	0.029	-0.024	0.053
Leukociti	0.017	-0.075	0	-0.046	-0.042
Eritrociti	-0.008	0.072	0.126	0.042	0.135
Hemoglobin	0.027	0.005	0.063	0.014	0.126
Hematokrit	0.002	0.039	0.046	0.043	0.161
Trombociti	.177*	0.007	0.029	0.016	0.132
Prokalcitonin	0.109	0.026	0.035	0.005	0.136
Limfociti	-0.11	0.063	-0.023	-0.1	.198*
Mokraćna kiselina	0.01	0.029	0.052	-0.092	.234**
Glukoza	-0.028	0.142	0.076	0.007	0.135
Urea	.284**	0.066	0.036	0.084	0.057
Kreatinin	0.065	0.007	0.07	-0.048	.250**
Holesterol	-0.077	0.086	.259**	-0.07	-0.026
Trigliceridi	0.019	0.099	0.05	-0.036	-0.032
HDL-holesterol	0.054	0.039	0.062	.174*	0.151
LDL-holesterol	-0.113	-0.023	-0.045	-0.094	0.015
Bilirubin	-0.046	0.094	-0.056	.191*	0.045
Proteini	0.052	0.06	0.034	0.077	0.129
Gvožđe	0.034	-0.028	-0.027	-0.048	0.155
Aspartat aminotransperaze	-0.011	-0.027	-0.059	0.027	0.089
Alanin aminotransperaze	0.035	-0.035	-0.018	0.001	0.093
Alkalna fosfataza	0.007	0.019	-0.004	0.015	0.08
Glutamil	-0.023	0.069	-0.039	-0.021	0.02
Keratin kineza	0.005	0.026	0.008	0.006	0.003
C reaktivni protein	0.011	0.011	-0.03	0.006	0.053
Kalijum	-0.022	0.004	0.037	-0.032	-0.082
Natrijum	-0.041	-0.054	-0.127	0.052	0.082
Sedimentacija	0.022	-0.005	0.027	-0.085	-0.059

* $p < 0.05$; ** $p < 0.01$

Citokin 3 je u pozitivnoj korelaciji sa parametrima pozitivne skale PANSS testa, opštim psihopatološkim parametrima PANSS testa, ukupnom vrednošću PANSS testa, eritrocitima, hemoglobinom, hematokritom, trombocitima, prokalcitoninom, mokraćnom kiselinom, glukozom, ureom, kreatininom, holesterolom ($r = 0.259$, $p = 0.002$), trigliceridima, HDL-holesterolom, proteinima, kreatin kinezom, kalijumom i sedimentacijom.

Citokin 3 je u negativnoj korelaciji sa starošću, parametrima negativne skale PANSS testa, limfocitima, LDL-holesterolom, bilirubinom, gvožđem, aspartat aminosperazom, alanin aminotransperazom, alkalnom fosfatazom, glutamilom, C reaktivnim proteinom i natrijumom. Leukociti i citokin 3 nisu u korelaciji.

Citokin 4 je u pozitivnoj korelaciji sa parametrima negativne skale PANSS testa, eritrocitima, hemoglobinom, hematokritom, trombocitima, prokalcitoninom, glukozom, ureom, HDL-holesterolom ($r = 0.174$, $p = 0.043$), bilirubinom ($r = 0.191$, $p = 0.026$), proteinima, aspartat aminotransperazom, alanin aminotransperazom, alkalnom fosfatazom, kreatin kinezom, C reaktivnim proteinom i natrijumom.

Citokin 4 je u negativnoj korelaciji sa starošću, parametrima pozitivne skale PANSS testa, opštim psihopatološkim parametrima PANSS testa, ukupnom vrednošću PANSS testa, leukocitima, limfocitima, mokraćnom kiselinom, kreatininom, holesterolom, trigliceridima, LDL-holesterolom, gvožđem, glutamilom, kalijumom i sedimentacijom.

Citokin 5 je u pozitivnoj korelaciji sa starošću, parametrima pozitivne skale PANSS testa, parametrima negativne skale PANSS testa, opštim psihopatološkim parametrima PANSS testa, ukupnom vrednošću PANSS testa, eritrocitima, hemoglobinom, hematokritom, trombocitima, prokalcitoninom, limfocitima ($r = 0.198$, $p = 0.021$), mokraćnom kiselinom ($r = 0.234$, $p = 0.006$), glukozom, ureom, kreatininom ($r = 0.25$, $p = 0.003$), HDL-holesterolom, LDL-holesterolom, bilirubinom, proteinima, aspartat aminotransperazom, alanin aminotransperazom, alkalnom fosfatazom, glutamilom, kreatin kinezom, C reaktivnim proteinom i natrijumom.

Citokin 5 je u negativnoj korelaciji sa leukocitima, holesterolom, trigliceridima, kalijumom i sedimentacijom.

