

UNIVERZITET U BEOGRADU

EKONOMSKI FAKULTET

Ivana B. Ivković

**NEPARAMETARSKE STATISTIČKE TEHNIKE
OCENJIVANJA REGRESIONIH
KOEFIČIJENATA I KOEFICIJENTA
VARIJACIJE U KORPORATIVNIM
FINANSIJAMA**

Doktorska disertacija

Beograd, 2021

UNIVERSITY OF BELGRADE

FACULTY OF ECONOMICS

Ivana B. Ivković

**NON-PARAMETRIC STATISTICAL
TECHNIQUES FOR ESTIMATION OF
REGRESSION COEFFICIENTS AND
COEFFICIENT OF VARIATION IN
CORPORATE FINANCE**

Doctoral Dissertation

Belgrade, 2021

Mentor:

dr Vesna Rajić,
Redovni profesor,
Univerzitet u Beogradu, Ekonomski fakultet

Članovi komisije:

dr Jelena Stanojević,
Docent,
Univerzitet u Beogradu, Ekonomski fakultet

dr Dragan Lončar,
Redovni profesor,
Univerzitet u Beogradu, Ekonomski fakultet

dr Irena Janković,
Vanredni profesor,
Univerzitet u Beogradu, Ekonomski fakultet

dr Milan Jovanović,
Docent,
Univerzitet u Beogradu, Matematički fakultet

Datum odbrane: _____

Doktorat posvećujem članovima moje porodice.

Zahvaljujem se mojoj mami na višegodišnjoj podršci i razumevanju.

Takođe, zahvaljujem se članovima komisije, a posebno mom mentoru na izdvojenom vremenu i svim korisnim savetima za unapređenje kvaliteta mog doktorata.

NEPARAMETARSKÉ STATISTIČKE TEHNIKE OCENJIVANJA REGRESIONIH KOEFICIJENATA I KOEFICIJENTA VARIJACIJE U KORPORATIVNIM FINANSIJAMA

APSTRAKT

Predmet istraživanja disertacije je intervalno ocenjivanje regresionih koeficijenata u prostom linearnom regresionom modelu i kvadratnom regresionom modelu ako slučajna greška nema normalnu raspodelu i intervalno ocenjivanje mera disperzije ako osnovni skup ne sledi normalnu raspodelu.

Ako su narušene polazne pretpostavke, proporcija simuliranih intervala za regresione koeficijente i mere disperzije može znatno odstupati od nominalnog nivoa pouzdanosti.

U radu su razvijene originalne metode za intervalno ocenjivanje regresionih koeficijenata u prostom linearnom i kvadratnom regresionom modelu zasnovane na Edgeworth-ovom razvoju raspodele t statistika koje se koriste u pomenutim modelima. Dalje, predložene su transformacije metoda koje se koriste za intervalno ocenjivanje koeficijenta varijacije. Reč je o transformaciji zasnovanoj na odsečenoj sredini i *bootstrap* transformaciji.

Validnost predloženih metoda proverena je kroz simulacije koristeći različite raspodele, kao i podatke u oblasti korporativnih finansija. Korporativne finansije obuhvataju praćenje efekata finansiranja kako bi se maksimizirala vrednost kompanije, kao i različite aspekte značajne za rast kompanije. Iz tog razloga, predmet razmatranja prilikom konstrukcije intervalnih ocena bili su podaci o indikatoru verovatnoće bankrotstva, količniku ukupnog duga, meri sistematskog rizika i dividendama.

Utvrđeno je da su, u većini razmatranih slučajeva, proporcije simuliranih intervala zasnovanih na predloženim metodama bliže nominalnom nivou pouzdanosti u poređenju sa proporcijama intervala proučavanih u literaturi. Na osnovu rezultata dobijenih u empirijskom delu, date su preporuke za intervalno ocenjivanje koje se mogu koristiti za donošenje pouzdanih zaključaka, pre svega, u oblasti korporativnih finansija.

Ključne reči: *proporcija simuliranih intervala poverenja, nominalni nivo pouzdanosti, regresioni koeficijent, prost linearni regresioni model, kvadratni regresioni model, koeficijent varijacije, Edgeworth-ov razvoj raspodele t statistike, odsečena sredina, korporativne finansije.*

Naučna oblast: Ekonomske nauke

Uža naučna oblast: Statistika

JEL klasifikacija: C13, C14, C15, C63

UDK broj: 519.233.5:658.14/.17(043.3)

NON-PARAMETRIC STATISTICAL TECHNIQUES FOR ESTIMATION OF REGRESSION COEFFICIENTS AND COEFFICIENT OF VARIATION IN CORPORATE FINANCE

ABSTRACT

The subject of the research of dissertation is the interval estimation of the regression coefficients in the simple linear regression model and quadratic regression model if an error term does not have normal distribution and interval estimation of the measures of dispersion if the population does not follow normal distribution.

If the initial assumptions are violated, the proportion of the simulated intervals for the regression coefficients and measures of dispersion can noticeably deviate from the nominal confidence level.

The original methods for the interval estimation of the regression coefficients in the simple linear and quadratic regression model, based on the Edgeworth's expansion of the distribution of the t statistics which are used in the mentioned models, were developed. Further, the transformations of the methods used for the interval estimation of the coefficient of variation were proposed. It is about the transformation based on the trimmed mean and the bootstrap transformation.

The validation of the proposed methods was checked through simulations using different distributions, as well as the data in the field of corporate finance. Corporate finance includes monitoring the effects of financing in order to maximize a company's value, as well as various aspects important to the company's growth. For that reason, the data on the bankruptcy probability indicator, total debt ratio, systematic risk measure and dividends were considered in order to construct the interval estimates.

It was found that, in most of the considered cases, the proportions of the simulated intervals based on the proposed methods are closer to the nominal confidence level compared with the proportions of the intervals studied so far in the literature. Based on the results obtained in the empirical part, the recommendations for the interval estimation, which can be used to draw the reliable conclusions, primarily, in the field of corporate finance, were given.

Key words: *proportion of the simulated confidence intervals, nominal confidence level, regression coefficient, simple linear regression model, quadratic regression model, coefficient of variation, Edgeworth's expansion of the distribution of the t statistic, trimmed mean, corporate finance.*

Scientific field: Economic sciences

Scientific subfield: Statistics

JEL classification: C13, C14, C15, C63

UDC number: 519.233.5:658.14/.17(043.3)

SADRŽAJ

UVOD	1
I METODE PONOVLJENIH UZORAKA.....	6
1.1 <i>Bootstrap</i> metoda.....	6
1.2 <i>Jackknife</i> metoda.....	7
1.3 Intervali poverenja zasnovani na ponovljenim uzorcima.....	9
1.4 Rangirano uzorkovanje.....	14
1.4.1 Balansirano rangirano uzorkovanje.....	14
1.4.2 Metode ponovljenih uzoraka u kombinaciji sa rangiranim uzorkovanjem	16
1.4.3 Parcijalno rangirano uzorkovanje	16
II INTERVALI POVERENJA ZA REGRESIONE KOEFICIJENTE	17
2.1 Asimetrične raspodele	17
2.1.1 Weibull-ova raspodela.....	17
2.1.2 Eksponecijalna raspodela.....	19
2.1.3 Gama raspodela.....	20
2.1.4 Log-normalna raspodela	20
2.1.5 <i>Skew</i> -normalna raspodela	21
2.1.6 <i>Skew</i> -Student-ova raspodela	23
2.1.7 <i>Skew</i> -Laplace-ova raspodela.....	26
2.2 Intervali poverenja za regresioni koeficijent β_1 u prostom linearnom regresionom modelu	27
2.2.1 Interval poverenja zasnovan na <i>t</i> -statistici.....	28
2.2.2 <i>Bootstrap-t</i> interval poverenja	28
2.2.3 Edgeworth-ov razvoj raspodele <i>t</i> -statistike	29
2.3 Intervali poverenja za regresione koeficijente β_1 i β_2 u kvadratnom regresionom modelu	37
2.3.1 Intervali poverenja zasnovani na <i>t</i> -statistikama.....	37
2.3.2 Intervali poverenja zasnovani na metodama ponovljenih uzoraka	39
2.3.3 Edgeworth-ov razvoj raspodele <i>t</i> -statistike	40
2.4 Analiza pouzdanosti intervala poverenja za regresione koeficijente	48
2.4.1 Analiza pouzdanosti intervala poverenja za regresioni koeficijent u prostom linearnom regresionom modelu	48
2.4.2 Analiza pouzdanosti intervala poverenja za regresione koeficijente u kvadratnom regresionom modelu.....	50
2.5 Analiza i primena u korporativnim finansijama	63
2.5.1 Korporativne finansije	63
2.5.2 Indikator verovatnoće bankrotstva.....	65
2.5.3 Količnik ukupnog duga	67
2.5.4 Analiza pouzdanosti intervala poverenja za regresioni koeficijent u prostom linearnom regresionom modelu na realnim podacima	68

2.5.5 Analiza pouzdanosti intervala poverenja za regresione koeficijente u kvadratnom regresionom modelu na realnim podacima	77
2.6 R kodovi za izabrane intervale	82
2.6.1 R kodovi za intervale poverenja za regresioni koeficijent u prostom linearnom regresionom modelu.....	82
2.6.2 R kodovi za intervale poverenja za regresione koeficijente u kvadratnom regresionom modelu.....	84

III INTERVALI POVERENJA ZA MERE DISPERZIJE..... 87

3.1 Intervali poverenja za varijansu	87
3.1.1 χ^2 interval poverenja.....	87
3.1.2 Bonett-ov interval poverenja	88
3.1.3 Steve-ov interval poverenja zasnovan na normalnoj aproksimaciji u slučaju velikih uzoraka	88
3.1.4 Interval poverenja zasnovan na logaritamskoj transformaciji	89
3.1.5 Interval poverenja sa korigovanim stepenima slobode	89
3.2 Intervali poverenja za koeficijent varijacije	89
3.2.1 Postojeći intervali poverenja.....	90
3.2.1.1 Interval poverenja zasnovan na t statistici	90
3.2.1.2 Miller-ov interval poverenja.....	91
3.2.1.3 Curto-ov i Pinto-ov interval poverenja	91
3.2.1.4 McKay-ev interval poverenja	91
3.2.2 Intervali poverenja zasnovani na metodama ponovljenih uzoraka	92
3.2.3 Intervali poverenja zasnovani na rangiranom uzorkovanju.....	93
3.2.4 Intervali poverenja zasnovani na metodama ponovljenih uzoraka u kombinaciji sa rangiranim uzorkovanjem	93
3.2.5 Intervali poverenja zasnovani na parcijalnom rangiranom uzorkovanju	94
3.2.6 Transformacije intervala poverenja.....	94
3.3 Analiza pouzdanosti intervala poverenja za mere disperzije.....	95
3.3.1 Analiza pouzdanosti intervala poverenja za varijansu u slučaju gama, log-normalne i eksponencijalne raspodele	95
3.3.2 Analiza pouzdanosti intervala poverenja za varijansu u slučaju <i>skew</i> -normalne, <i>skew</i> -Student-ove, <i>skew</i> -Laplace-ove i <i>skew</i> generalizovane t raspodele.....	98
3.3.3 Analiza pouzdanosti intervala poverenja za koeficijent varijacije.....	101
3.4 Analiza i primena u korporativnim finansijama	118
3.4.1 Mera sistematskog rizika	119
3.4.2 Dividenda.....	121
3.4.3 Analiza pouzdanosti intervala poverenja za varijansu	124
3.4.4 Analiza pouzdanosti intervala poverenja za koeficijent varijacije.....	127
3.5 R kodovi za izabrane intervale	131
3.5.1 R kodovi za intervale poverenja za varijansu	131
3.5.2 R kodovi za intervale poverenja za koeficijent varijacije.....	134

IV MARKOV CHAIN MONTE CARLO (MCMC) METODE KOD BAYES- OVSKOG OCENJIVANJA	142
4.1 Aposteriorna raspodela i Bayes-ovsko zaključivanje	142
4.1.1 Poređenje klasičnog i Bayes-ovskog pristupa	142
4.1.2 Osnovna ideja Bayes-ovskog zaključivanja	143
4.2 Osnovna ideja MCMC metoda	144
4.3 Vrste MCMC algoritama	146
4.3.1 Metropolis-Hastings-ov algoritam	147
4.3.2 Gibbs-ov algoritam.....	149
4.3.3 Dijagnostika lanca Markova	150
4.4 Primena MCMC algoritama kod Bayesovskog ocenjivanja.....	152
V POREĐENJE KLASIČNIH I BAYES-OVSKIH INTERVALA.....	154
5.1 Bayes-ovski intervali.....	154
5.2 Analiza klasičnih i Bayes-ovskih intervala za koeficijent varijacije log-normalne raspodele	155
5.2.1 Analiza lanaca Markova za koeficijent varijacije log-normalne raspodele.....	155
5.2.2 Analiza lanaca Markova za koeficijent varijacije neto dobitka	160
ZAKLJUČAK	164
LITERATURA	170
PRILOZI	182
SPISAK TABELA	198
SPISAK SLIKA.....	203
BIOGRAFIJA AUTORA.....	205

UVOD

Predmet istraživanja doktorske disertacije je intervalno ocenjivanje regresionih koeficijenata u prostom linearnom regresionom modelu i kvadratnom regresionom modelu ukoliko slučajna greška nije normalno raspodeljena, kao i intervalno ocenjivanje mera disperzije (varijansa i koeficijent varijacije) ukoliko je narušena pretpostavka o normalnosti osnovnog skupa.

Poznato je da se intervalno ocenjivanje regresionih koeficijenata u prostom linearnom regresionom modelu i kvadratnom regresionom modelu zasniva na t statistici ako su pretpostavke modela ispunjene. Takođe, poznato je da se intervalno ocenjivanje mera disperzije zasniva na χ^2 i t statistici ako nije narušena pretpostavka o normalnosti. U tim slučajevima će proporcije simuliranih intervala za regresione koeficijente odnosno mere disperzije koji sadrže nepoznati parametar biti najbliže nominalnom nivou pouzdanosti.

Moguće je definisati tri glavna cilja disertacije. Prvi cilj je analiza različitih metoda za intervalno ocenjivanje regresionih koeficijenata u prostom linearnom regresionom modelu i kvadratnom regresionom modelu, kao i mera disperzije u slučaju neispunjenosti određenih pretpostavki. Drugi cilj je razvoj originalnih metoda za intervalno ocenjivanje regresionih koeficijenata u prostom linearnom regresionom modelu i kvadratnom regresionom modelu koji su zasnovani na t statistici. Originalne metode zasnovane su na Edgeworth-ovom (Edgeworth, 1898) razvoju raspodele t statistike. Na osnovu razvoja predlažu se transformacije statistike koje se zatim koriste u kombinaciji sa *bootstrap-t* intervalom poverenja prilikom konstrukcije novih intervala poverenja za regresione koeficijente. Treći cilj odnosi se na predlaganje transformacija postojećih metoda koje se koriste za intervalno ocenjivanje relativne mere disperzije. Pored navedenog, koristeći originalne kodove koji su napisani u programskom jeziku R analiziraće se podaci u oblasti korporativnih finansija. Na osnovu empirijske analize biće predložene metode intervalnog ocenjivanja na osnovu kojih se mogu doneti pouzdani zaključci u oblasti korporativnih finansija.

Korporativne finansije obuhvataju planiranje, investiranje i praćenje efekata finansiranja kako bi se maksimizirala vrednost kompanije. Zato se smatra da se analizom podataka o indikatoru verovatnoće bankrotstva i količniku ukupnog duga ostvaruje značajan doprinos u oblasti korporativnih finansija. Pored navedenog, korporativne finansije obuhvataju različite aspekte koji su od značaja za rast kompanije. Reč je o načinima prikupljanja kapitala za investicije, izdavanju akcija i drugih hartija od vrednosti, itd. Analiza hartija od vrednosti ima veliki značaj u oblasti korporativnih finansija i iz tog razloga će u disertaciji akcenat biti na ispitivanju mere sistematskog rizika i dividende.

Uzimajući u obzir predmet i cilj istraživanja, u nastavku će biti definisana ključna istraživačka pitanja.

U doktorskoj disertaciji će se razmatrati sledeća istraživačka pitanja: koje metode će omogućiti da proporcije simuliranih intervala (za regresioni koeficijent) koji sadrže parametar (u daljem tekstu: proporcije simuliranih intervala) budu najbliže nominalnom nivou pouzdanosti ako slučajna greška u regresionom modelu ne sledi normalnu raspodelu, odnosno koje metode će omogućiti da proporcije simuliranih intervala (za mere disperzije) koji sadrže parametar (u daljem tekstu: proporcije simuliranih intervala) budu najbliže nominalnom nivou pouzdanosti ukoliko osnovni skup nije normalno raspodeljen? Pored glavnih istraživačkih pitanja, pažnja će biti posvećena i širinama klasičnih i Bayes-ovskih intervala.

Polazeći od definisanih istraživačkih pitanja, moguće je postaviti sledeće istraživačke hipoteze:

H₁: Proporcije simuliranih intervala poverenja za regresioni koeficijent u prostom linearnom regresionom modelu koji su zasnovani na novoj transformaciji t statistike, T_3 , za velike uzorke, bliže su nominalnom nivou pouzdanosti u odnosu na proporcije ostalih proučavanih intervala.

H₂: Proporcije simuliranih intervala poverenja za regresioni koeficijent β_2 u kvadratnom regresionom modelu koji su zasnovani na novoj transformaciji t statistike, T_2^ , za velike uzorke, bliže su nominalnom nivou pouzdanosti u odnosu na proporcije ostalih proučavanih intervala.*

H₃: Proporcije simuliranih intervala poverenja za varijansu osnovnog skupa, konstruisanih na osnovu rangiranog uzorkovanja, bliže su nominalnom nivou pouzdanosti od proporcija intervala poverenja konstruisanih na osnovu prostog slučajnog uzorkovanja.

H₄: Proporcije simuliranih intervala poverenja za koeficijent varijacije osnovnog skupa, zasnovanih na transformaciji sa odsečenom sredinom, bliže su nominalnom nivou pouzdanosti u odnosu na proporcije ostalih ispitivanih intervala.

H₅: Analiza pouzdanosti na realnim podacima, u oblasti korporativnih finansija, pokazuje da su proporcije predloženih intervala poverenja bliže nominalnom nivou pouzdanosti u odnosu na proporcije intervala koji su do sada korišćeni u literaturi.

U cilju postizanja adekvatnog odgovora na postavljena istraživačka pitanja, proučavaće se različite metode. Biće reči o metodama ponovljenih uzoraka, metodi rangiranog uzorkovanja, metodama ponovljenih uzoraka u kombinaciji sa rangiranim uzorkovanjem, metodi parcijalnog uzorkovanja i Markov Chain Monte Carlo (MCMC) metodama.

Metode ponovljenih uzoraka (engl. resampling methods) predstavljaju metode statističkog zaključivanja koje se koristi da se na osnovu informacija iz ponovljenih uzoraka (engl. resamples) izvedu zaključci o karakteristikama osnovnog skupa. Primena metoda ponovljenih uzoraka se preporučuje ako je raspodela osnovnog skupa nepoznata ili je izabrani uzorak male veličine. Ove metode se zasnivaju na generisanju uzoraka sa ponavljanjem iz originalnog uzorka. U svakom uzorku sa ponavljanjem se izračunava vrednost statistike i određuje se raspodela replika statistika. Metode ponovljenih uzoraka koje će se razmatrati u ovom radu su *bootstrap* i *jackknife* metoda.

Bootstrap metodu je razvio Efron krajem sedamdesetih godina XX veka. Prilikom primene ove metode polazi se od pretpostavke da je iz originalnog uzorka moguće generisati veliki broj uzoraka sa ponavljanjem koji su iste veličine kao originalni uzorak. Značaj ove metode zasniva se na izvođenju raspodela statistika koristeći originalni uzorak.

Jackknife metodu je predložio Quenouille sredinom XX veka. Ova metoda se zasniva na isključivanju po jednog elementa iz originalnog uzorka. Koristi se za određivanje standardnih grešaka ocena ili pristrasnosti ocena.