6.11. Od čega zavisi pojava bolesti?

Binarna logistička regresija bila je sprovedena kako bi se ocenio uticaj citokina na verovatnoću da će se pojaviti bolest. Urađene su univarijantne binarne logističke regresije za promenljive pol, starost, terapiju i prva merenja citokina 1, citokina 2, citokina 3, citokina 4 i citokina 5. Kako samo citokin 5 značajno utiče na pojavu bolesti, u tabeli 6.23 prikazani su rezultati. Model je je statistički značajan, $\chi^2(1, N = 171) = 23.62$,

$p < 0.001$, što pokazuje da model razlikuje zdrave i bolesne ispitanike. Model u celini objašnjava između 12.9% (R^2 Cox-a i Shnell-a) i 20.3% (R^2 Nagalkerke-a) varijanse pojave bolesti i tačno klasifikuje 80.1% slučajeva. Količnik verovatnoće za citokin 5 je 1.035. To pokazuje da ako se citokin 5 poveća za 1, rizik da se osoba razboli se poveća za 3.5%.

Tabela 6.23. Predviđanje verovatnoće da će se bolest pojaviti

						95% C.I.za EXP(B)			
		Stand.				Donja	Gornja		
	B	Greška	Wald	df	p	Exp(B)	granica	granica	
Korak 1^a	Citokin 5	.034	.010	11.817	1	.001	1.035	1.015	1.055
	Konstanta	.199	.321	.383	1	.536	1.220		

a. Promenljiva uključena u prvi korak je citokin 5.

Dobijena jednačina binarne logističke regresije je: $\ln \frac{p}{1-p} = 0.199 + 0.034 * Citokin5$.

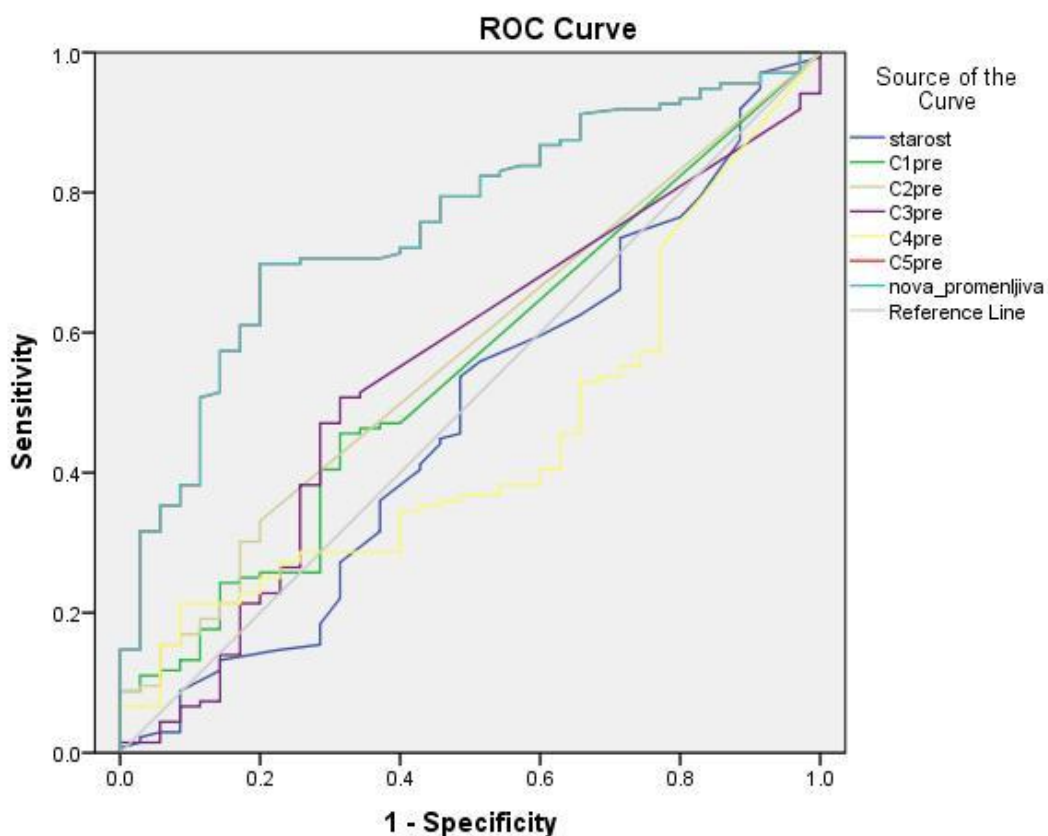
Eksplisitni oblik jednačine binarne logističke regresije je:

$$p = \frac{e^{0.199+0.034*Citokin5}}{1 + e^{0.199+0.034*Citokin5}}$$

6.12. Potencijalni markeri za pojavu bolesti

Kako bi se odredili potencijalni markeri za pojavu bolesti, sprovedena je ROC analiza. Analiza sadrži sedam nezavisnih promenljivih (starost, novu promenljivu, citokin 1, citokin 2, citokin 3, citokin 4 i citokin 5 – prvo merenje). Nova promenljiva formirana je koristeći jednačinu binarne logističke regresije (odeljak 6.11). Na grafiku 6.4 su prikazane ROC krive, koje će se dalje ispitivati.

Grafik 6.4. ROC krive



Diagonal segments are produced by ties.

Kako je u tabeli 6.24 dobijeno za citokin 5 (prvo merenje) površina = 0.755 > 0.7, $p < 0.0005$, citokin 5 može da bude marker za pojavu bolesti. Iz tabele 6.25 se vidi da je proizvod senzitivnosti i specifičnosti maksimalan, ako je vrednost citokina 5 jednaka 21.955 (senzitivnost = 0.699, a specifičnost = $1 - 0.2 = 0.8$). Sa grafika 6.4 kao i iz tabele 6.24 se vidi da se površine ispod krivih poklapaju za citokin 5 i novu promenljivu.

Tabela 6.24. Površina ispod ROC krive

Testirane promenljive	Površina	Std. greška ^a	p ^b	Asimptotski 95% interval poverenja	
				Donja granica	Gornja granica
Starost	0.480	0.057	0.712	0.368	0.591
Citokin 1	0.546	0.053	0.398	0.442	0.650
Citokin 2	0.567	0.052	0.222	0.465	0.669
Citokin 3	0.547	0.055	0.395	0.439	0.654
Citokin 4	0.446	0.052	0.323	0.344	0.547
Citokin 5	0.756	0.042	< 0.0005	0.673	0.839
Nova promenljiva	0.756	0.042	< 0.0005	0.673	0.839

a. Pod neparametarskom pretpostavkom

b. Nulta hipoteza: tačna površina = 0.5

Tabela 6.25. Koordinate ROC krive

	Pozitivno ako je veće ili jednako od ^a	Senzitivnost	1 - Specifičnost
Citokin 5	-60.386391	1.000	1.000
	-29.693196	1.000	0.971
	2.473251	0.985	0.971

	21.800739	0.699	0.257
	21.955579	0.699	0.200
	22.056583	0.691	0.200

	257.506341	0.007	0.000
	287.352382	0.000	0.000

a. Najmanja granična vrednost je minimalna posmatrana vrednost minus 1, a najveća posmatrana granična vrednost je maksimalna vrednost plus 1.

Po definiciji senzitivnost i specifičnost se mogu izračunati koristeći tabele kontigencije. Tabela 6.26 predstavlja tabelu kontigencije za pojavu bolesti i klasifikaciju. Klasifikacija je dihotomna promenljiva, koja uzima vrednost 0 = "negativni", ako je vrednost citokina 5 manja od 21.955, a vrednosti 1 = "pozitivni", ako je vrednost citokina 5 veća od 21.955.

Tabela 6.26. Tabela kontigencije

		pojava bolesti		Ukupno
		Zdravi	Bolesni	
Klasifikacija	Negativni	28	41	69
	Pozitivni	7	95	102
Ukupno		35	136	171

Po definiciji dobija se:

$$\text{senzitivnost} = \frac{95}{41 + 95} = 0.699$$

$$\text{specifičnost} = \frac{28}{28 + 7} = 0.8.$$

6.13. Promene citokina između dva merenja

Kako ispitujemo merenje istih citokina u dva navrata, potrebno je utvrditi da li razlike parova imaju normalnu raspodelu. Zbog toga se izračunavaju nove promenljive koje predstavljaju razlike citokina merenih pre terapije i 30 dana nakon početka terapije.

Tabela 6.27. Kolmogorov-Smirnov test normalnosti

	Kolmogorov-Smirnov ^a		
	Statistika	df	p
razlikaC1	0.273	136	< 0.0005
razlikaC2	0.279	136	< 0.0005
razlikaC3	0.207	136	< 0.0005
razlikaC4	0.221	136	< 0.0005
razlikaC5	0.115	136	< 0.0005

a. Lilliefors korekcija

Kolmogoriv – Smirnov test (tabela 6.27) pokazuje da razlike parova nemaju normalnu raspodelu, te se za poređenje srednjih vrednosti razlika između dva merenja koristi neparametarski Wilcoxon-ov test.

Tabela 6.28. Test Statistika^c

	C1posle - C1pre	C2posle - C2pre	C3posle - C3pre	C4posle - C4pre	C5posle - C5pre
Z	-0.578 ^a	-1.151 ^b	-2.771 ^b	-0.603 ^a	-2.709 ^a
p (dvostrano)	0.563	0.250	0.006	0.547	0.007

a. Zasnovan na negativnim rangovima.

b. Zasnovan na pozitivnim rangovima.

c. Wilcoxon-ov test rangova.