Intervali poverenja zasnovani na metodama ponovljenih uzoraka koriste se kao jedna od alternativa klasičnim intervalima poverenja u situacijama kada nisu ispunjene određene pretpostavke. U radu će se analizirati sledeći intervali poverenja: *bootstrap-t*, percentil, BC (skraćeno od engl. bias-corrected), BCa (skraćeno od engl. bias-corrected and accelerated) i

jackknife. BCa metoda zasniva se na percentilima *bootstrap* raspodele koji zavise od korekcije pristrasnosti i akceleracije. Percentil i BC metoda su specijalni slučajevi BCa metode kada su korekcija pristrasnosti i akceleracija jednaki nuli, odnosno kada je akceleracija jednaka nuli.

Rangirano uzorkovanje (engl. ranked set sampling) se prvi put pominje u radu McIntyre (1952). Predstavlja alternativni metod prikupljanja podataka koji je veoma ekonomičan. Primena rangiranog uzorkovanja u odnosu na prosto slučajno uzorkovanje preporučuje se ako je merenje na jedinicama skupo, zahteva mnogo vremena ili ga je komplikovano sprovesti. Pod ovom metodom podrazumeva se rangiranje slučajno izabranih jedinica posmatranja iz osnovnog skupa, a zatim i stvarna merenja na određenom broju izabranih jedinica. U disertaciji će se proučavati balansirano rangirano uzorkovanje sa savršenim procesom rangiranja, a koje se zasniva na izboru k prostih slučajnih uzoraka veličine k . Jedinice posmatranja se u svakom uzorku rangiraju prema promenljivoj od interesa. Zatim se iz prvog uzorka bira jedinica sa najmanjim rangom, iz drugog uzorka se bira jedinica sa drugim po redu rangom, i tako dalje. Postupak se ponavlja dok se iz k -tog uzorka ne izabere jedinica sa najvećim rangom.

Metoda parcijalnog rangiranog uzorkovanja (engl. partial ranked set sampling) predložena je u radu Haq et al. (2013). Predstavlja kombinaciju prostog slučajnog uzorkovanja i rangiranog uzorkovanja. Koristi se kada nije moguće ispitati sve jedinice koje su potrebne za formiranje balansirano rangiranog uzorka. Primenom ove metode se postiže veća preciznost ocena u poređenju sa prostim slučajnim uzorkovanjem, a zahteva da se ispita manji broj jedinica u odnosu na rangirano uzorkovanje.

Jedan od prvih Markov Chain Monte Carlo algoritama pod nazivom Metropolis-ov algoritam razvijen je sredinom XX veka (Metropolis et al., 1953). Zatim je Hastings (1970) izvršio uopštenje ovog algoritma koji je dobio naziv Metropolis-Hastings-ov algoritam. Markov Chain Monte Carlo metode predstavljaju metode generisanja slučajnih uzoraka iz ciljne raspodele. Ideja na kojoj su zasnovane MCMC metode podrazumeva da se za izabranu ciljnu raspodelu konstruiše lanac Markova čija ravnotežna raspodela dobro aproksimira ciljnu raspodelu. U radu će se razmatrati Metropolis-Hastings-ov algoritam, njegove varijante nezavisni Metropolis-Hastings-ov algoritam (engl. Independence Metropolis-Hastings algorithm) i Metropolis-Hastings-ov algoritam sa slučajnim hodom (engl. random walk Metropolis-Hastings), kao i Gibbs-ov algoritam.

Pored uvoda i zaključka, rad se sastoji od pet delova.

U prvom delu disertacije se razmatraju osnovne definicije i pojmovi koji se odnose na metode ponovljenih uzoraka: *bootstrap* metodu i *jackknife* metodu. Zatim se proučavaju intervali poverenja zasnovani na ponovljenim uzorcima: *bootstrap-t* interval, percentil interval, BC interval, BCa interval i *jackknife* interval. U nastavku se predstavlja koncept rangiranog uzorkovanja koje se posmatra kao alternativa prostom slučajnom uzorkovanju, opisuju se metode ponovljenih uzoraka u kombinaciji sa rangiranim uzorkovanjem i metoda parcijalnog rangiranog uzorkovanja.

Drugi deo disertacije odnosi se na intervale poverenja za regresione koeficijente u prostom linearnom regresionom modelu i kvadratnom regresionom modelu. Na početku ovog dela se ispituju asimetrične raspodele koje će biti osnova za sprovođenje simulacionih studija. Biće reči o sledećim raspedelama: Weibull-ovoj, eksponencijalnoj, gama, log-normalnoj, *skew*-normalnoj, *skew*-Student-ovoj, *skew* generalizovanoj t i *skew*-Laplace-ovoj raspodeli. Dalje, prikazuju se intervali poverenja za regresioni koeficijent u prostom linearnom regresionom modelu. U pitanju su postojeći interval poverenja zasnovan na t statistici, *bootstrap-t* interval i *bootstrap-t* intervali sa Hall-ovim (Hall, 1992a) transformacijama, T_1 i T_2 . Izvodi se Edgeworth-ov razvoj raspodele t

statistike i predlaže se nova transformacija t statistike, označena sa T_3 , na osnovu koje se konstruiše *bootstrap-t* interval sa transformacijom T_3 . U nastavku se predstavljaju intervali poverenja za regresione koeficijente u kvadratnom regresionom modelu: postojeći intervali poverenja zasnovani na t statistici, *bootstrap-t*, percentil, BC, BCa i *jackknife* interval. Izvodi se Edgeworth-ov razvoj raspodele t statistike za regresioni koeficijent β_2 i predlaže se nova transformacija t statistike, označena sa T_2^* , na osnovu koje se konstruiše *bootstrap-t* interval sa transformacijom T_2^* . U cilju ispitivanja proporcija simuliranih intervala sprovode se simulacione studije u kojima se polazi od pretpostavke da slučajna greška u regresionom modelu sledi neku od razmatranih asimetričnih raspodela. Na osnovu dobijenih rezultata vrši se poređenje proporcija simuliranih intervala za regresioni koeficijent u prostom linearnom regresionom modelu, a zatim i u kvadratnom regresionom modelu. Nakon sprovođenja simulacionih studija, analiza se vrši na realnim podacima u oblasti korporativnih finansija. Predmet istraživanja su podaci o indikatoru verovatnoće bankrotstva i količniku ukupnog duga, podaci koji se odnose na prosečnu cenu leta i udaljenost između gradova, kao i zbirni finansijski indikatori: neto dobitak, troškovi sirovina, broj zaposlenih i ukupna imovina. Na kraju drugog dela navode se kodovi iz programskog jezika R za izabrane intervale poverenja.

U trećem delu disertacije razmatraju se intervali poverenja za mere disperzije: varijansu i koeficijent varijacije. Na početku ovog dela navode se intervali poverenja za varijansu: interval zasnovan na χ^2 statistici, Bonett-ov interval (Bonett, 2006), Steve-ov interval (Steve, 1990) zasnovan na normalnoj aproksimaciji u slučaju velikih uzoraka, interval zasnovan na logaritamskoj transformaciji i interval sa korigovanim stepenima slobode. Dalje, predstavljaju se intervali poverenja za koeficijent varijacije. Prvo se razmatraju postojeći intervali poverenja: interval zasnovan na t statistici, Miller-ov interval (Miller, 1991), Curto-ov & Pinto-ov interval (Curto & Pinto, 2009) i McKay-ev interval (McKay, 1932), a zatim intervali zasnovani na metodama ponovljenih uzoraka: *bootstrap-t*, percentil, BC, BCa i *jackknife*. U nastavku se analiziraju Miller-ov, Curto-ov & Pinto-ov i McKay-ev interval poverenja zasnovani na rangiranom uzorkovanju i parcijalnom rangiranom uzorkovanju, kao i percentil, BC i BCa intervali u kombinaciji sa rangiranim uzorkovanjem. Predlažu se dve transformacije Miller-ovog, Curto-ovog & Pinto-ovog i McKay-evog intervala poverenja. Prva transformacija podrazumeva upotrebu odsečene sredine umesto aritmetičke sredine prilikom izračunavanja varijanse uzorka. Odsečena sredina se određuje nakon isključivanja opservacija sa ekstremnim vrednostima i zbog toga se smatra boljim izborom ukoliko se analiziraju podaci koji potiču iz asimetrične raspodele. Koriste se Bonett-ova odsečena sredina (Bonett, 2006) i odsečena sredina koja eliminiše 10% vrednosti. Druga transformacija se zasniva na primeni percentila koji se koriste prilikom konstrukcije *bootstrap-t* intervala. U cilju analize proporcija simuliranih intervala za razmatrane mere disperzije sprovode se simulacione studije u kojima se polazi od pretpostavke da osnovni skup sledi asimetričnu raspodelu, a zatim se ispitivanje vrši na realnim podacima u oblasti korporativnih finansija. Predmet istraživanja su podaci o meri sistematskog rizika, dividendama i štetama nastalim usled požara u Danskoj. Na kraju ovog dela prikazuju se kodovi iz programskog jezika R za izabrane intervale poverenja.

Na početku sledećeg, četvrtog dela disertacije razmatraju se razlike između klasičnog i Bayes-ovskog pristupa u statistici. Definišu se apriorna i aposteriorna raspodela i predstavlja se osnovna ideja Bayes-ovskog zaključivanja. Kako bi se omogućilo jednostavnije razumevanja ideje Markov Chain Monte Carlo metoda, objašnjavaju se pojmovi poput slučajnog procesa, procesa Markova i lanca Markova. U nastavku se izlažu svojstva lanaca Markova i osnovna ideja MCMC metoda. Detaljno se opisuju MCMC algoritmi poput Metropolis-Hastings-ovog algoritma, njegovih varijanti nezavisnog Metropolis-Hastings-ovog algoritma i Metropolis-Hastings-ovog algoritma sa slučajnim hodom, kao i Gibbs-ovog algoritma. Dalje se pažnja posvećuje dijagnostici lanca

Markova analizirajući konvergenciju lanca, prisustvo autokorelacije u lancu, efektivnu veličinu uzorka (engl. effective sample size), Gelman-Rubinov pokazatelj i stopu prihvatanja (engl. acceptance rate). Na kraju četvrtog dela se na hipotetičkom primeru objašnjava konstrukcija lanaca Markova koji za ravnotežnu raspodelu imaju ciljnu aposteriornu raspodelu.

Poslednji, peti deo odnosi se na klasične i Bayes-ovske intervale za nepoznati parametar. Na početku ovog dela predstavljaju se Bayes-ovski intervale. U pitanju su interval sa jednakim repovima (engl. equal-tailed interval) i interval sa najvećom aposteriornom gustinom (engl. highest posterior density interval). Dalje, konstruišu se dvodimenzionalni lanci Markova za sredinu i koeficijent varijacije log-normalne raspodele primenom Metropolis-Hastings-ovog algoritma. Analiziraju se širine intervala sa jednakim repovima i intervala sa najvećom aposteriornom gustinom za koeficijent varijacije log-normalne raspodele, a zatim se porede širine klasičnih intervala poverenja za koeficijent varijacije log-normalne raspodele sa širinama Bayes-ovskih intervala. Analiza se sprovodi na simuliranim podacima i na podacima o neto dobitku.

Na kraju rada dat je uvid u zaključke istraživanja, doprinose doktorske disertacije, kao i ograničenja istraživanja, a zatim su razmatrani pravci daljeg istraživanja.

Originalni naučni rezultati su: odeljak 2.2.3, odeljak 2.3.3, odeljak 2.4.1, odeljak 2.4.2, odeljak 2.5.4, odeljak 2.5.5, odeljak 3.2.5, odeljak 3.3.3 i odeljak 3.4.4.

I METODE PONOVLJENIH UZORAKA

Metode ponovljenih uzoraka (engl. resampling methods) su metode statističkog zaključivanja koje se koriste za izvođenje zaključaka o karakteristikama osnovnog skupa koristeći informacije iz ponovljenih uzoraka (engl. resamples). Ove metode najviše su proučavali Efron i Tibshirani (videti Efron, 1979, 1981b; Efron & Tibshirani, 1986, 1993), a pre njih i Bernard (1963), Marriott (1979), Simon (1969) i drugi.

Klasične metode se takođe koriste za izvođenje zaključaka o karakteristikama osnovnog skupa, ali postoji razlika u načinu na koji se pomenuti cilj postiže. Prednost metoda ponovljenih uzoraka u odnosu na klasične metode je u tome što zahtevaju minimalnu ispunjenost pretpostavki, pri čemu se njihovom primenom ponekad može postići i veća preciznost. Ukoliko je raspodela osnovnog skupa nepoznata ili je izabrani uzorak male veličine, preporučuje se upotreba metoda ponovljenih uzoraka. Ove metode se baziraju na podacima iz uzorka umesto na podacima osnovnog skupa što podrazumeva da se novi uzorci biraju iz originalnog uzorka. Jedna od najčešće korišćenih metoda ponovljenih uzoraka u praksi je *bootstrap* metoda.

Najpre će biti predstavljene dve metode ponovljenih uzoraka: *bootstrap* metoda i *jackknife* metoda, kao i intervali poverenja zasnovani na njima. Nakon toga će biti opisane: procedura rangiranog uzorkovanja, metode ponovljenih uzoraka u kombinaciji sa rangiranim uzorkovanjem i parcijalno rangirano uzorkovanje.

1.1 *Bootstrap* metoda

Bootstrap metodu je predložio Efron krajem sedamdesetih godina prošlog veka. Iako je on razvio ovu metodu, o njoj se pisalo i ranije u radovima Hope (1968), Hartigan (1971, 1975), Maritz i Jarrit (1978) i drugih.

Kod *bootstrap* metode, polazi se od pretpostavke da je iz originalnog uzorka moguće generisati veliki broj uzoraka sa ponavljanjem, koji su iste veličine kao i originalni uzorak. Ovaj originalni uzorak će reprezentovati osnovni skup iz koga je izabran. Na osnovu realizovanih vrednosti statistike iz ponovljenih uzoraka, može se dobiti informacija o raspodeli statistike koja se proučava. Ideja koja stoji iza *bootstrap* metode je zapravo sledeća: da bi se izvršila aproksimacija raspodela statistika, podaci iz originalnog uzorka se koriste kao “surogat osnovnog skupa” (Singh & Xie, n.d.).

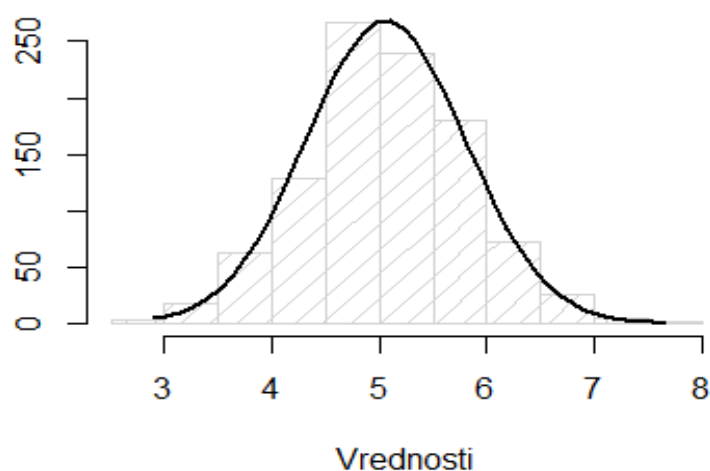
Neka je $\mathbf{X} = (X_1, X_2, \dots, X_n)$ slučajan uzorak generisan iz neke raspodele sa nepoznatim parametrom θ . Neka je statistika $\hat{\theta}$ ocena parametra θ , dobijena na osnovu slučajnog uzorka. *Bootstrap* uzorak $\mathbf{X}^* = (X_1^*, X_2^*, \dots, X_n^*)$ predstavlja slučajan uzorak veličine n , generisan sa ponavljanjem iz originalnog uzorka $\mathbf{X} = (X_1, X_2, \dots, X_n)$. Na primer, jedan *bootstrap* uzorak može biti formiran na sledeći način: $X_1^* = X_5$, $X_2^* = X_5$, $X_3^* = X_1$, $X_4^* = X_n$, $X_5^* = X_{12}$, ... , $X_n^* = X_6$. Primećuje se da se *bootstrap* uzorak sastoji od podataka originalnog uzorka, od kojih su neki zastupljeni jednom, neki dva puta ili više, dok neki uopšte ne moraju biti zastupljeni u izabranom *bootstrap* uzorku.

Sušтина *bootstrap* metode je izvođenje raspodela statistika na osnovu jednog (originalnog) uzorka. Ova procedura može se sprovesti na osnovu nekoliko sledećih koraka (videti Efron & Tibshirani, 1993, str. 45-46):

1. Iz originalnog uzorka $\mathbf{X} = (X_1, X_2, \dots, X_n)$ generiše se B nezavisnih *bootstrap* uzoraka $\mathbf{X}_1^*, \mathbf{X}_2^*, \dots, \mathbf{X}_B^*$, veličine n .
2. U svakom od B *bootstrap* uzoraka izračunava se vrednost statistike $\hat{\theta}$, koja se označava sa $\hat{\theta}^*$ (na primer, ako $\hat{\theta}$ predstavlja sredinu $\bar{X} = \sum_{i=1}^n X_i / n$ dobijenu na osnovu originalnog uzorka, tada $\hat{\theta}^*$ predstavlja sredinu $\bar{X}^* = \sum_{i=1}^n X_i^* / n$ dobijenu na osnovu *bootstrap* uzorka).
3. Na osnovu *bootstrap* vrednosti $\hat{\theta}_1^*, \hat{\theta}_2^*, \dots, \hat{\theta}_B^*$ određuje se *bootstrap* raspodela statistike $\hat{\theta}^*$.
4. *Bootstrap* raspodela statistike $\hat{\theta}^*$ reprezentuje raspodelu statistike $\hat{\theta}$, koja je zasnovana na velikom broju uzoraka.

Ideja *bootstrap* metode je objašnjena na sledećem ilustrativnom primeru. Neka je iz normalne raspodele sa sredinom 5 i varijansom 2 izabran originalni uzorak od deset elemenata: $\mathbf{x} = (6,299675; 7,643615; 7,811522; 5,905474; 4,932619; 5,164273; 7,378087; 3,480086; 3,030977; 7,635812)$.

Na osnovu izabranog uzorka izračunata je realizovana vrednost sredine uzorka koja iznosi 5,050145. Iz originalnog uzorka izabrano je 1000 uzoraka sa ponavljanjem (*bootstrap* uzoraka) veličine deset. U svakom od 1000 *bootstrap* uzoraka izračunata je sredina uzorka, tj. vrednost statistike \bar{X}^* . Na Slici 1.1 nalazi se histogram *bootstrap* raspodele statistike \bar{X}^* .



Slika 1.1: Histogram *bootstrap* raspodele statistike \bar{X}^*
Izvor: Izlaz iz programskog jezika R

1.2 Jackknife metoda

Jackknife metodu je predložio Quenouille (1949) godine sa ciljem da oceni pristrasnost ocene parametra. Tukey (1958) je dao naziv ovoj metodi i proširio je njenu upotrebu na izvođenje

varijansi ocena. O ovoj metodi je diskutovano u radovima Efron i Stein (1981), Efron (1982, 1992), Wu (1986), Shao i Tu (1995), Houdré (1997), Knight (2000), Abdi (2010), Chernick i LaBudde (2011), Zhang et al. (2016) i drugim.

Postoji veliki broj situacija u praksi kada je vrlo teško ili čak nemoguće izvesti standardne greške ocena. Na primer, ocenjivanje varijanse neke statistike korišćenjem uobičajene procedure može biti nemoguće ukoliko se radi o složenoj funkciji koju je nemoguće integraliti. U takvim situacijama se može koristiti *jackknife* metoda da bi se izvela standardna greška ocene ili pristrasnost ocene.