Kao što se vidi iz tabele 6.28, Wilcoxon-ov test ranga je otkrio statistički značajno smanjenje vrednosti citokina 3 nakon početka terapije, uz mali uticaj ($r = 0.17$). Test je takođe otkrio statistički značajno povećanje vrednosti citokina 5 30 dana nakon početka terapije, uz mali uticaj ($r = 0.16$). U tabeli 6.29 prikazane su razlike citokina između merenja.

Tabela 6.29. Razlike citokina između merenja

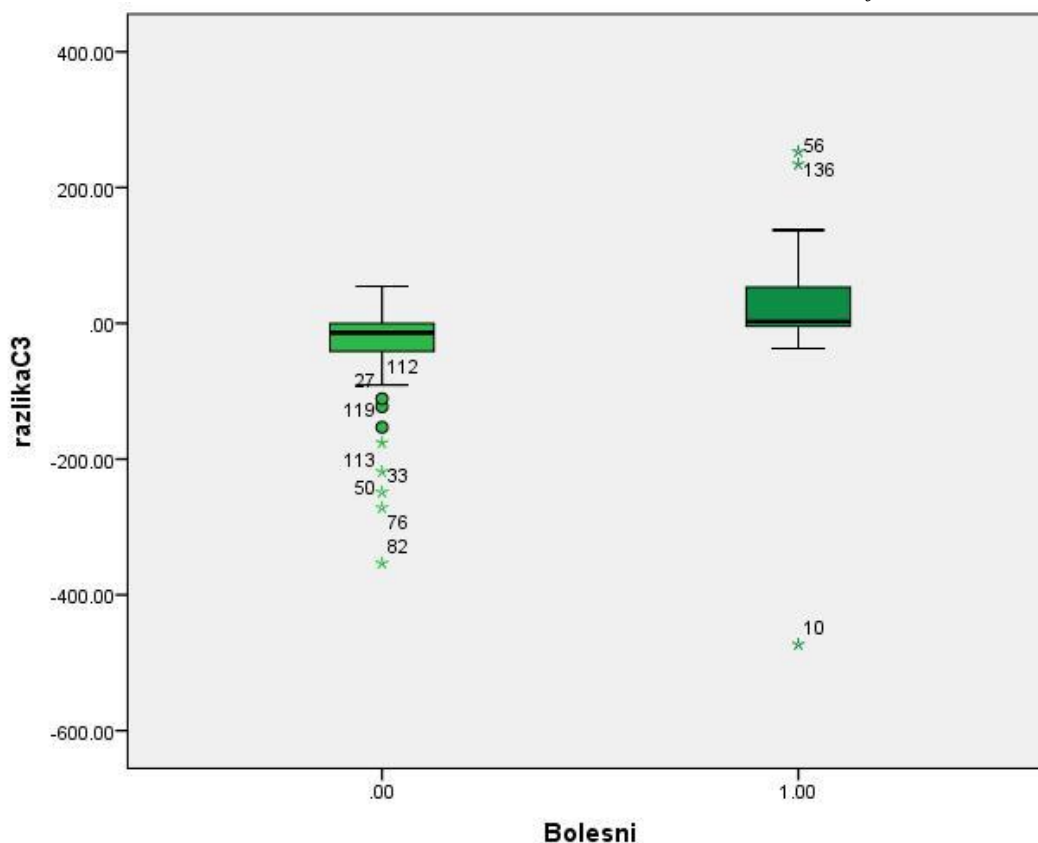
	Prvo merenje	Drugo merenje	p
	Medijana (25. perc – 75. perc)	Medijana (25. perc – 75. perc)	
Citokin 1	0.00 (0.00 – 38.33)	3.03 (0.00 – 59.34)	0.563
Citokin 2	0.00 (0.00 – 11.12)	0.00 (0.00 – 5.73)	0.250
Citokin 3	0.90 (0.00 – 14.98)	0.00 (-10.47 – 17.96)	0.006
Citokin 4	856.85 (0.00 – 4173.14)	1265.97 (0.00 – 4010.24)	0.547
Citokin 5	39.06 (18.75 – 81.65)	58.93 (25.90 – 96.60)	0.007

6.14. Šta utiče na promene citokina između dva merenja?

Promena citokina 3 i citokina 5 između muškaraca i žena ispitana je Mann-Whitney U testom. Mann – Whitney U test nije otkrio statistički značajnu razliku u promeni citokina 3 između žena -4.44 ($-25.93 - 7.79$) i muškaraca -10.11 ($-34.87 - 2.61$), $U = 1927$, $z = -0.945$, $p = 0.345$, $r = 0.08$. Isti test je pokazao da nije statistički značajna razlika u promeni citokina 5 između žena 12.01 ($-12.05 - 6.9$) i muškaraca 20.40 ($-19.13 - 57.84$), $U = 1992$, $z = -0.646$, $p = 0.518$, $r = 0.06$.