Neka je $\mathbf{X} = (X_1, X_2, \dots, X_n)$ slučajan uzorak generisan iz neke raspodele sa nepoznatim parametrom θ . Neka je statistika $\hat{\theta}$ ocena parametra θ , dobijena na osnovu slučajnog uzorka. *Jackknife* uzorci $\mathbf{X}_{(i)} = (X_1, X_2, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$, $i = 1, 2, \dots, n$ predstavljaju uzorke veličine $n-1$, izabrane iz originalnog uzorka $\mathbf{X} = (X_1, X_2, \dots, X_n)$. Prilikom izbora i -tog *jackknife* uzorka izostavlja se element X_i . Na primer, ako originalni uzorak ima pet elemenata $\mathbf{X} = (X_1, X_2, \dots, X_5)$, moguće je formirati sledećih pet *jackknife* uzoraka: $\mathbf{X}_{(1)} = (X_2, X_3, X_4, X_5)$, $\mathbf{X}_{(2)} = (X_1, X_3, X_4, X_5)$, $\mathbf{X}_{(3)} = (X_1, X_2, X_4, X_5)$, $\mathbf{X}_{(4)} = (X_1, X_2, X_3, X_5)$ i $\mathbf{X}_{(5)} = (X_1, X_2, X_3, X_4)$.

Jackknife procedura za ocenu standardne greške statistika se može predstaviti sledećim koracima (Efron & Tibshirani, 1993, str. 141-142).

1. Iz originalnog uzorka $\mathbf{X} = (X_1, X_2, \dots, X_n)$ bira se n *jackknife* uzoraka.
2. U svakom *jackknife* uzorku se ocenjuje parametar θ , odnosno računa se vrednost statistike $\hat{\theta}$, koja se označava sa $\theta_{(i)}$ (na primer, ako $\hat{\theta}$ predstavlja sredinu $\bar{X} = \sum_{i=1}^n X_i / n$ dobijenu na osnovu originalnog uzorka, tada $\theta_{(i)}$ predstavlja sredinu $\bar{X}_{(i)}$ dobijenu na osnovu *jackknife* uzorka).
3. *Jackknife* ocena standardne greške statistike $\hat{\theta}$ je definisana na sledeći način:

$$\left[\frac{n-1}{n} \sum_{i=1}^n \left(\theta_{(i)} - \sum_{i=1}^n \theta_{(i)} / n \right)^2 \right]^{1/2}. \quad (1.1)$$

Ideja *jackknife* metode je objašnjena na sledećem ilustrativnom primeru. Neka je iz normalne raspodele sa sredinom 5 i varijansom 2 izabran originalni uzorak od deset elemenata: $\mathbf{x} = (0,00194453; 6,88875643; 4,48961872; 5,60864398; 5,66182122; 6,38626244; 5,14790592; 4,96000652; 5,01502161; 4,48917942)$.

Na osnovu izabranog uzorka izračunata je realizovana vrednost sredine uzorka koja iznosi 4,864916. Iz originalnog uzorka izabrano je deset *jackknife* uzoraka veličine devet prema opisanom postupku. U svakom *jackknife* uzorku izračunata je sredina uzorka i dobijene sledeće vrednosti: 5,405246; 4,640045; 4,906616; 4,782280; 4,776371; 4,695878; 4,833473; 4,854350; 4,848238; 4,906665. Na osnovu izračunatih sredina, koristeći formulu (1.1), *jackknife* ocena standardne greške statistike \bar{X} iznosi 0,5929286.

Upotreba *jackknife* metode podrazumeva da se biraju uzorci bez ponavljanja, za razliku od *bootstrap* metode gde se biraju uzorci sa ponavljanjem. Ova metoda je manje računarski zahtevna u odnosu na *bootstrap* metodu, jer koristi manje informacija (prednost je posebno izražena ukoliko je veličina uzorka mala, jer se vrlo brzo i jednostavno računa n *jackknife* vrednosti razmatrane statistike). Primena *jackknife* metode, slično kao i *bootstrap* metode, ne zahteva da bude izvedena formula za ocenu standardne greške. Međutim, primenom *jackknife* metode asimptotska vrednost ocene standardne greške je veća od one dobijene primenom *bootstrap* metode.

1.3 Intervali poverenja zasnovani na ponovljenim uzorcima

Postoji više intervala poverenja koji su zasnovani na metodama ponovljenih uzoraka. Ovde će biti predstavljeni i analizirani sledeći intervali: *bootstrap-t*, percentil, BCa, BC i *jackknife* interval poverenja.

Bootstrap-t interval poverenja

Bootstrap-t metoda zasniva se na percentilima raspodele standardizovane statistike. Ovu metodu je predložio Efron (1981a), a ona je proučavana i u radovima sledećih autora: Bickel i Freedman (1981), Athreya (1986), DiCiccio i Romano (1988), Hinkley (1988), Mamman (1992), Efron i Tibshirani (1993), Barbe i Bertail (1995), Davison i Hinkley (1997), Carpenter (1999), Rajić (2007), Chernick (2008), Wang et al. (2010), Drikvandi et al. (2011) i drugih.

Neka je $\mathbf{X} = (X_1, X_2, \dots, X_n)$ slučajan uzorak generisan iz neke raspodele sa nepoznatim parametrom θ . Ocena parametra θ dobijena na osnovu izabranog uzorka obeležava se sa $\hat{\theta}$, a ocena njene standardne greške sa $S_{\hat{\theta}}$. *Bootstrap-t* procedura se sprovodi na sledeći način (Efron & Tibshirani, 1993, str. 160):

1. Iz originalnog uzorka $\mathbf{X} = (X_1, X_2, \dots, X_n)$ bira se B nezavisnih *bootstrap* uzoraka $\mathbf{X}_1^*, \mathbf{X}_2^*, \dots, \mathbf{X}_B^*$, veličine n .
2. U svakom od B *bootstrap* uzoraka određuje se $\hat{\theta}_b^*$, $b = 1, \dots, B$, koja predstavlja vrednost statistike $\hat{\theta}$ na osnovu *bootstrap* uzorka \mathbf{X}_b^* .
3. Izračunava se vrednost statistike $T_b^* = \frac{\hat{\theta}_b^* - \hat{\theta}}{S_{\hat{\theta}_b^*}}$, $b \in \{1, \dots, B\}$, gde $S_{\hat{\theta}_b^*}$ predstavlja ocenu standardne greške ocene $\hat{\theta}^*$ u *bootstrap* uzorku \mathbf{X}_b^* .
4. *Bootstrap-t* interval poverenja predstavljen je sledećom formulom:

$$I_{\theta} = \left[\hat{\theta} - \hat{t}_{1-\alpha/2} S_{\hat{\theta}}, \hat{\theta} - \hat{t}_{\alpha/2} S_{\hat{\theta}} \right], \quad (1.2)$$

gde se vrednosti $\hat{t}_{1-\alpha/2}$ i $\hat{t}_{\alpha/2}$ određuju na sledeći način:¹

$$\begin{aligned}\#\{T_b^* \leq \hat{t}_{1-\alpha/2}\} / B &= 1 - \alpha / 2, \\ \#\{T_b^* \leq \hat{t}_{\alpha/2}\} / B &= \alpha / 2.\end{aligned}$$

Percentil interval poverenja

Percentil metoda zasniva se na percentilima *bootstrap* raspodele. Prvi put se pominje u radu Efrona iz 1979. godine, a razmatrana je i u radovima Buckland (1985), Hall (1988a), DiCiccio i Efron (1996), Carpenter i Bithell (2000), Good (2006), Chernick i LaBudde (2011) i drugih.

Neka je θ nepoznati parametar. Ocena parametra θ dobijena na osnovu slučajnog uzorka, $\mathbf{X} = (X_1, X_2, \dots, X_n)$, označena je sa $\hat{\theta}$. Ova procedura se sprovodi na sledeći način (Efron & Tibshirani, 1993, str. 170-171):

1. Iz originalnog uzorka $\mathbf{X} = (X_1, X_2, \dots, X_n)$ bira se B nezavisnih *bootstrap* uzoraka $\mathbf{X}_1^*, \mathbf{X}_2^*, \dots, \mathbf{X}_B^*$.
2. U svakom od B *bootstrap* uzoraka određuje se $\hat{\theta}_b^*$, $b=1, \dots, B$, koja predstavlja vrednost statistike $\hat{\theta}$ u *bootstrap* uzorku \mathbf{X}_b^* .
3. Vrednosti $\hat{\theta}_b^*$, $b=1, \dots, B$, se rangiraju u neopadajućem poretku.

Percentil interval poverenja pouzdanosti $(1-\alpha)100\%$ za parametar θ definisan je na sledeći način:

$$I_\theta = [\hat{\theta}_{\alpha/2}^*, \hat{\theta}_{1-\alpha/2}^*], \quad (1.3)$$

gde su sa $\hat{\theta}_{\alpha/2}^*$ i $\hat{\theta}_{1-\alpha/2}^*$ označeni $\alpha/2$ i $1-\alpha/2$ percentili raspodele vrednosti $\hat{\theta}_b^*$. Oni se mogu izračunati koristeći sledeće formule:

$$\begin{aligned}\#\{\hat{\theta}_b^* \leq \hat{\theta}_{1-\alpha/2}^*\} / B &= 1 - \alpha / 2, \\ \#\{\hat{\theta}_b^* \leq \hat{\theta}_{\alpha/2}^*\} / B &= \alpha / 2.\end{aligned}$$

Ova metoda je vrlo jednostavna za upotrebu i iz tog razloga je široko primenljiva. Prednost percentil metode u odnosu na *bootstrap-t* metodu je što ima osobinu očuvanja transformacija².

¹ Simbol $\#\{T_b^* \leq \hat{t}_{1-\alpha/2}\}$ označava broj vrednosti T_b^* ($b=1, \dots, B$) manjih od $\hat{t}_{1-\alpha/2}$.

Međutim, za razliku od *bootstrap-t* metode koja je drugog reda preciznosti³, percentil metoda je prvog reda preciznosti. Nedostatak ove metode je u tome što je greška pokrivanja⁴ (engl. coverage error) često velika ukoliko raspodela statistike $\hat{\theta}$ nije približno simetrična (Carpenter & Bithell, 2000).

BCa interval poverenja

BCa metoda zasniva se na percentilima *bootstrap* raspodele, koji zavise od dve veličine: korekcije pristrasnosti (engl. bias-correction) i akceleracije (engl. acceleration). Razmatrali su je Efron (1987), Hall (1988b), DiCiccio i Romano (1988), Efron i Tibshirani (1993), Carpenter i Bithell (2000), Chernick i LaBudde (2011) i drugi.

Neka je $\hat{\theta}$ ocena nepoznatog parametra θ izvedena na osnovu slučajnog uzorka $\mathbf{X} = (X_1, X_2, \dots, X_n)$. Iz originalnog uzorka generiše se B *bootstrap* uzoraka. U svakom od B *bootstrap* uzoraka računa se vrednost statistike $\hat{\theta}$ i rangiraju se dobijene vrednosti $\hat{\theta}_b^*$, $b=1, \dots, B$, u neopadajućem poretku. BCa interval poverenja pouzdanosti $(1-\alpha)100\%$ za parametar θ je oblika (Efron & Tibshirani, 1993, str. 185):

$$I_{\theta} = \left[\hat{\theta}_{\alpha_1}^*, \hat{\theta}_{\alpha_2}^* \right], \quad (1.4)$$

gde se vrednosti α_1 i α_2 računaju na osnovu sledećih formula:

$$\alpha_1 = \Phi \left[\hat{z}_0 + \frac{\hat{z}_0 + Z_{\alpha/2}}{1 - \hat{a}(\hat{z}_0 + Z_{\alpha/2})} \right] \quad (1.5)$$

i

$$\alpha_2 = \Phi \left[\hat{z}_0 + \frac{\hat{z}_0 + Z_{1-\alpha/2}}{1 - \hat{a}(\hat{z}_0 + Z_{1-\alpha/2})} \right]. \quad (1.6)$$

² Neka se primenom određene metode ponovljenih uzoraka dobija sledeći interval poverenja $(\hat{\theta}_d, \hat{\theta}_g)$ za nepoznati parametar θ . Neka je $h(\theta)$ monotona transformacija parametra θ . Ako se interval za $h(\theta)$ može definisati kao $(h(\hat{\theta}_d), h(\hat{\theta}_g))$, metoda ima osobinu očuvanja transformacije.

³ Neka su $\hat{\theta}_d$ i $\hat{\theta}_g$ donja i gornja granica intervala poverenja za nepoznati parametar θ . Interval poverenja je prvog reda preciznosti (engl. first-order accurate) ako važi $P(\theta < \hat{\theta}_d) = \alpha + O(n^{-1/2})$ i $P(\theta > \hat{\theta}_g) = \alpha + O(n^{-1/2})$. Za interval se kaže da je drugog reda preciznosti (engl. second-order accurate) ako važi $P(\theta < \hat{\theta}_d) = \alpha + O(n^{-1})$ i $P(\theta > \hat{\theta}_g) = \alpha + O(n^{-1})$.

⁴ Greška pokrivanja intervala poverenja predstavlja razliku između stvarnog i nominalnog nivoa pouzdanosti.

U formulama (1.5) i (1.6), $\Phi(\cdot)$ je funkcija raspodele standardizovane normalne promenljive, \hat{z}_0 je korekcija pristrasnosti, \hat{a} je akceleracija, dok $Z_{\alpha/2}$ i $Z_{1-\alpha/2}$ predstavljaju $\alpha/2$ i $1-\alpha/2$ percentile standardizovane normalne raspodele. Pristrasnost predstavlja razliku između očekivane vrednosti ocene parametra i stvarne vrednosti parametra. Korekcija pristrasnosti određuje se na osnovu proporcije vrednosti $\hat{\theta}_b^*$ ($b=1, \dots, B$) koje su manje od ocene $\hat{\theta}$, po sledećoj formuli:

$$\hat{z}_0 = \Phi^{-1} \left(\frac{\#\{\hat{\theta}_b^* < \hat{\theta}\}}{B} \right), \quad (1.7)$$

gde je $\Phi^{-1}(\cdot)$ inverzna funkcija funkcije raspodele standardizovane normalne promenljive. Akceleracija meri brzinu promene standardne greške ocene parametra u odnosu na stvarnu vrednost parametra i računa se prema sledećoj formuli:

$$\hat{a} = \frac{\sum_{i=1}^n \left(\frac{\sum_{i=1}^n \theta_{(i)}}{n} - \theta_{(i)} \right)^3}{6 \left[\sum_{i=1}^n \left(\frac{\sum_{i=1}^n \theta_{(i)}}{n} - \theta_{(i)} \right)^2 \right]^{3/2}}, \quad (1.8)$$

gde $\theta_{(i)}$ predstavlja vrednost statistike $\hat{\theta}$ dobijenu na osnovu *jackknife* uzorka (više o tome pročitati u Efron & Tibshirani, 1993, str. 184-186).

Izračunavanje \hat{z}_0 i \hat{a} se može objasniti na sledećem ilustrativnom primeru. Neka je $\alpha = 0,05$. Tada vrednosti $Z_{\alpha/2}$ i $Z_{1-\alpha/2}$ iznose, redom, -1,96 i 1,96. Ako je proporcija vrednosti $\hat{\theta}_b^*$ ($b=1, \dots, B$) koje su manje od ocene $\hat{\theta}$ jednaka npr. 0,65, tada je prema formuli (1.7), korekcija pristrasnosti jednaka $\hat{z}_0 = \Phi^{-1}(0,65) = 0,3853205$. Prilikom formiranja intervala poverenja npr. za sredinu skupa u formuli (1.8) simbol $\theta_{(i)}$ bi trebalo zameniti simbolom $\bar{X}_{(i)}$. Ako bi se iskoristile vrednosti statistike na osnovu *jackknife* uzoraka iz odeljka 1.2 (one iznose 5,405246; 4,640045; 4,906616; 4,782280; 4,776371; 4,695878; 4,833473; 4,854350; 4,848238; 4,906665) dobila bi se vrednost akceleracije -0,09584. Sledi, $\alpha_1 = \Phi(2,2399) = 0,9875$ i $\alpha_2 = \Phi(2,3002) = 0,9893$.

BCa metoda ima osobinu očuvanja transformacija, drugog je reda preciznosti i greška pokrivanja je manja nego kod percentil⁵ i BC metode. Nedostatak je u tome što računanje akceleracije može biti veoma komplikovano pri rešavanju kompleksnih problema (Carpenter & Bithell, 2000).

⁵ Percentil metoda je specijalni slučaj BCa metode koja se dobija kada su korekcija pristrasnosti i akceleracija jednaki nuli (Ghosh et al., 2017, str. 1695).

BC interval poverenja

BC metoda je specijalni slučaj BCa metode kada je akceleracija jednaka nuli (videti Ghosh et al., 2017, str. 1695). Interval poverenja pouzdanosti $(1-\alpha)100\%$ za parametar θ , zasnovan na BC metodi, predstavljen je sledećom formulom:

$$I_{\theta} = [\hat{\theta}_{\alpha_1}^*, \hat{\theta}_{\alpha_2}^*]. \quad (1.9)$$

Vrednosti α_1 i α_2 , za percentile $\hat{\theta}_{\alpha_1}^*$ i $\hat{\theta}_{\alpha_2}^*$, izračunate su primenom sledećih formula:

$$\alpha_1 = \Phi[2\hat{z}_0 + Z_{\alpha/2}] \text{ i} \\ \alpha_2 = \Phi[2\hat{z}_0 + Z_{1-\alpha/2}].$$

BC metoda takođe poseduje osobinu očuvanja transformacija. Uzimajući u obzir nedostatak percentil metode, zamišljeno je da BC metoda predstavlja poboljšanje percentil metode. Ukoliko je raspodela statistike $\hat{\theta}^*$ simetrična u odnosu na $\hat{\theta}$, BC i percentil interval poverenja su jednaki. Greška pokrivanja je manja nego kod percentil metode (videti Carpenter & Bithell, 2000).

Jackknife interval poverenja

Neka je θ nepoznati parametar. Ocena parametra θ dobijena na osnovu slučajnog uzorka, $\mathbf{X} = (X_1, X_2, \dots, X_n)$ je statistika $\hat{\theta}$. Procedura je sledeća:

1. Iz originalnog uzorka $\mathbf{X} = (X_1, X_2, \dots, X_n)$ bira se n jedinstvenih *jackknife* uzoraka (videti odeljak 1.2).
2. U svakom *jackknife* uzorku se računa vrednost statistike $\hat{\theta}$ (u oznaci $\theta_{(i)}$).
3. Određuju se pseudo-vrednosti: $\tilde{\theta}_i = n\hat{\theta} - (n-1)\hat{\theta}_{(i)}$ u cilju korigovanja pristrasnosti ocene.
4. Računa se sredina pseudo-vrednosti: $\tilde{\theta} = \sum_{i=1}^n \tilde{\theta}_i / n$, koja predstavlja *jackknife* ocenu parametra θ .
5. Na osnovu standardne devijacije pseudo-vrednosti: $S_{jack} = \left\{ \sum_{i=1}^n (\tilde{\theta}_i - \tilde{\theta})^2 / [(n-1) \cdot n] \right\}^{1/2}$ određuje se *jackknife* standardna greška ocene parametra.

Jackknife interval poverenja pouzdanosti $(1-\alpha)100\%$ za parametar θ je oblika:

$$I_{\theta} = \left[\tilde{\theta} - t_{1-\alpha/2, n-1} S_{jack}, \tilde{\theta} + t_{1-\alpha/2, n-1} S_{jack} \right], \quad (1.10)$$

gde $t_{1-\alpha/2}$ predstavlja $1-\alpha/2$ percentil t raspodele sa $n-1$ stepeni slobode.

1.4 Rangirano uzorkovanje

Procedura rangiranog uzorkovanja (u oznaci RSS, skraćeno od engl. ranked set sampling) predstavlja alternativni metod prikupljanja podataka i vrlo je ekonomična procedura uzorkovanja. Kada je merenje na jedinicama teško izvodljivo, vremenski zahtevno ili skupo, a rangiranje može da se sprovede na jednostavan način, bez stvarnih merenja na izabranim jedinicama, preporučuje se primena rangiranog uzorkovanja u odnosu na prosto slučajno uzorkovanje (u oznaci SRS, skraćeno od engl. simple random sampling). U navedenim situacijama, ova procedura se koristi da bi se obezbedile ocene sa boljim osobinama (Wolfe, 2004, str. 636).