Promena citokina 3 i citokina 5 između bolesnih (akutnih i shizofreničnih) takođe je ispitana Mann-Whitney U testom. Mann-Whitney U test je otkrio statistički značajnu razliku u promeni citokina 3 između akutnih -13.93 ($-42.12 - 0.00$) i shizofreničnih 2.67 ($-4.44 - 54.03$), $U = 882$, $z = -5.54$, $p < 0.0005$, $r = 0.47$.

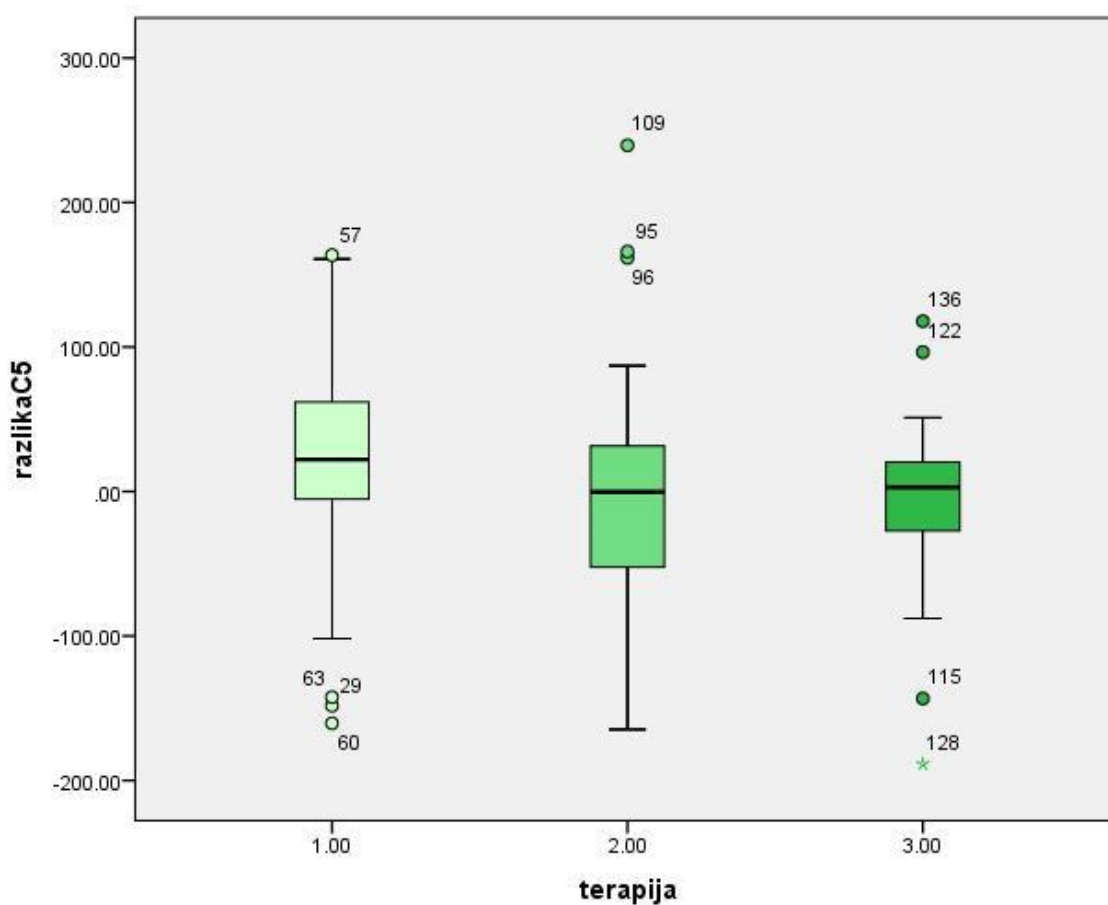
Grafik 6.5. Promena citokina 3 između akutnih i shizofreničnih



Isti test nije otkrio statistički značajnu razliku u promeni citokina 5 kod akutnih 12.01 (−11.08 – 6.47) i shizofreničnih 20.45 (−36.88 – 67.76), $U = 2044$, $z = -0.217$, $p = 0.828$, $r = 0.2$.

Kruskal-Wallis-ov test nije otkrio statistički značajnu razliku u promeni citokina 3 između terapija (terapija 1, $N = 66$; terapija 2, $N = 40$; terapija 3, $N = 25$), $\chi^2(2, 131) = 3.779$, $p = 0.151$. Kruskal-Wallis-ov test otkrio je statistički značajnu razliku u promeni citokina 5 između terapija (terapija 1, $N = 66$; terapija 2, $N = 40$; terapija 3, $N = 25$) $\chi^2(2, 131) = 8.441$, $p = 0.015$. Srednja vrednost promene citokina 5 kod prve terapije je 22.16 (− 5.31 – 62.09). Srednja vrednost promene citokina 5 kod druge terapije je −0.29 (−54.40 – 32.23). Srednja vrednost promene citokina 5 kod treće terapije je 2.82 (−38.57 – 22.83). Na grafiku 6.6 prikazana je promena citokina 5 između terapija.

Grafik 6.6. Promena citokina 5 između terapija



Primenom Pirsonovog koeficijenta korelacije, utvrđeno je da li biohemijski parametri, parametri PANSS testa, starost i početne vrednosti citokina 3 i citokina 5 utiču na promene citokina 3 i citokina 5 između dva merenja. U tabeli 6.30 su prikazani Pirsonovi koeficijenti korelacije.

Tabela 6.30. Pirsonovi koeficijenti korelacija

	razlikaC3	razlikaC5
P1	-0.102	-0.039
N1	0.071	-0.039
G1	-0.121	0.022
PANSS1	-0.077	-0.005
starost	0.051	-0.151
C3pre	-0.828**	-0.018
C5pre	0.023	-0.630**
Leukociti	0.145	0.133
Eritrociti	-0.206*	-0.252**
Hemoglobin	-0.098	-0.179*
Hematokrit	-0.071	-0.208*
Trombociti	0.036	0.097
Prokalcitonin	0.106	0.104
Limfociti	-0.040	0.072
Mokraćna kiselina	-0.035	-0.159
Glukoza	0.040	-0.081
Urea	-0.063	-0.137
Kreatinin	-0.072	-0.111
Holesterol	-0.268**	-0.083
Trigliceridi	-0.035	0.028
HDL-holesterol	0.017	-0.084
LDL-holesterol	-0.121	0.106
Bilirubin	-0.003	-0.181*
Proteini	0.022	-0.369**
Gvožđe	0.072	-0.174*
Aspartat aminotransperaze	0.067	-0.101
Alanin aminotransperaze	0.059	-0.091
Alkalna fosfataza	0.095	-0.076
Glutamil	0.042	0.088
Keratin kineza	0.016	-0.027
C reaktivni protein	-0.089	-0.158
Kalijum	0.032	0.117
Natrijum	0.074	-0.059
Sedimentacija	0.117	-0.048

**p < 0.01 (dvostrano) *p < 0.05 (dvostrano)

Promena citokina 3 između dva merenja je u pozitivnoj korelaciji sa parametrima negativne skale PANSS testa, starošću, početnom vrednošću citokina 5, leukocitima, trombocitima, prokalcitoninom, glukozom, HDL-holesterolom, proteinima, gvožđem, aspartat aminotransperazom, alanin aminotransperazom, alkalnom fosfatazom, glutamilom, kreatin kinezom, kalijumom, natrijumom i sedimentacijom.

Promena citokina 3 između dva merenja je u negativnoj korelaciji sa parametrima pozitivne skale PANSS testa, opštim psihopatološkim parametrima PANSS testa, ukupnom vrednošću PANSS testa, početnom vrednošću citokina 3 ($r = -0.828, p < 0.0005$) eritrocitima ($r = -0.206, p = 0.016$), hemoglobinom, hematokritom, limfocitima, mokraćnom kiselinom, ureom, kreatininom, holesterolom ($r = -0.268, p = 0.002$), trigliceridima, LDL-holesterolom, bilirubinom i C reaktivnim proteinom.

Promena citokina 5 između dva merenja je u pozitivnoj korelaciji sa opštim psihopatološkim parametrima PANSS testa, leukocitima, trombocitima, prokalcitoninom, limfocitima, trigliceridima, LDL-holesterolom, glutamilom i kalijumom.

Promena citokina 5 između dva merenja je u negativnoj korelaciji sa parametrima pozitivne skale PANSS testa, parametrima negativne skale PANSS testa, ukupnom vrednošću PANSS testa, starošću, početnom vrednošću citokina 3, početnom vrednošću citokina 5 ($r = -0.63, p < 0.0005$), eritrocitima ($r = -0.252, p = 0.003$), hemoglobinom ($r = -0.179, p = 0.037$), hematokritom ($r = -0.208, p = 0.015$), mokraćnom kiselinom, glukozom, ureom, kreatininom, holesterolom, HDL-holesterolom, bilirubinom ($r = -0.181, p = 0.035$), proteinima ($r = -0.369, p < 0.0005$), gvoždem ($r = -0.174, p = 0.042$), aspartat aminotransperazom, alanin aminotransperazom, alkalnom fosfatazom, keratin kinezom, C reaktivnim proteinom, natrijumom i sedimentacijom.

7. Diskusija

Tabele kontigencije za dijagnozu i pol pokazuju da kontrolnu grupu čine samo muškarci, dok je u grupama akutnih i bolesnih bilo više žena. Kako je žena bilo 47% od ukupnog broja ispitanika, a muškaraca 53%, sledi da je ispitivani uzorak homogen po polu. Među akutnima bilo je za 31% žena i 22% muškaraca, dok je shizofreničnih bilo 16.10% žena i 10.10% muškaraca.