Rangirano uzorkovanje predstavlja proceduru koja se sastoji iz dve etape. U prvoj etapi jedinice se identifikuju i rangiraju, dok se u drugoj etapi vrše merenja na određenom broju rangiranih jedinica.

Rangirano uzorkovanje je prvi put objasnio McIntyre (1952) na primeru ocenjivanja sredine osnovnog skupa. Koristeći rangirano uzorkovanje, Stokes (1980) je zaključio da se dobija ocena varijanse koja je asimptotski nepristrasna a da je asimptotska efikasnost ocene bolja u poređenju sa ocenom koja je zasnovana na istom broju jedinica iz prostog slučajnog uzorka. Terpstra i Wang (2008) su proučavali nekoliko metoda za konstruisanje intervala poverenja za proporciju osnovnog skupa zasnovanih na rangiranom uzorkovanju, dok je Wolfe (2012) napisao pregledni članak o uticaju rangiranog uzorkovanja na statističko zaključivanje. Više o proceduri rangiranog uzorkovanja i njenoj primeni može se naći u Dell i Clutter (1972), Samawi (1999), Husby et al. (2005), Chen (2007), Albatineh et al. (2014), Zamanzade i Vock (2015), Zhang et al. (2016), Albatineh et al. (2017), Zamanzade i Mahdizadeh (2017), Ozturk (2018), i drugih.

1.4.1 Balansirano rangirano uzorkovanje

Balansirano rangirano uzorkovanje podrazumeva da se iz osnovnog skupa bira k^2 jedinica na slučajan način (Ganeslingam & Ganesh, 2006; Chen, 2007; Wolfe, 2012). Procedura se sastoji u sledećem (Ghosh et al., 2017; Ivković & Rajić, 2018):

1. Od k^2 izabranih jedinica formira se k uzoraka sa po k jedinica;
2. U svakom prostom slučajnom uzorku vrši se rangiranje jedinica u zavisnosti od istraživačkog zadatka;
3. Iz prvog uzorka bira se jedinica sa najmanjim rangom $X_{(1)}$. Zatim se iz drugog uzorka bira jedinica sa sledećim po redu rangom $X_{(2)}$. Procedura se nastavlja dok se ne izabere jedinica sa najvećim rangom iz k -tog uzorka, $X_{(k)}$. Kao rezultat, dobijen je uzorak koji se sastoji od k jedinica $X_{(1)}, X_{(2)}, \dots, X_{(k)}$;
4. U cilju generisanja balansiranog rangiranog uzorka veličine $n = mk$, prethodno opisani postupak (odnosno ciklus) se ponavlja m puta (videti Tabelu 1.1). Balansirani rangirani uzorak sa m ciklusa i k jedinica u svakom ciklusu se zapisuje na sledeći način:
$$\{X_{(j)i} : j = 1, 2, \dots, k; i = 1, 2, \dots, m\}.$$

Tabela 1.1: Balansirani rangirani uzorak sa m ciklusa i k jedinica u svakom ciklusu

Ciklus 1	$X_{(1)1}$	$X_{(2)1}$...	$X_{(k)1}$
Ciklus 2	$X_{(1)2}$	$X_{(2)2}$...	$X_{(k)2}$
...
Ciklus m	$X_{(1)m}$	$X_{(2)m}$...	$X_{(k)m}$

Izvor: Preuzeto iz Wolfe (2012)

Proces formiranja rangiranog uzorka može se objasniti na sledećem primeru. Neka je predmet istraživanja verovatnoća bankrotstva kompanije, koja se meri indikatorom verovatnoće bankrotstva. U tu svrhu, slučajnim putem se bira 25 kompanija. Indikator verovatnoće bankrotstva je složen pokazatelj za čije izračunavanje je potrebno poznavati nekoliko različitih pokazatelja. Iz tog razloga, vremenski je zahtevno izračunati ovaj pokazatelj za svaku od 25 izabranih kompanija. Međutim, na osnovu racia pokrića kamata zaradom, koji je dosta jednostavniji pokazatelj, može se dati gruba procena verovatnoće bankrotstva kompanije. Kako bi se formirao rangirani uzorak, na osnovu izabranih kompanija formira se pet uzoraka veličine pet. U prvom uzorku od pet kompanija vrši se rangiranje kompanija, tako da se kompaniji za koju se smatra da ima najmanju verovatnoću da bankrotira (analizirajući racio pokrića kamata zaradom) dodeljuje rang 1, kompaniji za koju se smatra da ima veću verovatnoću za bankrot dodeljuje se rang 2, dok se kompaniji za koju se smatra da ima najveću verovatnoću za bankrot dodeljuje rang 5. Iz ovog uzorka se bira kompanija za koju se smatra da ima najmanju verovatnoću za bankrot (odnosno kompanija kojoj je dodeljen rang 1) i za ovu kompaniju se određuje indikator verovatnoće bankrotstva. U drugom uzorku od pet kompanija vrši se rangiranje po istom principu. Iz drugog uzorka bira se kompanija kojoj je dodeljen rang 2 i za ovu kompaniju se računa indikator verovatnoće bankrotstva. Postupak se ponavlja sve dok se iz poslednjeg uzorka ne izabere kompanija za koju se smatra da ima najveću verovatnoću za bankrot (kompanija kojoj je dodeljen rang 5). Za ovu kompaniju se određuje indikator verovatnoće bankrotstva. Sprovedenjem ove procedure završen je jedan ciklus koji se sastoji od pet jedinica, odnosno odabranih kompanija, za koje će se izračunati indikator verovatnoće bankrotstva. Ove kompanije čine rangirani uzorak. Ukoliko je potrebno generisati rangirani uzorak veličine 15, ciklus se ponavlja tri puta (u svakom od tri ciklusa biće izabrano pet kompanija za koje će se određivati indikator verovatnoće bankrotstva, što rezultira rangiranim uzorkom veličine 15).

Formule (1.11) i (1.12), zasnovane na rangiranom uzorkovanju, primenjuju su prilikom računanja ocene sredine i ocene varijanse osnovnog skupa (Stokes, 1980, str. 36-37):

$$\bar{X}_{RSS} = \frac{1}{km} \sum_{j=1}^k \sum_{i=1}^m X_{(j)i}, \quad (1.11)$$

$$S^2_{RSS} = \frac{1}{km-1} \sum_{j=1}^k \sum_{i=1}^m \left(X_{(j)i} - \bar{X}_{RSS} \right)^2. \quad (1.12)$$

Moguće je navesti nekoliko razlika između balansiranog rangiranog uzorka veličine n i prostog slučajnog uzorka iste veličine. Prost slučajan uzorak se bira tako da su jedinice u uzorku identično raspodeljene, pri čemu mogu, a ne moraju biti nezavisne (u zavisnosti od toga da li je uzorkovanje bez ili sa ponavljanjem). Kod prostog slučajnog uzorka svaka jedinica se može smatrati reprezentativnom, što podrazumeva da dobro odražava osobine osnovnog skupa. Kod balansiranog rangiranog uzorka jedinice su nezavisne, ali nisu identično raspodeljene, zbog čega jedinice koje čine ovaj uzorak ne reprezentuju dobro osnovni skup. Međutim, individualne poređane

statistike dobijene rangiranim uzorkovanjem veoma dobro reprezentuju različite delove osnovnog skupa. Zbog ove osobine, ocene dobijene na osnovu balansirano rangiranog uzorka su efikasnije od ocena dobijenih prostim slučajnim uzorkovanjem.

1.4.2 Metode ponovljenih uzoraka u kombinaciji sa rangiranim uzorkovanjem

Modarres et al. (2006), Amiri et al. (2013) i Ghosh et al. (2017) proučavali su proceduru *bootstrap* rangirano uzorkovanje po redovima (engl. bootstrap ranked set sampling by row). Ova procedura će u odeljku 3.3.3 biti primenjena prilikom konstrukcije percentil, BCa i BC intervala poverenja i sprovodi se na sledeći način (Ivković & Rajić, 2019, str. 6):

1. Iz k nezavisnih slučajnih uzoraka veličine m $(X_{(1)1}, \dots, X_{(1)m}), \dots, (X_{(k)1}, \dots, X_{(k)m})$ generiše se k *bootstrap* uzoraka, svaki veličine m .
2. *Bootstrap* rangirani uzorak predstavljen je na sledeći način:

$$X_{RSS}^* = (X_{(j)i}^*, j=1, \dots, k; i=1, \dots, m).$$

1.4.3 Parcijalno rangirano uzorkovanje

Parcijalno rangirano uzorkovanje (engl. partial ranked set sampling) se sprovodi na sledeći način. Iz skupa se bira A jedinica koristeći prosto slučajno uzorkovanje i B jedinica koristeći rangirano uzorkovanje, na osnovu kojih se dobija parcijalni rangirani uzorak veličine $A+B$. Uočava se da je parcijalno rangirano uzorkovanje kombinacija prostog slučajnog uzorkovanja i rangiranog uzorkovanja. Predstavljena metoda se preporučuje kada nije moguće ispitati sve jedinice koje su potrebne za formiranje balansirano rangiranog uzorka ili kada je ispitivanje po jedinici veoma skupo. Procedura parcijalnog rangiranog uzorkovanja zahteva ispitivanje manjeg broja jedinica u poređenju sa rangiranim uzorkovanjem i obezbeđuje preciznije ocene u odnosu na prosto slučajno uzorkovanje (Haq et al., 2013, str. 207).

U cilju generisanja parcijalnog rangiranog uzorka veličine k sprovodi se procedura opisana u nastavku (Haq et al., 2013; Ivković & Rajić, 2019):

1. Definiše se r , tako da $r = \alpha k, 0 \leq \alpha < 0,5$ ⁶ i bira se $2r$ prostih slučajnih uzoraka veličine jedan iz populacije.
2. Bira se $k - 2r$ uzoraka veličine k i u svakom uzorku se jedinice rangiraju.
3. Iz i -tog uzorka, za $i = r+1, \dots, k-r$, bira se i -ta rangirana jedinica. Ovim se završava jedan ciklus i dobija se parcijalni rangirani uzorak veličine k .

Ukoliko je za dalju analizu potreban parcijalni rangirani uzorak veličine $n = mk$, ciklus se ponavlja m puta. Kada je $r=0$, parcijalno rangirano uzorkovanje je identično rangiranom uzorkovanju.

⁶ α je ponder koji se koristi da se izračuna koeficijent na osnovu kog se određuje koji deo parcijalnog rangiranog uzorka će činiti prost slučajni uzorak.

II INTERVALI POVERENJA ZA REGRESIJE KOEFICIJENTE

U ovom delu biće razmatrani intervali poverenja za regresione koeficijente u prostom linearnom regresionom modelu i kvadratnom regresionom modelu. Pitanje koje se javlja je sledeće: kada će proporcije simuliranih intervala (za regresioni koeficijent) koji sadrže parametar biti najbliže nominalnom nivou pouzdanosti kada slučajna greška u regresionom modelu ne sledi normalnu raspodelu⁷, već asimetričnu raspodelu. Proporcija simuliranih intervala koji sadrže parametar (u daljem tekstu: proporcija simuliranih intervala) predstavlja proporciju intervala poverenja koji sadrže vrednost parametra osnovnog skupa (određuje se kao n_1/n , gde je n_1 broj intervala poverenja koji sadrže vrednost parametra osnovnog skupa, a n je ukupan broj formiranih intervala poverenja).

Najpre će biti razmatrane asimetrične raspodele koje će biti korišćene prilikom sprovođenja simulacionih studija, a zatim i intervali poverenja za regresione koeficijente u prostom linearnom i kvadratnom regresionom modelu. U nastavku će biti prikazani rezultati simulacionih studija, kao i rezultati dobijeni na realnim podacima. Na kraju će biti navedeni kodovi iz programskog jezika *R* za izabrane intervale.

2.1 Asimetrične raspodele

Podaci se, u praksi, često mogu opisati nekom asimetričnom raspodelom. Neke od asimetričnih raspodela koje će se razmatrati u ovom odeljku su: Weibull-ova, eksponencijalna, gama, log-normalna, *skew*-normalna, *skew*-Student-ova, *skew*-generalizovana *t* i *skew*-Laplace-ova raspodela.

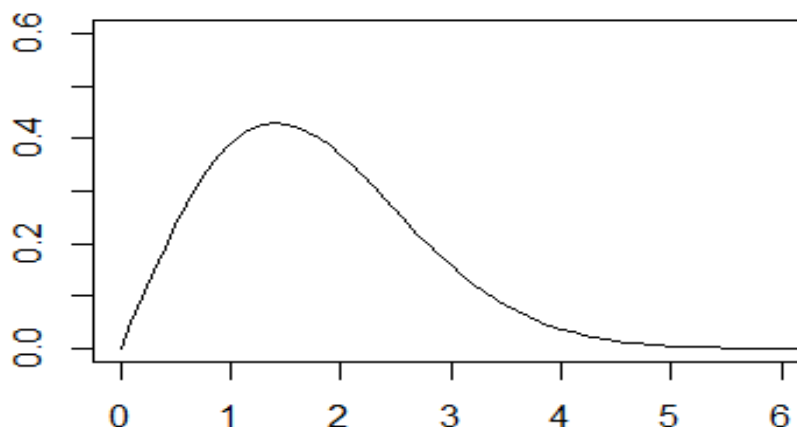
2.1.1 Weibull-ova raspodela

Neprekidna slučajna promenljiva X ima Weibull-ovu raspodelu sa parametrom oblika $a > 0$ i parametrom skaliranja $b > 0$ (u oznaci $X \sim W(a, b)$), ako je njena funkcija gustine oblika (Đorić et al., 2007, str. 71):

$$f(x) = \frac{a}{b} \left(\frac{x}{b}\right)^{a-1} \exp\left\{-\left(\frac{x}{b}\right)^a\right\}, x > 0. \quad (2.1)$$

Na Slici 2.1 prikazana je funkcija gustine Weibull-ove raspodele za proizvoljno izabrane vrednosti parametara $a = 2$ i $b = 2$.

⁷ Pretpostavlja se da su ostale pretpostavke regresionog modela ispunjene.



Slika 2.1: Funkcija gustine Weibull-ove raspodele

Izvor: Izlaz iz programskog jezika R

Na osnovu parametra oblika utvrđuje se oblik Weibull-ove “krive”, a parametar skaliranja ukazuje na raspršenost Weibull-ove raspodele. Matematičko očekivanje, varijansa, koeficijent asimetrije⁸ i koeficijent spljoštenosti⁹ Weibull-ove raspodele su predstavljeni sledećim formulama (Đorić et al., 2007, str. 72):

$$E(X) = \mu = b\Gamma\left(1 + \frac{1}{a}\right),$$

$$\text{Var}(X) = \sigma^2 = b^2 \left[\Gamma\left(1 + \frac{2}{a}\right) - \Gamma^2\left(1 + \frac{1}{a}\right) \right],$$

$$\gamma_3(X) = \frac{\Gamma\left(1 + \frac{3}{a}\right)b^3 - 3\mu\sigma^2 - \mu^3}{\sigma^3},$$

⁸ Koeficijent asimetrije i koeficijent spljoštenosti predstavljaju mere oblika raspodele. Raspodela neprekidne slučajne promenljive X je simetrična ako postoji tačka c tako da važi: $g(c-x) = g(c+x)$, za svaki realan broj x . Ukoliko navedeno nije ispunjeno, raspodela je asimetrična. Koeficijent asimetrije $\gamma_3(X)$ definiše se na sledeći način:

$\gamma_3(X) = \frac{\mu_3}{\mu_2^{3/2}}$, gde su μ_2 i μ_3 centralni momenti drugog i trećeg reda, respektivno. Za $\gamma_3(X) = 0$ raspodela je simetrična, za $\gamma_3(X) > 0$ raspodela je asimetrična udesno, dok je za $\gamma_3(X) < 0$ raspodela asimetrična ulevo (Đorić et al., 2007, str. 8).

⁹ Spljoštenost raspodele govori o brzini konvergencije krajeva gustine raspodele ka nuli, kao i o koncentraciji oko srednje vrednosti. Koeficijent spljoštenosti $\gamma_4(X)$ određuje se sledećom formulom: $\gamma_4(X) = \frac{\mu_4}{\mu_2^2}$, gde je μ_4 centralni

momenat četvrtog reda. Za $\gamma_4(X) = 3$ raspodela je normalne spljoštenosti, za $\gamma_4(X) > 3$ raspodela je manje spljoštenosti od normalne, dok je za $\gamma_4(X) < 3$ raspodela veće spljoštenosti od normalne (Đorić et al., 2007, str. 8).

$$\gamma_4(X) = \frac{b^4 \Gamma\left(1 + \frac{4}{a}\right) - 3\sigma^4 - 4\gamma_3\sigma^3\mu - 6\sigma^2\mu^2 - \mu^4}{\sigma^4},$$

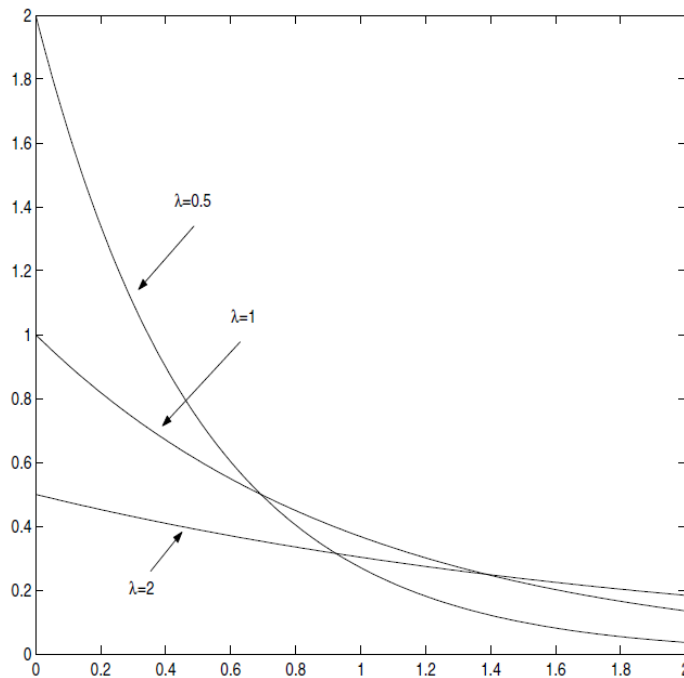
gde $\Gamma(\cdot)$ predstavlja gama funkciju¹⁰.

2.1.2 Eksponecijalna raspodela

Neprekidna slučajna promenljiva X ima eksponencijalnu raspodelu sa parametrom skaliranja $\lambda > 0$ (u oznaci $X \sim E(\lambda)$), ako je njena funkcija gustine (Đorić et al., 2007, str. 121):

$$f(x) = \frac{1}{\lambda} \exp\left(-\frac{x}{\lambda}\right), x \geq 0. \quad (2.2)$$

Na Slici 2.2 predstavljena je funkcija gustine eksponencijalne raspodele za proizvoljno izabrane vrednosti parametra λ .



Slika 2.2: Funkcija gustine eksponencijalne raspodele

Izvor: Preuzeto iz Đorić et al. (2007)

Parametar skaliranja λ opisuje centar raspodele. Matematičko očekivanje, varijansa, koeficijent asimetrije i koeficijent spljoštenosti eksponencijalne raspodele su predstavljeni u sledećem obliku (Đorić et al., 2007, str. 122):

$$E(X) = \lambda, \quad \text{Var}(X) = \lambda^2, \quad \gamma_3(X) = 2, \quad \gamma_4(X) = 6.$$

¹⁰ Gama funkcija se definiše na sledeći način: $\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt, x > 0$.