Poređenjem citokina između zdravih i bolesnih, pokazano je da se srednje vrednosti citokina 5 statistički značajno razlikuju između zdravih i bolesnih. Ostali citokini se ne razlikuju značajno između te dve populacije.

Ako se posmatra pojava bolesti između polova, tabele kontigencije pokazuju da bolesnih ima više među ženama (47%) nego među muškarcima (32.10%).

Starost se statistički značajno ne razlikuje ni između dijagnoza (zdravi, akutni i shizofrenija), ni između zdravih i bolesnih.

Poređenjem srednjih vrednosti biohemijskih parametara između bolesnih, pokazano je da se hemoglobin, trombociti, prokalcinotin, mokraćna kiselina, bilirubin i sedimentacija statistički značajno razlikuju između akutnih i shizofreničnih.

Srednje vrednosti parametara PANSS testa se statistički značajno ne razlikuju između akutnih i shizofreničnih, ali se statistički značajno razlikuju u promenama između merenja pre početka terapije i 30 dana nakon početka terapije.

Citokin 1 (prvo merenje) je u statistički značajnoj korelaciji sa trombocitima i ureom. Citokin 3, takođe prvo merenje, je značajno povezan sa holesterolom. Citokin 4 (prvo merenje) značajno korelira sa HDL-om i bilirubinom. Citokin 5 je u statistički značajnoj korelaciji sa limfocitima, mokraćnom kiselinom i kreatininom. Citokin 2 ne korelira značajno ni sa jednim od ostalih parametara. Sve statistički značajne korelacije su pozitivne, tako da se povećavanjem vrednosti ostalih parametara povećavaju i vrednosti citokina.

Binarna logistička regresija pokazala je da citokin 5 jedini značajno utiče na pojavu bolesti. Potencijalan marker za pojavu bolesti može biti citokin 5. Kako je senzitivnost jednaka 0.699, znači da naš test 69.9% pogađa da je ispitanik bolestan. Specifičnost je 0.8, što implicira da naš test 80% pogađa da je ispitanik zdrav.

Promene citokina između dva merenja su statistički značajne kod citokina 3 i citokina 5. Razlika u promeni citokina 3 između akutnih i shizofreničkih je statistički značajna. Promena citokina 5 između terapija je statistički značajna. Na promenu citokina 3 značajno utiče početna vrednost citokina 3 i biohemijski parametri (eritrociti i holesterol). Korelacija između razlike citokina 3 i početne vrednosti citokina 3, eritrocita i holesterola je negativna, što implicira da se povećavanjem početne vrednosti citokina 3, eritrocita i holesterola smanjuje razlika između dva merenja citokina 3. Na promenu citokina 5 značajno utiče početna vrednost citokina 5 i biohemijski parametri (eritrociti,

hemoglobin, hematokrit, bilirubin, proteini i gvožđe). Kako je korelacija razlike citokina 5 i početne vrednosti citokina 5, eritrocita, hemoglobina, hematokrita, bilirubina, proteina i gvožđa negativna, sledi da se povećavanjem početne vrednosti citokina 5, eritrocita, hemoglobina, hematokrita, bilirubina, proteina i gvožđa smanjuje razlika između dva merenja.

8. Zaključak

Statističke metode koje su primenjene u radu, pokazale su da je pol značajno povezan sa dijagnozom, kao i sa pojavom bolesti. Otkriveno je da se citokin 5 statistički značajno razlikuje i između dijagnoza i između zdravih i bolesnih, kao i da citokin 5 može biti marker za otkrivanje bolesti. Kako je citokin 5 u značajnoj korelaciji sa limfocitima, mokraćnom kiselinom i kreatininom, može se zaključiti da limfociti, mokraćna kiselina i kreatinin mogu da ukažu na pojavu bolesti.

Kako se hemoglobin, trombociti, prokalcitonin, mokraćna kiselina, bilirubin i sedimentacija statistički značajno razlikuju između akutnih i shizofreničnih, treba ih imati u vidu prilikom određivanja dijagnoze. Vrednosti na PANSS skali se statistički značajno ne razlikuje između akutnih i hroničnih, ali postoji statistički značajna razlika između dva merenja. Imajući u vidu da je došlo do pada vrednosti na PANSS skali, zaključuje se da se stanje kod pacijenata, u proseku, popravilo.

Imajući u vidu da su vrednosti citokina 5 kod bolesnih pacijenata veće nego kod zdravih pacijenata, zaključuje se da nije dobro da citokin 5 raste. Promene citokina 5 između terapija su statistički značajne, odakle sledi da promene citokina 5 zavise od terapija. Imajući u vidu srednje vrednosti ovih promena, zaključujemo da je prvi lek najbolji, jer kod njega vrednosti citokina 5 u proseku opadaju (22.16), lek 3 je dobar, jer i kod njega vrednosti citokina 5 u proseku opadaju (2.82), dok kod leka 2 rastu (-0.29).