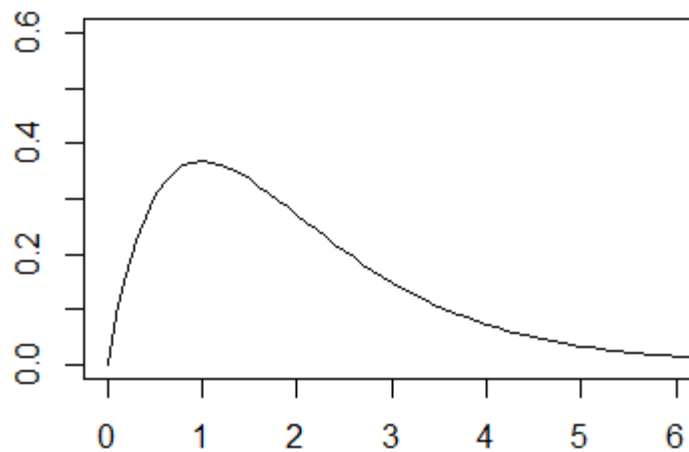
2.1.3 Gama raspodela

Neprekidna slučajna promenljiva X ima gama raspodelu sa parametrom oblika $a > 0$ i parametrom skaliranja $b > 0$ (u oznaci $X \sim G(a, b)$), ako je njena funkcija gustine (Đorić et al., 2007, str. 96):

$$f(x) = \frac{1}{b\Gamma(a)} \left(\frac{x}{b}\right)^{a-1} \exp\left(-\frac{x}{b}\right), x \geq 0, \quad (2.3)$$

gde $\Gamma(\cdot)$ predstavlja gama funkciju.

Na Slici 2.3 prikazana je funkcija gustine gama raspodele za proizvoljno izabrane vrednosti parametara $a = 2$ i $b = 1$.



Slika 2.3: Funkcija gustine gama raspodele

Izvor: Izlaz iz programskog jezika R

Matematičko očekivanje, varijansa, koeficijent asimetrije i koeficijent spljoštenosti gama raspodele su (Đorić et al., 2007, str. 98):

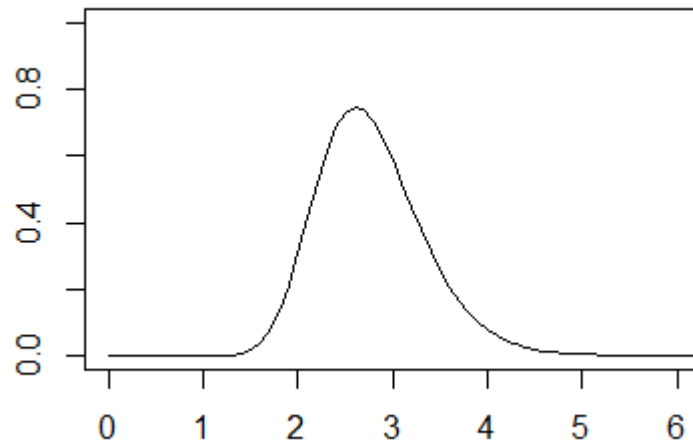
$$E(X) = ab, \quad \text{Var}(X) = ab^2, \quad \gamma_3(X) = \frac{2}{\sqrt{a}}, \quad \gamma_4(X) = \frac{6}{a}.$$

2.1.4 Log-normalna raspodela

Neprekidna slučajna promenljiva X ima log-normalnu raspodelu sa parametrom oblika $\mu \in \mathbf{R}$ i parametrom skaliranja $\sigma > 0$ (u oznaci $X \sim LN(\mu, \sigma^2)$), ako je njena funkcija gustine (Đorić et al., 2007, str. 240):

$$f(x) = \frac{1}{x} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2\right\}, x > 0. \quad (2.4)$$

Na Slici 2.4 predstavljena je funkcija gustine log-normalne raspodele za proizvoljno izabrane vrednosti parametara $\mu = 1$ i $\sigma = 0,2$.



Slika 2.4: Funkcija gustine log-normalne raspodele

Izvor: Izlaz iz programskog jezika R

Matematičko očekivanje, varijansa, koeficijent asimetrije i koeficijent spljoštenosti log-normalne raspodele su prikazani sledećim formulama (Đorić et al., 2007, str. 241):

$$E(X) = \exp\left(\mu + \frac{1}{2}\sigma^2\right),$$

$$\text{Var}(X) = \exp(2\mu + \sigma^2)(\exp(\sigma^2) - 1),$$

$$\gamma_3(X) = (\omega - 1)^{1/2}(\omega + 2),$$

$$\gamma_4(X) = \omega^4 + 2\omega^3 + 3\omega^2 - 6,$$

gde je $\omega = \exp(\sigma^2)$.

2.1.5 Skew-normalna raspodela

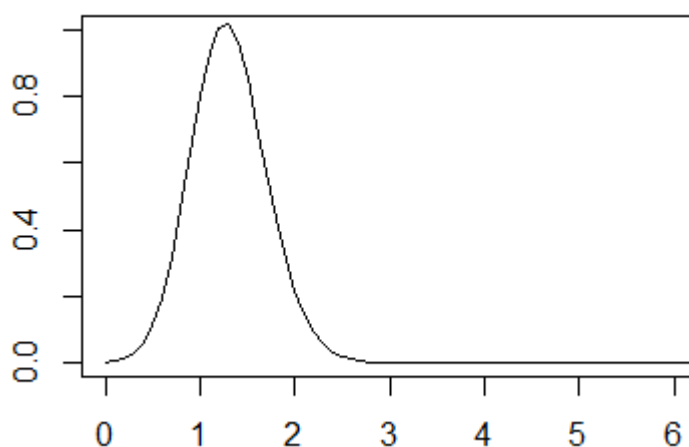
Skew-normalna raspodela izvedena je iz normalne raspodele. Neprekidna slučajna promenljiva X ima skew-normalnu raspodelu sa parametrom oblika $\alpha \in \mathbf{R}$ (u oznaci $X \sim SN(0,1,\alpha)$), ako je njena funkcija gustine:

$$f(x) = 2\phi(x)\Phi(\alpha x), x \in \mathbf{R}, \quad (2.5)$$

gde je $\phi(x)$ funkcija gustine standardizovane normalne promenljive, a $\Phi(x)$ je funkcija raspodele standardizovane normalne promenljive (Azzalini, 1985; Eling, 2012). Kada je $\alpha = 0$ skew-normalna raspodela postaje standardizovana normalna raspodela $N(0,1)$.

Ako $X \sim SN(0,1,\alpha)$ i ako $Y = \beta + \gamma X$, tada slučajna promenljiva Y ima skew-normalnu raspodelu sa parametrima β , γ i α , gde je $\beta \in \mathbf{R}$ parametar lokacije, $\gamma > 0$ predstavlja parametar skaliranja i $\alpha \in \mathbf{R}$ je parametar oblika, što može biti zapisano u formi $Y \sim SN(\beta, \gamma^2, \alpha)$. Parametar lokacije predstavlja apcisu neke mere lokacije, dok parametri skaliranja i oblika imaju već pomenuto značenje.

Na Slici 2.5 predstavljena je funkcija gustine skew-normalne raspodele za $\beta = 1$, $\gamma = 0,5$ i $\alpha = 1,2$.



Slika 2.5: Funkcija gustine skew-normalne raspodele
Izvor: Izlaz iz programskog jezika R

Matematičko očekivanje, varijansa, koeficijent asimetrije i koeficijent spljoštenosti skew-normalne raspodele su (Eling, 2012, str. 240):

$$E(Y) = \beta + \gamma \sqrt{2/\pi} \delta,$$

$$\text{Var}(Y) = \gamma^2 (1 - 2\delta^2 / \pi),$$

$$\gamma_3(Y) = (4 - \pi) / 2 (\delta(2/\pi)^{1/2})^3 / (1 - 2\delta^2 / \pi)^{3/2},$$

$$\gamma_4(Y) = 2(\pi - 3)(\delta(2/\pi)^{1/2})^4 / (1 - 2\delta^2 / \pi)^2,$$

gde je $\delta = \alpha / \sqrt{1 + \alpha^2}$.

2.1.6 Skew-Student-ova raspodela

Skew-Student-ova raspodela izvedena je iz Student-ove raspodele. Neka slučajna promenljiva Z ima skew-normalnu raspodelu sa parametrom oblika $\alpha \in \mathbf{R}$ (označeno sa $Z \sim SN(0,1,\alpha)$) i neka slučajna promenljiva W ima hi-kvadrat raspodelu sa ν stepeni slobode (u oznaci $W \sim \chi^2(\nu)$). Tada slučajna promenljiva X , definisana na sledeći način:

$$X = \frac{Z}{\sqrt{W/\nu}},$$

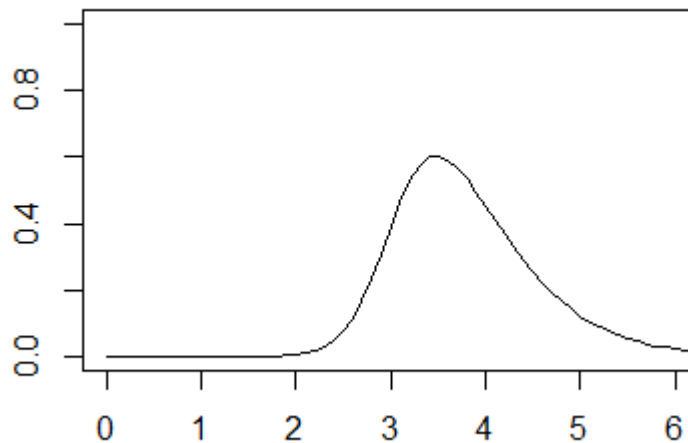
ima skew-Student-ovu raspodelu sa parametrom $\alpha \in \mathbf{R}$ (u oznaci $X \sim ST(0,1,\alpha)$), ako je njena funkcija gustine:

$$f(x) = 2t(x;\nu)T(\alpha x;\nu), x \in \mathbf{R}, \quad (2.6)$$

gde je $t(x;\nu)$ funkcija gustine T promenljive koja ima Studentovu raspodelu, a $T(\alpha x;\nu)$ je funkcija raspodele T promenljive (Azzalini & Capitanio, 2003, str. 101).

Linearna transformacija $Y = \beta + \gamma X$ ima skew-Student-ovu raspodelu sa parametrom lokacije $\beta \in \mathbf{R}$, parametrom skaliranja $\gamma > 0$ i parametrom oblika $\alpha \in \mathbf{R}$, u oznaci $Y \sim ST(\beta, \gamma^2, \alpha)$.

Na Slici 2.6 prikazana je funkcija gustine skew-Student-ove raspodele za proizvoljno izabrane vrednosti parametara $\beta = 3$, $\gamma = 1$, $\alpha = 2,5$ i proizvoljan broj stepeni slobode 10.



Slika 2.6: Funkcija gustine skew-Student-ove raspodele

Izvor: Izlaz iz programskog jezika R

Parametri lokacije, skaliranja i oblika imaju slično značenje kao i u slučaju skew-normalne raspodele. Matematičko očekivanje, varijansa i koeficijent asimetrije skew-Student-ove raspodele, za $\beta = 0$, predstavljeni su u sledećem obliku (Azzalini & Capitanio, 2003, str. 103-104):

$$E(Y) = \gamma\eta\delta, \text{ za } \nu > 1,$$

$$\text{Var}(Y) = \gamma^2 \left(\frac{\nu}{\nu-2} - \eta^2 \delta^2 \right), \text{ za } \eta = \sqrt{\frac{\nu}{\pi}} \frac{\Gamma\left(\frac{1}{2}(\nu-1)\right)}{\Gamma\left(\frac{1}{2}\nu\right)},$$

$$\gamma_3(Y) = \eta\delta \left\{ \frac{\nu(3-\delta^2)}{\nu-3} - \frac{3\nu}{\nu-2} + 2\eta^2\delta^2 \right\} \left(\frac{\nu}{\nu-2} - \eta^2\delta^2 \right)^{-3/2}, \text{ za } \nu > 3.$$

Koeficijent spljoštenosti definisan je za $\nu > 4$ i jednak je:

$$\gamma_4(Y) = \left\{ \frac{3\nu^2}{(\nu-2)(\nu-4)} - \frac{4\eta^2\delta^2\nu(3-\delta^2)}{\nu-3} + \frac{6\eta^2\delta^2\nu}{\nu-2} - 3\eta^4\delta^4 \right\} \left(\frac{\nu}{\nu-2} - \eta^2\delta^2 \right)^{-2} - 3.$$

Moguće je izvršiti i generalizaciju ove raspodele. Neprekidna slučajna promenljiva X ima *skew* generalizovanu t raspodelu sa parametrom lokacije $\mu \in \mathbf{R}$, parametrom skaliranja $\sigma > 0$, parametrom asimetrije $-1 \leq \lambda \leq 1$ i dva parametra $p > 0$ i $q > 0$ koji kontrolišu spljoštenost¹¹, ako je njena funkcija gustine (Davis, 2015):

$$f(x) = \frac{p}{2\nu\sigma q^{1/p} B\left(\frac{1}{p}, q\right) \left(\frac{|x-\mu+m|^p}{q(\nu\sigma)^p (\lambda \text{sign}(x-\mu+m)+1)^p} + 1 \right)^{\frac{1}{p+q}}}, \quad x \in \mathbf{R}, \quad (2.7)$$

gde $B(\cdot)$ predstavlja beta funkciju¹² i

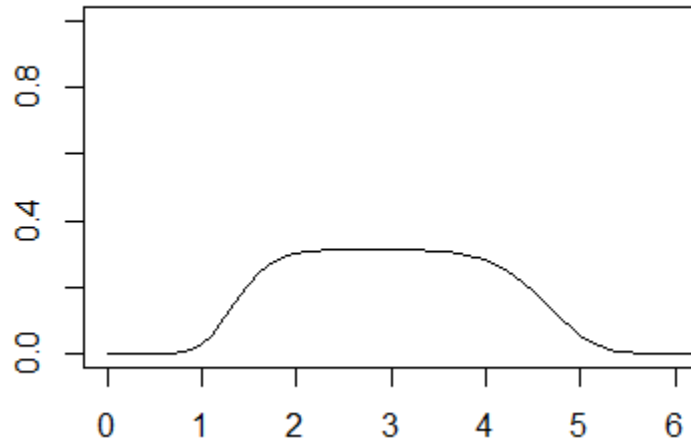
$$m = \frac{2\nu\sigma\lambda q^{1/p} B\left(\frac{2}{p}, q - \frac{1}{p}\right)}{B\left(\frac{1}{p}, q\right)},$$

$$\nu = \frac{q^{-1/p}}{\sqrt{(3\lambda^2+1) \frac{B\left(\frac{3}{p}, q - \frac{2}{p}\right)}{B\left(\frac{1}{p}, q\right)} - 4\lambda^2 \frac{B\left(\frac{2}{p}, q - \frac{1}{p}\right)^2}{B\left(\frac{1}{p}, q\right)^2}}}$$

¹¹ Za manje vrednosti p i q veće su vrednosti koeficijenta spljoštenosti raspodele. Takođe, dozvoljeno je da p i q budu beskonačni.

¹² Beta funkcija je definisana sledećom formulom: $B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$.

Na Slici 2.7 predstavljena je funkcija gustine *skew* generalizovane *t* raspodele za $\mu = 3$, $\sigma = 1$, $\lambda = 0,2$, $p = 5$ i $q = Inf$.



Slika 2.7: Funkcija gustine *skew* generalizovane *t* raspodele
Izvor: Izlaz iz programskog jezika R

Matematičko očekivanje, varijansa, treći i četvrti centralni momenat *skew* generalizovane *t* raspodele su predstavljeni sledećim formulama (Davis, 2015):

$$E(X) = \mu + \frac{2\nu\sigma\lambda q^{1/p} B\left(\frac{2}{p}, q - \frac{1}{p}\right)}{B\left(\frac{1}{p}, q\right)} - m, \text{ za } pq > 1,$$

$$\text{Var}(X) = (\nu\sigma)^2 q^{2/p} \left((3\lambda^2 + 1) \frac{B\left(\frac{3}{p}, q - \frac{2}{p}\right)}{B\left(\frac{1}{p}, q\right)} - 4\lambda^2 \frac{B\left(\frac{2}{p}, q - \frac{1}{p}\right)^2}{B\left(\frac{1}{p}, q\right)^2} \right), \text{ za } pq > 2,$$

$$E[X - E(X)]^3 = \frac{2(\nu\sigma)^3 \lambda q^{3/p}}{B\left(\frac{1}{p}, q\right)^3} \times \left(8\lambda^2 B\left(\frac{2}{p}, q - \frac{1}{p}\right)^3 - 3(3\lambda^2 + 1) B\left(\frac{1}{p}, q\right) B\left(\frac{2}{p}, q - \frac{1}{p}\right) B\left(\frac{3}{p}, q - \frac{2}{p}\right) + \right. \\ \left. + 2(\lambda^2 + 1) B\left(\frac{1}{p}, q\right)^2 B\left(\frac{4}{p}, q - \frac{3}{p}\right) \right), \text{ za } pq > 3,$$

$$E[X - E(X)]^4 = \frac{(v\sigma)^4 q^{4/p}}{B\left(\frac{1}{p}, q\right)^4} \times \left(-48\lambda^4 B\left(\frac{2}{p}, q - \frac{1}{p}\right)^4 + 24\lambda^2 (3\lambda^2 + 1) B\left(\frac{1}{p}, q\right) B\left(\frac{2}{p}, q - \frac{1}{p}\right)^2 B\left(\frac{3}{p}, q - \frac{2}{p}\right) - \right. \\ \left. - 32\lambda^2 (\lambda^2 + 1) B\left(\frac{1}{p}, q\right)^2 B\left(\frac{2}{p}, q - \frac{1}{p}\right) B\left(\frac{4}{p}, q - \frac{3}{p}\right) + (5\lambda^4 + 10\lambda^2 + 1) B\left(\frac{1}{p}, q\right)^3 B\left(\frac{5}{p}, q - \frac{4}{p}\right) \right), \text{ za } pq > 4.$$

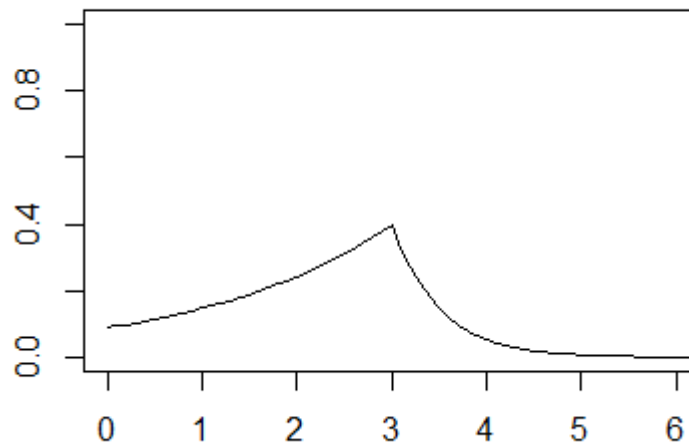
2.1.7 Skew-Laplace-ova raspodela

Slučajna promenljiva X ima *skew-Laplace-ovu* raspodelu sa parametrima oblika $\alpha > 0$ i $\beta > 0$ i parametrom lokacije $\mu \in \mathbf{R}$, ako je njena funkcija gustine (Puig & Stephens, 2007, str. 47):

$$f(x) = \begin{cases} \exp\left(\frac{x-\mu}{\alpha}\right) / (\alpha + \beta), & x \leq \mu \\ \exp\left(\frac{\mu-x}{\beta}\right) / (\alpha + \beta), & x > \mu. \end{cases} \quad (2.8)$$

Ako je $\alpha = \beta$, *skew-Laplace-ova* raspodela postaje klasična Laplace-ova raspodela, a ako α ili β teže 0 dobija se dvoparametarska eksponencijalna ili negativna eksponencijalna raspodela.

Na Slici 2.8 prikazana je funkcija gustine *skew-Laplace-ove* raspodele za proizvoljno izabrane vrednosti parametara $\alpha = 2$, $\beta = 0,5$ i $\mu = 3$.



Slika 2.8: Funkcija gustine *skew-Laplace-ove* raspodele

Izvor: Izlaz iz programskog jezika R

Matematičko očekivanje, varijansa, koeficijent asimetrije i koeficijent spljoštenosti *skew-Laplace-ove* raspodele su definisani na sledeći način (Puig & Stephens, 2007, str. 47):

$$E(X) = \mu + \beta - \alpha,$$

$$\text{Var}(X) = \alpha^2 + \beta^2,$$

$$\gamma_3(X) = \frac{2(\beta^3 - \alpha^3)}{(\alpha^2 + \beta^2)^{3/2}},$$

$$\gamma_4(X) = 3 + \frac{6(\beta^4 + \alpha^4)}{(\alpha^2 + \beta^2)^2}.$$

2.2 Intervali poverenja za regresioni koeficijent β_1 u prostom linearnom regresionom modelu

Postojeći interval poverenja za regresioni koeficijent β_1 u prostom linearnom regresionom modelu zasnovan je na t -statistici. Proporcija simuliranih t intervala biće bliska nominalnom nivou pouzdanosti ako slučajna greška u regresionom modelu sledi normalnu raspodelu. Međutim, postavlja se pitanje kada će proporcija simuliranih intervala biti bliska nominalnom nivou pouzdanosti ukoliko slučajna greška sledi neku asimetričnu raspodelu. Različite alternative mogu biti primenjene u tim situacijama.

U ovom odeljku biće razmatrani sledeći intervali poverenja za regresioni koeficijent β_1 u prostom linearnom regresionom modelu: interval poverenja zasnovan na t -statistici, *bootstrap-t* interval, *bootstrap* intervali poverenja zasnovani na postojećim transformacijama statistika, kao i interval poverenja zasnovan na novoj transformaciji statistike kombinovan sa *bootstrap* metodom (Ivković et al., 2020).

U literaturi se može naći nekoliko radova u kojima su se autori bavili ocenjivanjem nepoznatih regresionih parametara, kao npr. Antille i Milasević (1988), Wilcox (1996), Buchholz et al. (2008), Hall i Mayo (2008), Adrover i Salibian-Barrera (2010), George et al. (2013), Sobotka et al. (2013), Tóth i Somorčík (2017), i drugih. Sen (1968) je proučavao klasične i robusne ocene regresionih koeficijenata.

Malo je radova u kojima se proučavaju intervalne ocene regresionih parametara. Tako je Douglas (1996) koristio *bootstrap* metodu za konstrukciju intervala poverenja za regresioni koeficijent, dok su Jones i Waller (2013) proučavali intervale poverenja za standardizovane regresione koeficijente.

Raspodelom slučajne greške i odstupanjem te raspodele od normalnosti bavili su se autori Looney i Gullledge (1985), Jarque i Bera (1987), Yazici i Yolacan (2007) i drugi. Nekoliko autora je ponudilo rešenje za situacije u kojima slučajna greška nije normalno raspodeljena (Bartolucci & Scaccia, 2005; Soffritti & Galimberti, 2011).

Kada slučajna greška ne sledi normalnu raspodelu, jedan od načina za konstrukciju intervalnih ocena podrazumeva transformaciju t -statistike (Ivković et al., 2020). Nekoliko autora se bavilo transformacijom t -statistike ali prilikom konstrukcije intervalnih ocena sredine i varijanse. Johnson (1978) i Hall (1992a) razmatrali su transformacije t -statistike prilikom ocenjivanja sredine, zasnovano na jednom uzorku. Čojbašić i Lončar (2011) su se bavili jednostranim intervalima

poverenja za varijansu skupa koji sledi asimetričnu raspodelu. Takođe, ovom tematikom bavili su se Chen (1995), Zhou i Dinh (2005), Čojbašić i Tomović (2007) i Čojbašić i Stanojević (2013).

2.2.1 Interval poverenja zasnovan na t -statistici

Neka je:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i,$$

prost linearni regresioni model osnovnog skupa sa parametrima β_0 i β_1 . U ovom modelu X_i je objašnjavajuća promenljiva, Y_i je zavisna promenljiva a ε_i je stohastički član odnosno slučajna greška (pri čemu je $E(\varepsilon_i) = 0$ i $\text{Var}(\varepsilon_i) = \sigma_\varepsilon^2$).

Ukoliko su ispunjene sve pretpostavke prostog linearnog regresionog modela interval poverenja za parametar β_1 zasnovan je na t -statistici:

$$T = \frac{\mathbf{b}_1 - \beta_1}{S_{\mathbf{b}_1}}, \quad (2.9)$$

gde je \mathbf{b}_1 ocena¹³ parametra β_1 a $S_{\mathbf{b}_1}$ je ocena standardne greške ocene \mathbf{b}_1 . Interval poverenja pouzdanosti $(1-\alpha)100\%$ za regresioni koeficijent β_1 u prostom linearnom regresionom modelu, zasnovan na statistici (2.9), sledećeg je oblika:

$$I_{\beta_1} = \left[\mathbf{b}_1 - t_{n-2; \alpha/2} S_{\mathbf{b}_1}, \mathbf{b}_1 + t_{n-2; \alpha/2} S_{\mathbf{b}_1} \right], \quad (2.10)$$

gde je t_{n-2} slučajna promenljiva koja ima t raspodelu sa $n-2$ stepena slobode, a $t_{n-2; \alpha/2}$ određeno je sa $P(t_{n-2} > t_{n-2; \alpha/2}) = \alpha/2$. Kao što je već navedeno u uvodnom delu odeljka, kada je slučajna greška normalno raspodeljena proporcija simuliranih t intervala bliska je nominalnom nivou pouzdanosti $1-\alpha$.

2.2.2 Bootstrap- t interval poverenja

Ukoliko slučajna greška sledi neku asimetričnu raspodelu, moguće je konstruisati *bootstrap- t* interval za nepoznati parametar β_1 . On je zasnovan na sledećoj statistici:

¹³ Ocena \mathbf{b}_1 zasnovana na uzorku $(X_1, Y_1), \dots, (X_n, Y_n)$ data je formulom $\mathbf{b}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$.

$$T^* = \frac{\mathbf{b}_1^* - \mathbf{b}_1}{S_{\mathbf{b}_1^*}}, \quad (2.11)$$

gde je \mathbf{b}_1^* vrednost ocene \mathbf{b}_1 na osnovu *bootstrap* uzoraka a $S_{\mathbf{b}_1^*}$ je ocena standardne greške statistike \mathbf{b}_1^* . Ovaj interval može se predstaviti na sledeći način:

$$I_{\beta_1} = \left[\mathbf{b}_1 - \hat{t}_{1-\alpha/2} S_{\mathbf{b}_1}, \mathbf{b}_1 - \hat{t}_{\alpha/2} S_{\mathbf{b}_1} \right], \quad (2.12)$$

gde $\hat{t}_{1-\alpha/2}$ i $\hat{t}_{\alpha/2}$ predstavljaju $1-\alpha/2$ i $\alpha/2$ percentile *bootstrap* raspodele statistike T^* .

2.2.3 Edgeworth-ov razvoj raspodele t -statistike

Ukoliko slučajna greška nije normalno raspodeljena, proporcije simuliranih t intervala mogu se znatno razlikovati od nominalnog nivoa pouzdanosti. Jedan od načina za prevazilaženje ovog problema jeste da se izvrši transformacija t -statistike. Postupak koji je potrebno sprovesti sastoji se u sledećem: izvesti Edgeworth-ov razvoj (Edgeworth, 1898) raspodele t -statistike (2.9) do reda $n^{-1/2}$ i na osnovu razvoja predložiti transformaciju t -statistike. Pored nove transformacije t -statistike biće predstavljene i dve postojeće transformacije. Sve navedene transformacije biće upotrebljene za konstruisanje *bootstrap-t* intervala poverenja za regresioni koeficijent (Ivković et al., 2020).

Teorema 1. Pod odgovarajućim uslovima regularnosti¹⁴ raspodela t -statistike date izrazom (2.9) ima sledeći razvoj:

$$P(T \leq x) = \Phi(x) + \frac{1}{\sqrt{n}} q(x) \phi(x) + O(n^{-1}). \quad (2.13)$$

U izrazu (2.13), $q(x) = \frac{M_3'}{6}(2x^2 + 1)$, a $\phi(\cdot)$ i $\Phi(\cdot)$ predstavljaju funkciju gustine i funkciju raspodele standardizovane normalne promenljive. Treći momenat M_3' računa se koristeći formulu $M_3' = E[(1/n) \sum_{i=1}^n X_i'^3]$, a slučajne promenljive $X_i', i=1, 2, \dots, n$ definisane su na sledeći način (Ivković et al., 2020):

$$X_i' = \frac{\frac{(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} - \beta_1}{\sqrt{V}}, \quad i = 1, 2, \dots, n, \quad (2.14)$$

¹⁴Analogno Hall (1992b, str.71) uslovi regularnosti su:

1) Cramer-ov uslov $\limsup_{|t_1|+|t_2| \rightarrow \infty} \left| E \exp(it_1 X_j + it_2 X_j^2) \right| < 1$ i 2) $E(X_i'^9) < \infty$ gde je X_i' definisano u (2.14).

gde je:

$$\bar{X} = (1/n) \sum_{i=1}^n X_i, \quad \bar{Y} = (1/n) \sum_{i=1}^n Y_i$$

i

$$V = E \left\{ \left[\frac{(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} - \frac{\beta_1}{n} \right]^2 \right\}.$$

Dokaz. Kako je ocena parametra β_1 oblika:

$$b_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2},$$

i $E(b_1) = \beta_1$ sledi da je:

$$E \left(\frac{(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} \right) = \frac{\beta_1}{n}, i = 1, 2, \dots, n.$$

Dalje će biti izvršena standardizacija slučajnih promenljivih:

$$\frac{(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}, \quad i = 1, 2, \dots, n.$$

Njihova očekivana vrednost je $\frac{\beta_1}{n}$, a varijansa je $V = E \left\{ \left[\frac{(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} - \frac{\beta_1}{n} \right]^2 \right\}$. Nove

standardizovane promenljive biće obeležene sa X'_i ($i = 1, 2, \dots, n$) i one su oblika:

$$X'_i = \frac{\frac{(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} - \frac{\beta_1}{n}}{\sqrt{V}}, i = 1, 2, \dots, n. \quad (2.15)$$

Zbog načina na koji su dobijene (standardizacijom) za ove promenljive važi da je $E(X'_i) = 0$ i $Var(X'_i) = 1$. Prosek promenljivih (2.15) je oblika:

$$\frac{1}{n} \sum_{i=1}^n X'_i = \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) - \frac{1}{n} \beta_1}{\sqrt{V}}, \quad (2.16)$$

a u daljem tekstu biće označen sa Z_1 . Rešavajući jednakost (2.16) po oceni \mathbf{b}_1 dobija se (Ivković et al., 2020):

$$\mathbf{b}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} = n(Z_1 \sqrt{V} + \frac{1}{n} \beta_1) = Z_1 \sqrt{V} n + \beta_1. \quad (2.17)$$

Iz (2.17) sledi da je varijansa ocene \mathbf{b}_1 jednaka:

$$\text{Var}(\mathbf{b}_1) = \text{Var}(Z_1 \sqrt{V} n + \beta_1) = n^2 V \text{Var}(Z_1) = n^2 V \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X'_i\right) = n V \text{Var}(X'_i),$$

odnosno njena standardna greška je oblika $\sqrt{n V \text{Var}(X'_i)}$. Ocena standardne greške ocene \mathbf{b}_1 određuje se na sledeći način:

$$S_{\mathbf{b}_1} = \sqrt{n V \text{Var}(X'_i)}. \quad (2.18)$$

Koristeći formulu za varijansu uzorka dobija se

$$S_{\mathbf{b}_1} = \sqrt{n} \sqrt{V} \sqrt{\frac{\sum_{i=1}^n X_i'^2 - \frac{\left(\sum_{i=1}^n X_i'\right)^2}{n}}{n-1}} = n \sqrt{V} \sqrt{\frac{1}{n-1}} \sqrt{Z_2 - Z_1^2}, \quad (2.19)$$

gde je $Z_2 = \frac{1}{n} \sum_{i=1}^n X_i'^2$.

Zamenom (2.17) i (2.19) u (2.9), statistika T postaje (Ivković et al., 2020):

$$T = \frac{n Z_1 \sqrt{V} + \beta_1 - \beta_1}{n \sqrt{V} \sqrt{\frac{1}{n-1}} \sqrt{Z_2 - Z_1^2}} = \sqrt{n-1} \frac{Z_1}{\sqrt{Z_2 - Z_1^2}}. \quad (2.20)$$

U sledećem koraku se definiše statistika W_n :

$$\begin{aligned}
W_n = & \sqrt{n-1} \left(\frac{\partial g}{\partial Z_1}(U) \cdot (Z_1 - U_1) + \frac{\partial g}{\partial Z_2}(U) \cdot (Z_2 - U_2) + \right. \\
& \left. + \frac{1}{2} \left[\frac{\partial^2 g}{\partial Z_1^2}(U) \cdot (Z_1 - U_1)^2 + 2 \frac{\partial^2 g}{\partial Z_1 \partial Z_2}(U) \cdot (Z_1 - U_1)(Z_2 - U_2) + \frac{\partial^2 g}{\partial Z_2^2}(U) \cdot (Z_2 - U_2)^2 \right] \right), \quad (2.21)
\end{aligned}$$

gde je $g(Z) = \frac{Z_1}{\sqrt{Z_2 - Z_1^2}}$ ¹⁵ i $E(Z_1, Z_2) \equiv U \equiv (U_1, U_2) = (0, 1)$. Izračunavanjem izvoda u izrazu (2.21), dobija se sledeće:

$$\frac{\partial g}{\partial Z_1}(0, 1) = (Z_2 - Z_1^2)^{-1/2} + Z_1^2 (Z_2 - Z_1^2)^{-3/2} \Big|_{Z_1=0, Z_2=1} = 1,$$

$$\frac{\partial g}{\partial Z_2}(0, 1) = -\frac{1}{2} Z_1 (Z_2 - Z_1^2)^{-3/2} \Big|_{Z_1=0, Z_2=1} = 0,$$

$$\frac{\partial^2 g}{\partial Z_1^2}(0, 1) = 3Z_1 (Z_2 - Z_1^2)^{-3/2} + 3Z_1^3 (Z_2 - Z_1^2)^{-5/2} \Big|_{Z_1=0, Z_2=1} = 0,$$

$$\frac{\partial^2 g}{\partial Z_1 \partial Z_2}(0, 1) = -\frac{1}{2} (Z_2 - Z_1^2)^{-3/2} - \frac{3}{2} Z_1^2 (Z_2 - Z_1^2)^{-5/2} \Big|_{Z_1=0, Z_2=1} = -\frac{1}{2}$$

i

$$\frac{\partial^2 g}{\partial Z_2^2}(0, 1) = \frac{3}{4} Z_1 (Z_2 - Z_1^2)^{-5/2} \Big|_{Z_1=0, Z_2=1} = 0.$$

Zamenom dobijenih vrednosti u (2.21) statistika W_n se može izraziti na sledeći način:

$$W_n = \sqrt{n-1} \left(\frac{3}{2} Z_1 - \frac{1}{2} Z_1 Z_2 \right).$$

Dalje, potrebno je odrediti prva tri momenta statistike W_n . Prvi momenat se može predstaviti sledećom formulom:

$$EW_n = E\left[\sqrt{n-1} \left(\frac{3}{2} Z_1 - \frac{1}{2} Z_1 Z_2 \right)\right] = \sqrt{n-1} \left[\frac{3}{2} E(Z_1) - \frac{1}{2} E(Z_1 Z_2) \right] = -\frac{1}{2} \frac{1}{\sqrt{n}} M_3' + O(n^{-1}),$$

gde je $\sqrt{n-1} = \sqrt{n} - \frac{1}{2\sqrt{n}} + o(n^{-1/2})$, $E(Z_1) = E\left(\frac{1}{n} \sum_{i=1}^n X_i'\right) = \frac{1}{n} E\left(\sum_{i=1}^n X_i'\right) = \frac{1}{n} \sum_{i=1}^n E(X_i') = 0$

¹⁵ Pretpostavlja se da je $g(Z)$ dva puta diferencijabilna u okolini tačke U .

i

$$\begin{aligned}
 E(Z_1 Z_2) &= E\left(\frac{1}{n} \sum_{i=1}^n X_i' \cdot \frac{1}{n} \sum_{i=1}^n X_i'^2\right) = \frac{1}{n^2} E\left(\sum_{i=1}^n X_i' \cdot \sum_{i=1}^n X_i'^2\right) = \frac{1}{n^2} E\left(\sum_{i=1}^n X_i'^3 + \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n X_i' X_j'^2\right) \\
 &= \frac{1}{n^2} \left[E\left(\sum_{i=1}^n X_i'^3\right) + \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n E(X_i') E(X_j'^2) \right] = \frac{1}{n^2} E\left(\sum_{i=1}^n X_i'^3\right) = \frac{1}{n} E\left(\frac{1}{n} \sum_{i=1}^n X_i'^3\right) = \frac{1}{n} M_3'.
 \end{aligned}$$

Zamenom $E(Z_1)$ i $E(Z_1 Z_2)$ u EW_n dobija se predstavljeni rezultat.

Drugi momenat se određuje na sledeći način:

$$EW_n^2 = E\left[\sqrt{n-1} \left(\frac{3}{2} Z_1 - \frac{1}{2} Z_1 Z_2\right)^2\right] = (n-1) \left[\frac{9}{4} E(Z_1^2) - \frac{3}{2} E(Z_1^2 Z_2) + \frac{1}{4} E(Z_1^2 Z_2^2) \right] = 1 + O(n^{-1}),$$

gde je

$$\begin{aligned}
 E(Z_1^2) &= E\left(\frac{1}{n} \sum_{i=1}^n X_i'\right)^2 = \frac{1}{n^2} E\left(\sum_{i=1}^n X_i'\right)^2 = \frac{1}{n^2} \left[E\left(\sum_{i=1}^n X_i'^2\right) + E\left(\sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n X_i' X_j'\right) \right] \\
 &= \frac{1}{n^2} \left[\sum_{i=1}^n E(X_i'^2) + \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n E(X_i') E(X_j') \right] = \frac{1}{n},
 \end{aligned}$$

$$\begin{aligned}
 E(Z_1^2 Z_2) &= E\left[\left(\frac{1}{n} \sum_{i=1}^n X_i'\right)^2 \cdot \frac{1}{n} \sum_{i=1}^n X_i'^2\right] = \frac{1}{n^3} E\left[\left(\sum_{i=1}^n X_i'\right)^2 \cdot \sum_{i=1}^n X_i'^2\right] = \frac{1}{n^3} E\left[\left(\sum_{i=1}^n X_i'^2 + \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n X_i' X_j'\right) \cdot \sum_{i=1}^n X_i'^2\right] \\
 &= \frac{1}{n^3} E\left(\sum_{i=1}^n X_i'^4 + \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n X_i'^2 X_j'^2 + \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n X_i'^3 X_j' + \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n \sum_{\substack{k=1 \\ i \neq j \neq k}}^n X_i'^2 X_j' X_k'\right) \\
 &= \frac{1}{n^3} \left[\sum_{i=1}^n E(X_i'^4) + \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n E(X_i'^2) E(X_j'^2) + \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n E(X_i'^3) E(X_j') + \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n \sum_{\substack{k=1 \\ i \neq j \neq k}}^n E(X_i'^2) E(X_j') E(X_k') \right] \\
 &= \frac{1}{n^3} \left[E\left(\sum_{i=1}^n X_i'^4\right) + n(n-1) \right]
 \end{aligned}$$

i

$$\begin{aligned}
E(Z_1^2 Z_2^2) &= E\left[\left(\frac{1}{n} \sum_{i=1}^n X_i'\right)^2 \cdot \left(\frac{1}{n} \sum_{i=1}^n X_i'^2\right)^2\right] = \frac{1}{n^4} E\left[\left(\sum_{i=1}^n X_i'\right)^2 \cdot \left(\sum_{i=1}^n X_i'^2\right)^2\right] \\
&= \frac{1}{n^4} E\left[\left(\sum_{i=1}^n X_i'^2 + \sum_{\substack{i=1 \\ j=1 \\ i \neq j}}^n \sum_{j=1}^n X_i' X_j'\right) \cdot \left(\sum_{i=1}^n X_i'^4 + \sum_{\substack{i=1 \\ j=1 \\ i \neq j}}^n \sum_{j=1}^n X_i'^2 X_j'^2\right)\right] \\
&= \frac{1}{n^4} E\left[\sum_{i=1}^n X_i'^6 + \sum_{\substack{i=1 \\ j=1 \\ i \neq j}}^n \sum_{j=1}^n X_i'^2 X_j'^4 + \sum_{\substack{i=1 \\ j=1 \\ i \neq j}}^n \sum_{j=1}^n X_i'^4 X_j'^2 + \sum_{\substack{i=1 \\ j=1 \\ k=1 \\ i \neq j \neq k}}^n \sum_{j=1}^n \sum_{k=1}^n X_i'^2 X_j'^2 X_k'^2 + \right. \\
&\quad \left. + \sum_{\substack{i=1 \\ j=1 \\ i \neq j}}^n \sum_{j=1}^n X_i'^5 X_j' + \sum_{\substack{i=1 \\ j=1 \\ k=1 \\ i \neq j \neq k}}^n \sum_{j=1}^n \sum_{k=1}^n X_i'^4 X_j' X_k' + \sum_{\substack{i=1 \\ j=1 \\ k=1 \\ l=1 \\ i \neq j \neq k \neq l}}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n X_i'^2 X_j'^2 X_k' X_l' + \right. \\
&\quad \left. + \sum_{\substack{i=1 \\ j=1 \\ i \neq j}}^n \sum_{j=1}^n X_i'^3 X_j'^3 + \sum_{\substack{i=1 \\ j=1 \\ k=1 \\ i \neq j \neq k}}^n \sum_{j=1}^n \sum_{k=1}^n X_i'^3 X_j'^2 X_k'\right) \\
&= \frac{1}{n^4} \left[E\left(\sum_{i=1}^n X_i'^6\right) + 2E\left(\sum_{\substack{i=1 \\ j=1 \\ i \neq j}}^n \sum_{j=1}^n X_i'^4 X_j'^2\right) + n(n-1)(n-2) + E\left(\sum_{\substack{i=1 \\ j=1 \\ i \neq j}}^n \sum_{j=1}^n X_i'^3 X_j'^3\right) \right].
\end{aligned}$$