Literatura

- [1] Andreasen, N. C. (1995). *Symptoms, signs and diagnosis of schizophrenia*. Lancet, 346, 108-112.
- [2] Aviva Petrie, Caroline Sabin, *Medical statistics at a glance*, Wiley-Blackwell, London 2007.
- [3] Bjelica Artur L., Zorić, D., Kovačević-Petljanski, V. (2003). *Značaj prenatalnih i perinatalnih faktora u etiologiji shizofrenije*. Aktualnosti iz neurologije, psihijatrije i graničnih područja, 11(2), 77-81.
- [4] Bošković, A.,(2009). *Krv*, Medicinska fiziologija – skripta za opštu medicinu, Medicinski fakultet, Novi Sad.97-122.
- [5] Čulić, S. (2005). *Citokini i autoimunosne bolesti*. Paediatr Croat, 49 (Supl 1), 148-161.
- [6] http://library.uvm.edu/services/statistics/SPSS20Manuals/IBM_SPSS_Statistics_Core_System_User_Guide.pdf
- [7] http://www.emotionalwellbeing.southcentral.nhs.uk/resources/doc_download/62-panss-positive-and-negative-syndrome-scale-pdf-document
- [8] <http://www.moj-imunitet.com/tumacenje-rezultata-analize-krvi-i-urina.html>
- [9] <http://www.stetoskop.info/Kada-je-povisen-C-reaktivni-protein-CRP-2889-s3-content.htm>
- [10] Janjičić, B., Novović, Z., (2011). *Procena uspešnosti u klasifikovanju rezultata na osnovu graničnih (cut-off) skorova: receiver operating characteristic curve*. Primenjena psihologija, 4, 335-351.
- [11] Jevremović, V., *Verovatnoća i statistika*, Matematički fakultet, Beograd 2009.
- [12] Joaquim P. Marques de Sá, *Applied Statistics Using SPSS, STATISTICA, MATLAB and R*, Springer, Porto 2007.
- [13] Kanji, Gopal K., *100 Statistical Tests* (3th edn), SAGE Publications, London 2006.
- [14] Landau S., Everitt B., *A Handbook of Statistical Analyses using SPSS*, Chapman & Hall/CRC, United States 2004.
- [15] Mehta, Cyrus, R., Patel, Nitin, R., *IBM SPSS Exact Tests*, Cytel Software Corporation and Harvard School of Public Health, Cambridge 2011.

- [16] Milosavljević, T., *Smernice Dobre kliničke prakse u kliničkom ispitivanju*, Službeni glasnik RS, broj 28/2008 od 13.03.2008.
- [17] Nedić, A., Živanović, O., Lisulov R., Vučković, N., Knežević, A., Soldatović-Stajić, B., Cvjetković-Bošnjak, M., Mitrović, D., Borišev, Lj., Drezgić-Vukić, S., Mišić-Pavkov, G., Dickov, A., Vasić, V., Silađi-Mladenović, Đ., Ivezić, A., Doroški, M., Ivanović-Kovačević, S., Marković, J. (2007). *Aktuelnosti u psihijatriji*. Medicinski Pregled, 60 (11-12), 637-641.
- [18] Ostojčić, D. (2012). *Prva epizoda shizofrenije – važnost ranog otkrivanja bolesti šta o tome trebaju znati socijalni radnici*. Ljetopis socijalnog rada, 19(1), 53-72.
- [19] Pallant, J., *SPSS priručnik za preživljavanje*, Mikro knjiga, Beograd 2009.
- [20] Spiegel, M., Stephens, L., *Statistics* (4th edn), Schaum's outline, United States 2007.
- [21] *Statistički kalendar Republike Srbije 2011*, Republički zavod za statistiku, Beograd.
- [22] Tabachnick, B. G., Fidell, L. S., *Using Multivariate Statistics* (5th edn). Pearson Education, Boston 2007.
- [23] Tenjović, L., *Statistika u psihologiji Priručnik* (Dopunjeno izdanje), Centar za primenjenu psihologiju, Beograd 2002.

Biografija

Snežana Todorović rođena 14. decembra 1989. godine u Bijeljini, u Republici Srpskoj. Godine 2008. završava srednju Ekonomsku školu u Bijeljini sa odličnim uspehom. Školovanje nastavlja iste godine na Matematičkom fakultetu, Univerziteta u Beogradu, na smeru statistika, aktuarska i finansijska matematika. Zvanje diplomirani matematičar stiče u julu 2012. godine. Iste godine upisuje Master akademske studije primenjene statistike, modul biomedicina na Univerzitetkom centru za primenjenu statistiku, Univerziteta u Novom Sadu. Zaključno sa septembarskim ispitnim rokom 2013. godine polože sve ispite predviđene planom i programom sa prosečnom ocenom 8.87.