Treći momenat predstavljen je sledećom formulom:

$$\begin{aligned}
EW_n^3 &= E\left[\sqrt{n-1} \left(\frac{3}{2} Z_1 - \frac{1}{2} Z_1 Z_2\right)\right]^3 \\
&= (n-1) \sqrt{n-1} \left[\frac{27}{8} E(Z_1^3) - \frac{27}{8} E(Z_1^3 Z_2) + \frac{9}{8} E(Z_1^3 Z_2^2) - \frac{1}{8} E(Z_1^3 Z_2^3) \right] \\
&= -\frac{7}{2} \frac{1}{\sqrt{n}} M_3' + O(n^{-1}),
\end{aligned}$$

gde se posle niza izračunavanja dobija¹⁶:

$$E(Z_1^3) = E\left(\frac{1}{n} \sum_{i=1}^n X_i'\right)^3 = \frac{1}{n^2} M_3',$$

$$E(Z_1^3 Z_2) = \frac{1}{n^3} M_5' + 4 \frac{1}{n^2} M_3' - 4 \frac{1}{n^3} M_3',$$

$$E(Z_1^3 Z_2^2) = \frac{1}{n^4} M_7' + 7 \frac{1}{n^3} M_4' M_3' - 7 \frac{1}{n^4} M_4' M_3' + 5 \frac{1}{n^3} M_5' - 5 \frac{1}{n^4} M_5' + 7 \frac{1}{n^2} M_3' - 21 \frac{1}{n^3} M_3' + 14 \frac{1}{n^4} M_3'$$

i

¹⁶ Na primer, peti momenat je jednak $M_5' = E\left[(1/n) \sum_{i=1}^n X_i'^5\right]$. Na sličan način određuju se i ostali momenti.

$$\begin{aligned}
E(Z_1^3 Z_2^3) &= 10 \frac{1}{n^4} M_6' M_3' - 10 \frac{1}{n^5} M_6' M_3' + \frac{1}{n^5} M_9' + 12 \frac{1}{n^3} M_5' - 36 \frac{1}{n^4} M_5' + 24 \frac{1}{n^5} M_5' + 10 \frac{1}{n^2} M_3' - 60 \frac{1}{n^3} M_3' \\
&+ 110 \frac{1}{n^4} M_3' - 60 \frac{1}{n^5} M_3' + 6 \frac{1}{n^4} M_7' - 6 \frac{1}{n^5} M_7' + 13 \frac{1}{n^3} M_4' M_3' - 39 \frac{1}{n^4} M_4' M_3' + 26 \frac{1}{n^5} M_4' M_3' \\
&+ 12 \frac{1}{n^4} M_5' M_4' - 12 \frac{1}{n^5} M_5' M_4' + \frac{1}{n^3} M_3'^3 - 3 \frac{1}{n^4} M_3'^3 + 2 \frac{1}{n^5} M_3'^3.
\end{aligned}$$

Na osnovu izračunatih momenata određuju se kumulante statistike W_n :

$$K_{1N} = EW_n = -\frac{1}{2} \frac{1}{\sqrt{n}} M_3' + O(n^{-1}),$$

$$K_{2N} = EW_n^2 - (EW_n)^2 = [1 + O(n^{-1})] - \left[-\frac{1}{2} \frac{1}{\sqrt{n}} M_3' + O(n^{-1})\right]^2 = 1 + O(n^{-1})$$

i

$$\begin{aligned}
K_{3N} &= E(W_n - EW_n)^3 = EW_n^3 - 3EW_n^2 EW_n + 2(EW_n)^3 \\
&= \left[-\frac{7}{2} \frac{1}{\sqrt{n}} M_3' + O(n^{-1})\right] - 3[1 + O(n^{-1})] \left[-\frac{1}{2} \frac{1}{\sqrt{n}} M_3' + O(n^{-1})\right] \\
&\quad + 2\left[-\frac{1}{2} \frac{1}{\sqrt{n}} M_3' + O(n^{-1})\right]^3 \\
&= -2 \frac{1}{\sqrt{n}} M_3' + O(n^{-1}).
\end{aligned}$$

Kumulante statistike W_n koriste se za izračunavanje karakteristične funkcije statistike W_n :

$$\begin{aligned}
\psi_n(t) &= \exp \left\{ K_{1N}(it) + K_{2N} \frac{(it)^2}{2} + K_{3N} \frac{(it)^3}{6} + \dots \right\} \\
&= \exp \left(-\frac{t^2}{2} \right) \left\{ 1 - \frac{1}{2} \frac{1}{\sqrt{n}} M_3'(it) - \frac{2}{\sqrt{n}} M_3' \frac{(it)^3}{6} + O(n^{-1}) \right\} \\
&= \exp \left(-\frac{t^2}{2} \right) \left\{ 1 + \frac{1}{\sqrt{n}} M_3' \left(-\frac{1}{2}(it) - \frac{1}{3}(it)^3 \right) + O(n^{-1}) \right\}.
\end{aligned}$$

Prema Čojbašić i Tomović (2007), uzimajući u obzir da je $\psi_n(t) = \int_{-\infty}^{+\infty} e^{itx} dP(W_n \leq x)$, kao i da važi $e^{-t^2/2} = \int_{-\infty}^{+\infty} e^{itx} d\Phi(x)$ i $T = W_n + O(n^{-1})$, sledi Edgeworth-ov razvoj raspodele t -statistike predstavljen izrazom (2.13). \square

U nastavku će biti predstavljene tri kubne transformacije t -statistike, koje su konstruisane da bi se korigovala asimetrija raspodele. Hall (1992a) je predložio sledeće dve transformacije t -statistike koje su oblika:

$$T_1 = T_1(U) = U + \frac{1}{3} \hat{M}_3' U^2 + \frac{1}{27} \hat{M}_3'^2 U^3 + \frac{1}{6} \frac{1}{n} \hat{M}_3',$$

$$T_2 = T_2(U) = \left(2 \frac{1}{3} \frac{1}{\sqrt{n}} \hat{M}_3' \right)^{-1} \left\{ \exp \left(2 \frac{1}{3} \frac{1}{\sqrt{n}} \hat{M}_3' U \right) - 1 \right\} + \frac{1}{6} \frac{1}{n} \hat{M}_3',$$

gde je $U = T / \sqrt{n}$, a \hat{M}_3' je ocena momenta M_3' ,

Hall (1992a) je dodao član $\frac{1}{27} \hat{M}_3'^2 U^3$ Johnson-ovoj kvadratnoj transformaciji¹⁷ $U + \frac{1}{3} \hat{M}_3' U^2 + \frac{1}{6} \frac{1}{n} \hat{M}_3'$ i dobio kubnu transformaciju T_1 , koja je invertibilna i primenljiva prilikom konstruisanja intervala poverenja. Transformacija T_2 je član familije transformacija koje su razmatrali Mallows i Tukey (1982).

Navedene transformacije zadovoljavaju sledeće uslove (Hoaglin, 1985):

- $T_{transf}(U) - \hat{M}_3' / 6n$ je približno U za U blisko nuli¹⁸,
- $T_{transf}(U) - \hat{M}_3' / 6n$ je jednako nuli za U jednako nuli.

Nakon analize prethodnih uslova, uvedena je nova kubna transformacija (Ivković et al., 2020):

$$T_3 = T_3(U) = U + \frac{1}{2} U^2 + \frac{1}{12} U^3 + \frac{1}{6} \frac{1}{n} \hat{M}_3'. \quad (2.22)$$

Sve transformacije su monotone i imaju eksplicitne inverzne formule, na osnovu čega se mogu primeniti na probleme konstrukcije intervalnih ocena. U cilju konstrukcije intervala poverenja potrebno je odrediti inverznu funkciju transformacije. Inverzna funkcija od $T_i(\cdot)$, $i=1,2,3$ može biti rešena analitički i ima sledeće oblike (Hall, 1992a; Ivković et al., 2020):

$$T_1^{-1}(t) = \left(\frac{1}{3} \hat{M}_3' \right)^{-1} \left[1 + 3 \frac{1}{3} \hat{M}_3' \left(t - \frac{1}{6} \frac{1}{n} \hat{M}_3' \right) \right]^{1/3} - \left(\frac{1}{3} \hat{M}_3' \right)^{-1},$$

$$T_2^{-1}(t) = \left(2 \frac{1}{3} \frac{1}{\sqrt{n}} \hat{M}_3' \right)^{-1} \log \left[2 \frac{1}{3} \frac{1}{\sqrt{n}} \hat{M}_3' \left(t - \frac{1}{6} \frac{1}{n} \hat{M}_3' \right) + 1 \right],$$

¹⁷ Ova kvadratna transformacija je predložena u radu Johnson (1978).

¹⁸ T_{transf} predstavlja bilo koju transformaciju t -statistike.

$$T_3^{-1}(t) = \left[8 + 12 \left(t - \frac{1}{6} \frac{1}{n} \hat{M}_3' \right) \right]^{1/3} - 2.$$

Koristeći inverznu funkciju, transformisani $(1-\alpha)100\%$ *bootstrap-t* interval poverenja za regresioni koeficijent je oblika (Ivković et al., 2020):

$$I_{\beta_i} = \left[\mathbf{b}_1 - \sqrt{n} T_i^{-1} \left(\frac{\hat{t}^{(1-\alpha/2)}}{\sqrt{n}} \right) S_{b_1}, \mathbf{b}_1 - \sqrt{n} T_i^{-1} \left(\frac{\hat{t}^{(\alpha/2)}}{\sqrt{n}} \right) S_{b_1} \right], \quad (2.23)$$

gde je $i = 1, 2, 3$. Intervali poverenja zasnovani na transformacijama T_1 , T_2 i T_3 u daljem radu označeni su kao T_1 , T_2 i T_3 intervali.

2.3 Intervali poverenja za regresione koeficijente β_1 i β_2 u kvadratnom regresionom modelu

Postojeći intervali poverenja za regresione koeficijente β_1 i β_2 u kvadratnom regresionom modelu zasnovani su na t -statistici i adekvatni su ako slučajna greška sledi normalnu raspodelu. Intervali poverenja koji mogu biti bolji izbor ukoliko slučajna greška nije normalno raspodeljena zasnovani su na metodama ponovljenih uzoraka. Takođe, moguće je izvršiti transformaciju t -statistike, koja će se koristiti za konstrukciju novog intervala poverenja. Transformisana statistika se predlaže na osnovu Edgeworth-ovog razvoja raspodele t -statistike do reda n^{-1} i ima za cilj korigovanje asimetrije raspodele, zbog čega se očekuje da će proporcije simuliranih novih intervala biti bliže nominalnom nivou pouzdanosti u odnosu na proporcije simuliranih t intervala.

Kvadratnim regresionim modelima i ocenama parametara modela bavili su se sledeći autori: Piegorsch (1986), Galindo et al. (2001), Jarušková (2001), Kuha i Temple (2003), Hirschberg i Lye (2004), Košmelj et al. (2005), Shacham et al. (2007), Yao i Müller (2010), Horváth i Reeder (2013), Han et al. (2016) i drugi.

2.3.1 Intervali poverenja zasnovani na t -statistikama

Kvadratni regresioni model je oblika:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i.$$

U ovom modelu, β_0 , β_1 i β_2 predstavljaju parametre, Y_i je zavisna promenljiva, X_i je objašnjavajuća promenljiva a ε_i je slučajna greška (pri čemu je $E(\varepsilon_i) = 0$ i $\text{Var}(\varepsilon_i) = \sigma_\varepsilon^2$).

Postojeći intervali poverenja za regresione koeficijente, β_1 i β_2 , zasnovani na sledećim statistikama:

$$T_1 = \frac{\mathbf{b}_1 - \beta_1}{S_{b_1}} \quad (2.24)$$

i

$$T_2 = \frac{\mathbf{b}_2 - \beta_2}{S_{b_2}}, \quad (2.25)$$

pogodni su za primenu ukoliko slučajna greška sledi normalnu raspodelu. U izrazima (2.24) i (2.25), \mathbf{b}_1 i \mathbf{b}_2 predstavljaju ocene¹⁹ parametara β_1 i β_2 , a S_{b_1} i S_{b_2} su ocene standardnih grešaka ocena \mathbf{b}_1 i \mathbf{b}_2 . Na statistikama (2.24) i (2.25) zasnovani su intervali poverenja za regresione koeficijente:

$$I_{\beta_1} = \left[\mathbf{b}_1 - t_{n-3; \alpha/2} S_{b_1}, \mathbf{b}_1 + t_{n-3; \alpha/2} S_{b_1} \right] \quad (2.26)$$

i

$$I_{\beta_2} = \left[\mathbf{b}_2 - t_{n-3; \alpha/2} S_{b_2}, \mathbf{b}_2 + t_{n-3; \alpha/2} S_{b_2} \right], \quad (2.27)$$

gde je $t_{n-3; \alpha/2}$ određeno sa $P(t_{n-3} > t_{n-3; \alpha/2}) = \alpha/2$ i t_{n-3} je slučajna promenljiva koja ima Studentovu raspodelu sa $n-3$ stepeni slobode.

¹⁹ Ocena \mathbf{b}_1 data je formulom

$$\mathbf{b}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) \sum_{i=1}^n (X_i^2 - \bar{X}^2) - \sum_{i=1}^n (X_i^2 - \bar{X}^2)(Y_i - \bar{Y}) \sum_{i=1}^n (X_i - \bar{X})(X_i^2 - \bar{X}^2)}{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (X_i^2 - \bar{X}^2) - \left(\sum_{i=1}^n (X_i - \bar{X})(X_i^2 - \bar{X}^2) \right)^2},$$

dok se ocena \mathbf{b}_2 računa na sledeći način

$$\mathbf{b}_2 = \frac{\sum_{i=1}^n (X_i^2 - \bar{X}^2)(Y_i - \bar{Y}) \sum_{i=1}^n (X_i - \bar{X})^2 - \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) \sum_{i=1}^n (X_i - \bar{X})(X_i^2 - \bar{X}^2)}{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (X_i^2 - \bar{X}^2) - \left(\sum_{i=1}^n (X_i - \bar{X})(X_i^2 - \bar{X}^2) \right)^2},$$

a posle niza sređivanja dobija se

$$\mathbf{b}_2 = \frac{\sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n (X_i - \bar{X})(Y_i - \bar{Y})(X_j - \bar{X})^2 (X_i - X_j)}{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (X_i^2 - \bar{X}^2) - \left(\sum_{i=1}^n (X_i - \bar{X})(X_i^2 - \bar{X}^2) \right)^2}.$$

2.3.2 Intervali poverenja zasnovani na metodama ponovljenih uzoraka

Ukoliko slučajna greška nije normalno raspodeljena, intervali poverenja za regresione koeficijente kvadratnog regresionog modela koji su zasnovani na metodama ponovljenih uzoraka se koriste kao alternativa intervalima poverenja koji su zasnovani na t -statistici.

Konstruisanje intervala poverenja zasnovanih na *bootstrap* metodi podrazumeva da se iz prostog slučajnog uzorka izabere B *bootstrap* uzoraka i u svakom od B *bootstrap* uzoraka izračunaju vrednosti statistika b_1 i b_2 , označene sa b_1^* i b_2^* . U nastavku je opisana procedura samo za konstruisanje intervala poverenja za regresioni koeficijent β_2 , a sličan postupak se može primeniti za konstruisanje intervala za β_1 .

Bootstrap-t interval poverenja

Prilikom konstruisanja *bootstrap-t* intervala poverenja u svakom od B *bootstrap* uzoraka izračunava se vrednost statistike $T^* = \frac{b_2^* - b_2}{S_{b_2^*}}$, gde $S_{b_2^*}$ predstavlja ocenu standardne greške statistike b_2^* . Interval poverenja za regresioni koeficijent β_2 pouzdanosti $(1-\alpha)100\%$ definisan je sledećom formulom:

$$I_{\beta_2} = \left[b_2 - \hat{t}_{1-\alpha/2} S_{b_2}, b_2 - \hat{t}_{\alpha/2} S_{b_2} \right], \quad (2.28)$$

gde $\hat{t}_{1-\alpha/2}$ i $\hat{t}_{\alpha/2}$ predstavljaju $1-\alpha/2$ i $\alpha/2$ percentile *bootstrap* raspodele statistike T^* .

Percentil interval poverenja

Za konstrukciju percentil intervala poverenja prvo se vrši rangiranje vrednosti statistike b_2^* (na osnovu *bootstrap* uzoraka) u neopadajućem poretku. Zatim se određuju $\alpha/2$ i $1-\alpha/2$ percentili raspodele *bootstrap* replika, označeni sa $b_{2_{\alpha/2}}^*$ i $b_{2_{1-\alpha/2}}^*$. Interval poverenja pouzdanosti $(1-\alpha)100\%$ za regresioni koeficijent β_2 je oblika:

$$I_{\beta_2} = \left[b_{2_{\alpha/2}}^*, b_{2_{1-\alpha/2}}^* \right]. \quad (2.29)$$

BCa interval poverenja

BCa interval poverenja pouzdanosti $(1-\alpha)100\%$ za regresioni koeficijent β_2 predstavljen je formulom:

$$I_{\beta_2} = \left[b_{2_{\alpha_1}}^*, b_{2_{\alpha_2}}^* \right]. \quad (2.30)$$

Za percentile $b_{2_{\alpha_1}}^*$ i $b_{2_{\alpha_2}}^*$, vrednosti α_1 i α_2 izračunate su na osnovu sledećih formula:

$\alpha_1 = \Phi \left[\hat{z}_0 + \frac{\hat{z}_0 + Z_{\alpha/2}}{1 - \hat{a}(\hat{z}_0 + Z_{\alpha/2})} \right]$ i $\alpha_2 = \Phi \left[\hat{z}_0 + \frac{\hat{z}_0 + Z_{1-\alpha/2}}{1 - \hat{a}(\hat{z}_0 + Z_{1-\alpha/2})} \right]$ (pogledati odeljak 1.3). Ovde \hat{z}_0 predstavlja korekciju pristrasnosti, \hat{a} je akceleracija, dok $Z_{\alpha/2}$ i $Z_{1-\alpha/2}$ predstavljaju $\alpha/2$ i $1-\alpha/2$ percentile standardizovane normalne raspodele.

BC interval poverenja

BC interval poverenja pouzdanosti $(1-\alpha)100\%$ je oblika:

$$I_{\beta_2} = \left[b_{2_{\alpha_1}}^*, b_{2_{\alpha_2}}^* \right]. \quad (2.31)$$

Vrednosti α_1 i α_2 za percentile $b_{2_{\alpha_1}}^*$ i $b_{2_{\alpha_2}}^*$ izračunavaju se korišćenjem istih formula kao u slučaju BCa intervala.

Jackknife interval poverenja

Za potrebe formiranja *jackknife* intervala poverenja u svakom *jackknife* uzorku izračunava se vrednost statistike b_2 , koja je označena sa $b_{2(i)}$. *Jackknife* interval poverenja pouzdanosti $(1-\alpha)100\%$ za regresioni koeficijent β_2 definisan je formulom:

$$I_{\beta_2} = \left[\tilde{b}_2 - t_{1-\alpha/2, n-3} S_{jack}, \tilde{b}_2 + t_{1-\alpha/2, n-3} S_{jack} \right], \quad (2.32)$$

gde je $\tilde{b}_{2_i} = nb_2 - (n-1)b_{2(i)}$, $\tilde{b}_2 = \sum_{i=1}^n \tilde{b}_{2_i} / n$, $S_{jack} = \left\{ \sum_{i=1}^n (\tilde{b}_{2_i} - \tilde{b}_2)^2 / [(n-1)n] \right\}^{1/2}$ i $t_{1-\alpha/2}$ je $1-\alpha/2$ percentil t raspodele sa $n-3$ stepeni slobode.

2.3.3 Edgeworth-ov razvoj raspodele t -statistike

U cilju transformisanja t -statistike koja se koristi za konstrukciju intervalnih ocena za regresioni koeficijent β_2 , izvešće se Edgeworth-ov razvoj raspodele t -statistike (2.25) do reda n^{-1} . Na osnovu razvoja predložiće se nova transformacija t -statistike, koja je upotrebljena za konstrukciju *bootstrap-t* intervala.

Teorema 2. Pod odgovarajućim uslovima regularnosti²⁰ raspodela t -statistike date izrazom (2.25) ima sledeći razvoj:

²⁰Analogno Hall (1992b, str.71) uslovi regularnosti su:

1) Cramer-ov uslov $\limsup_{|t_1|+|t_2| \rightarrow \infty} \left| E \exp(it_1 X_j + it_2 X_j^2) \right| < 1$ i 2) $E(X_i^9) < \infty$ gde je $X_{i,j}'$ definisano u (2.34).

$$P(T_2 \leq x) = \Phi(x) + \frac{1}{6} \frac{1}{n-1} M_3' (2x^2 + 1) \phi(x) + O(n^{-2}), \quad (2.33)$$

gde su $\phi(\cdot)$ i $\Phi(\cdot)$ funkcije gustine i raspodele standardizovane normalne promenljive. Na osnovu formule $M_3' = E[(1/n(n-1)) \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n X_{i,j}^3]$ izračunat je treći momenat M_3' , a slučajne promenljive

$X_{i,j}'$, $i=1,2,\dots,n$; $j=1,2,\dots,n$ predstavljene su sledećim oblikom:

$$X_{i,j}' = \frac{(X_i - \bar{X})(Y_i - \bar{Y})(X_j - \bar{X})^2 (X_i - X_j)}{\frac{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (X_i^2 - \bar{X}^2)^2 - \left(\sum_{i=1}^n (X_i - \bar{X})(X_i^2 - \bar{X}^2)\right)^2}{n(n-1)}} - \frac{\beta_2}{\sqrt{V}}, \quad (2.34)$$

gde je

$$\bar{X} = (1/n) \sum_{i=1}^n X_i, \quad \bar{Y} = (1/n) \sum_{i=1}^n Y_i$$

i

$$V = E \left\{ \left[\frac{(X_i - \bar{X})(Y_i - \bar{Y})(X_j - \bar{X})^2 (X_i - X_j)}{\frac{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (X_i^2 - \bar{X}^2)^2 - \left(\sum_{i=1}^n (X_i - \bar{X})(X_i^2 - \bar{X}^2)\right)^2}{n(n-1)}} - \frac{\beta_2}{\sqrt{V}} \right]^2 \right\}.$$

Dokaz. Kako je ocena parametra β_2 oblika:

$$b_2 = \frac{\sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n (X_i - \bar{X})(Y_i - \bar{Y})(X_j - \bar{X})^2 (X_i - X_j)}{\frac{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (X_i^2 - \bar{X}^2)^2 - \left(\sum_{i=1}^n (X_i - \bar{X})(X_i^2 - \bar{X}^2)\right)^2}{n(n-1)}},$$

i $E(b_2) = \beta_2$ sledi da je:

$$E \left(\frac{(X_i - \bar{X})(Y_i - \bar{Y})(X_j - \bar{X})^2 (X_i - X_j)}{\frac{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (X_i^2 - \bar{X}^2)^2 - \left(\sum_{i=1}^n (X_i - \bar{X})(X_i^2 - \bar{X}^2)\right)^2}{n(n-1)}} \right) = \frac{\beta_2}{n(n-1)}, i=1,\dots,n; j=1,\dots,n, i \neq j.$$

U nastavku će biti izvršena standardizacija slučajnih promenljivih:

$$\frac{(X_i - \bar{X})(Y_i - \bar{Y})(X_j - \bar{X})^2(X_i - X_j)}{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (X_i^2 - \bar{X}^2)^2 - \left(\sum_{i=1}^n (X_i - \bar{X})(X_i^2 - \bar{X}^2) \right)^2}, i=1, \dots, n; j=1, \dots, n, i \neq j.$$

Njihova očekivana vrednost je $\frac{\beta_2}{n(n-1)}$, dok je varijansa

$$V = E \left\{ \left[\frac{(X_i - \bar{X})(Y_i - \bar{Y})(X_j - \bar{X})^2(X_i - X_j)}{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (X_i^2 - \bar{X}^2)^2 - \left(\sum_{i=1}^n (X_i - \bar{X})(X_i^2 - \bar{X}^2) \right)^2} - \frac{\beta_2}{n(n-1)} \right]^2 \right\}.$$

Novo standardizovane promenljive biće označene sa $X'_{i,j}$ ($i=1,2,\dots,n; j=1,2,\dots,n$) i mogu se predstaviti sledećim oblikom:

$$X'_{i,j} = \frac{\frac{(X_i - \bar{X})(Y_i - \bar{Y})(X_j - \bar{X})^2(X_i - X_j)}{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (X_i^2 - \bar{X}^2)^2 - \left(\sum_{i=1}^n (X_i - \bar{X})(X_i^2 - \bar{X}^2) \right)^2} - \frac{\beta_2}{n(n-1)}}{\sqrt{V}}, i=1,2,\dots,n; j=1,2,\dots,n, i \neq j. \quad (2.35)$$

Pošto su ove promenljive dobijene standardizacijom, za njih važi da je $E(X'_{i,j}) = 0$ i $Var(X'_{i,j}) = 1$. Prosek promenljivih (2.35) je sledećeg oblika:

$$\frac{1}{n(n-1)} \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n X'_{i,j} = \frac{\frac{1}{n(n-1)} \frac{\sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n (X_i - \bar{X})(Y_i - \bar{Y})(X_j - \bar{X})^2(X_i - X_j)}{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (X_i^2 - \bar{X}^2)^2 - \left(\sum_{i=1}^n (X_i - \bar{X})(X_i^2 - \bar{X}^2) \right)^2} - \frac{1}{n(n-1)} \beta_2}{\sqrt{V}}, \quad (2.36)$$

a u nastavku će biti označen sa Z_1 . Rešavajući jednakost (2.36) po oceni b_2 dobija se:

$$\begin{aligned}
b_2 &= \frac{\sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n (X_i - \bar{X})(Y_i - \bar{Y})(X_j - \bar{X})^2 (X_i - X_j)}{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (X_i^2 - \bar{X}^2)^2 - \left(\sum_{i=1}^n (X_i - \bar{X})(X_i^2 - \bar{X}^2) \right)^2} \\
&= n(n-1)(Z_1 \sqrt{V} + \frac{1}{n(n-1)} \beta_2) = Z_1 \sqrt{V} n(n-1) + \beta_2.
\end{aligned} \tag{2.37}$$

Iz (2.37) sledi da je varijansa ocene b_2 jednaka:

$$\begin{aligned}
\text{Var}(b_2) &= \text{Var}(Z_1 \sqrt{V} n(n-1) + \beta_2) = n^2 (n-1)^2 V \text{Var}(Z_1) \\
&= n^2 (n-1)^2 V \text{Var}\left(\frac{1}{n(n-1)} \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n X'_{i,j}\right) = n(n-1) V \text{Var}(X'_{i,j}),
\end{aligned}$$

odnosno njena standardna greška je oblika $\sqrt{n(n-1)V \text{Var}(X'_{i,j})}$. Ocena standardne greške ocene b_2 određuje se na sledeći način:

$$S_{b_2} = \sqrt{n(n-1)V \text{Var}(X'_{i,j})}. \tag{2.38}$$

Koristeći formulu za varijansu uzorka dobija se

$$S_{b_2} = \sqrt{n(n-1)} \sqrt{V} \sqrt{\frac{\left(\sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n X'_{i,j} \right)^2}{\sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n X'^2_{i,j} - \frac{(\sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n X'_{i,j})^2}{n(n-1)}}} - 1} = \sqrt{\frac{1}{n(n-1)-1}} n(n-1) \sqrt{V} \sqrt{Z_2 - Z_1^2}, \tag{2.39}$$

gde je $Z_2 = \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n X'^2_{i,j}$.

Zamenom (2.37) i (2.39) u (2.25), sledi da se statistika (2.25) može predstaviti formulom:

$$T_2 = \frac{Z_1 \sqrt{V} n(n-1) + \beta_2 - \beta_2}{\sqrt{\frac{1}{n(n-1)-1}} n(n-1) \sqrt{V} \sqrt{Z_2 - Z_1^2}} = \sqrt{n(n-1)-1} \frac{Z_1}{\sqrt{Z_2 - Z_1^2}}. \tag{2.40}$$

Dalje, definiše se statistika W_n :

$$W_n = \sqrt{n(n-1)-1} \left(\frac{\partial g}{\partial Z_1}(U) \cdot (Z_1 - U_1) + \frac{\partial g}{\partial Z_2}(U) \cdot (Z_2 - U_2) + \right. \\ \left. + \frac{1}{2} \left[\frac{\partial^2 g}{\partial Z_1^2}(U) \cdot (Z_1 - U_1)^2 + 2 \frac{\partial^2 g}{\partial Z_1 \partial Z_2}(U) \cdot (Z_1 - U_1)(Z_2 - U_2) + \frac{\partial^2 g}{\partial Z_2^2}(U) \cdot (Z_2 - U_2)^2 \right] \right), \quad (2.41)$$

gde je $g(Z) = \frac{Z_1}{\sqrt{Z_2 - Z_1^2}}$ ²¹ i $E(Z_1, Z_2) \equiv U \equiv (U_1, U_2) = (0, 1)$. Izračunavanjem izvoda u izrazu (2.41) dobija se sledeće:

$$\frac{\partial g}{\partial Z_1}(0, 1) = (Z_2 - Z_1^2)^{-1/2} + Z_1^2 (Z_2 - Z_1^2)^{-3/2} \Big|_{Z_1=0, Z_2=1} = 1,$$

$$\frac{\partial g}{\partial Z_2}(0, 1) = -\frac{1}{2} Z_1 (Z_2 - Z_1^2)^{-3/2} \Big|_{Z_1=0, Z_2=1} = 0,$$

$$\frac{\partial^2 g}{\partial Z_1^2}(0, 1) = 3Z_1 (Z_2 - Z_1^2)^{-3/2} + 3Z_1^3 (Z_2 - Z_1^2)^{-5/2} \Big|_{Z_1=0, Z_2=1} = 0,$$

$$\frac{\partial^2 g}{\partial Z_1 \partial Z_2}(0, 1) = -\frac{1}{2} (Z_2 - Z_1^2)^{-3/2} - \frac{3}{2} Z_1^2 (Z_2 - Z_1^2)^{-5/2} \Big|_{Z_1=0, Z_2=1} = -\frac{1}{2}$$

i

$$\frac{\partial^2 g}{\partial Z_2^2}(0, 1) = \frac{3}{4} Z_1 (Z_2 - Z_1^2)^{-5/2} \Big|_{Z_1=0, Z_2=1} = 0.$$

Zamenom dobijenih vrednosti u (2.41) statistika W_n se može izraziti na sledeći način:

$$W_n = \sqrt{n(n-1)-1} \left(\frac{3}{2} Z_1 - \frac{1}{2} Z_1 Z_2 \right),$$

gde je $\sqrt{n(n-1)-1} = \sqrt{n(n-1)} - \frac{1}{2\sqrt{n(n-1)}} + o(n^{-1})$.

Zatim, potrebno je izračunati prva tri momenta statistike W_n .

Momenti se određuju na analogan način kao u dokazu Teoreme 1, samo što umesto n treba da stoji $n(n-1)$ i jednaki su:

²¹ Pretpostavlja se da je $g(Z)$ dva puta diferencijabilna u okolini tačke U .

$$EW_n = -\frac{1}{2} \frac{1}{\sqrt{n(n-1)}} M_3' + O(n^{-2}),$$

$$EW_n^2 = 1 + O(n^{-2}) \text{ i}$$

$$EW_n^3 = -\frac{7}{2} \frac{1}{\sqrt{n(n-1)}} M_3' + O(n^{-2}).$$

Koristeći izračunate momente određuju se kumulante statistike W_n na analogan način kao u dokazu Teoreme 1 i jednaki su:

$$K_{1N} = EW_n = -\frac{1}{2} \frac{1}{\sqrt{n(n-1)}} M_3' + O(n^{-2}),$$

$$K_{2N} = 1 + O(n^{-2}) \text{ i}$$

$$K_{3N} = -\frac{2}{\sqrt{n(n-1)}} M_3' + O(n^{-2}).$$

U članu $O(n^{-2})$ eksponent je -2 , jer se izvodi razvoj raspodele t -statistike do reda n^{-1} . Zatim se na osnovu izračunatih kumulanti, koristeći Tejlorov razvoj, određuje karakteristična funkcija statistike W_n :

$$\begin{aligned} \psi_n(t) &= \exp \left\{ K_{1N}(it) + K_{2N} \frac{(it)^2}{2} + K_{3N} \frac{(it)^3}{6} + \dots \right\} \\ &= \exp \left(-\frac{t^2}{2} \right) \left\{ 1 - \frac{1}{2} \frac{1}{\sqrt{n(n-1)}} M_3'(it) - \frac{2}{\sqrt{n(n-1)}} M_3' \frac{(it)^3}{6} + O(n^{-2}) \right\} \\ &= \exp \left(-\frac{t^2}{2} \right) \left\{ 1 + \frac{1}{\sqrt{n(n-1)}} M_3' \left(-\frac{1}{2}(it) - \frac{1}{3}(it)^3 \right) + O(n^{-2}) \right\}. \end{aligned}$$

Dalje, poznato je da važi:

$$e^{-t^2/2} = \int_{-\infty}^{+\infty} e^{itx} \phi(x) dx,$$

gde je $\phi(x)$ funkcija gustine standardizovane normalne promenljive, kao i da je (ako postoji) funkcija gustine slučajne promenljive W_n jednaka:

$$f(x) = \frac{1}{2\pi} \lim_{T \rightarrow +\infty} \int_{-T}^{+T} e^{-itx} \psi_n(t) dt,$$

odnosno:

$$\begin{aligned}
 f(x) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} e^{-\frac{t^2}{2}} \left(1 - \frac{1}{2} \frac{1}{\sqrt{n(n-1)}} M_3' it - \frac{1}{3} \frac{1}{\sqrt{n(n-1)}} M_3' (it)^3\right) dt = \\
 &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} e^{-\frac{t^2}{2}} dt - \frac{1}{4\pi} \frac{1}{\sqrt{n(n-1)}} M_3' \int_{-\infty}^{+\infty} e^{-itx} e^{-\frac{t^2}{2}} it dt - \frac{1}{6\pi} \frac{1}{\sqrt{n(n-1)}} M_3' \int_{-\infty}^{+\infty} e^{-itx} e^{-\frac{t^2}{2}} (it)^3 dt \\
 &= \phi(x) - \frac{1}{4\pi} \frac{1}{\sqrt{n(n-1)}} M_3' \int_{-\infty}^{+\infty} e^{-itx} e^{-\frac{t^2}{2}} it dt - \frac{1}{6\pi} \frac{1}{\sqrt{n(n-1)}} M_3' \int_{-\infty}^{+\infty} e^{-itx} e^{-\frac{t^2}{2}} (it)^3 dt.
 \end{aligned}$$

Integral $\int_{-\infty}^{+\infty} e^{-itx} e^{-\frac{t^2}{2}} it dt$ se izračunava na sledeći način:

$$\begin{aligned}
 \int_{-\infty}^{+\infty} e^{-itx} e^{-\frac{t^2}{2}} it dt &= \int_{-\infty}^{+\infty} e^{-\left(\frac{t^2}{2} + 2 \frac{t}{\sqrt{2}} \frac{ix}{\sqrt{2}} + \frac{i^2 x^2}{2}\right)} e^{-\frac{x^2}{2}} it dt = e^{-\frac{x^2}{2}} \int_{-\infty}^{+\infty} e^{-\left(\frac{t}{\sqrt{2}} + \frac{ix}{\sqrt{2}}\right)^2} it dt = e^{-\frac{x^2}{2}} \int_{-\infty}^{+\infty} e^{-\frac{(t+ix)^2}{2}} it dt = \\
 &= \left\{ \begin{array}{l} \text{smena} \\ t+ix = u \end{array} \right\} = e^{-\frac{x^2}{2}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} i(u-ix) du = \sqrt{2\pi} \phi(x) \left(i \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} u du + x \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} du \right) = \\
 &2\pi \phi(x) \left(i \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} u du + x \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} du \right) = 2\pi \phi(x) (0 + x) = 2\pi x \phi(x).
 \end{aligned}$$

Koristeći analogan postupak izračunava se i sledeći integral:

$$\begin{aligned}
 \int_{-\infty}^{+\infty} e^{-itx} e^{-\frac{t^2}{2}} (it)^3 dt &= \int_{-\infty}^{+\infty} e^{-\left(\frac{t^2}{2} + 2 \frac{t}{\sqrt{2}} \frac{ix}{\sqrt{2}} + \frac{i^2 x^2}{2}\right)} e^{-\frac{x^2}{2}} (it)^3 dt = e^{-\frac{x^2}{2}} \int_{-\infty}^{+\infty} e^{-\left(\frac{t}{\sqrt{2}} + \frac{ix}{\sqrt{2}}\right)^2} (it)^3 dt = e^{-\frac{x^2}{2}} \int_{-\infty}^{+\infty} e^{-\frac{(t+ix)^2}{2}} (it)^3 dt = \\
 &= \left\{ \begin{array}{l} \text{smena} \\ t+ix = u \end{array} \right\} = e^{-\frac{x^2}{2}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} i(ix-u)^3 du = \\
 &= \sqrt{2\pi} \phi(x) i \left(\int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} (-ix^3) du + 3ix^2 \int_{-\infty}^{+\infty} u e^{-\frac{u^2}{2}} du + 3ix \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} u^2 du - \int_{-\infty}^{+\infty} u^3 e^{-\frac{u^2}{2}} du \right) = \\
 &= 2\pi \phi(x) i \left(-ix^3 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} du + 0 + 3ix \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} u^2 du - 0 \right) = 2\pi \phi(x) i (-ix^3 + 3ix) = 2\pi \phi(x) (x^3 - 3x).
 \end{aligned}$$

Sledi da je funkcija gustine jednaka:

$$f(x) = \phi(x) - \frac{1}{2} \frac{1}{\sqrt{n(n-1)}} M_3' x \phi(x) - \frac{1}{3} \frac{1}{\sqrt{n(n-1)}} (x^3 - 3x) M_3' \phi(x) + O(n^{-2}). \quad (2.42)$$

Na osnovu izraza (2.42) može se izračunati

$$P(W_n \leq x) = \int_{-\infty}^x \phi(t) dt + \frac{1}{2} \frac{1}{\sqrt{n(n-1)}} M_3' \int_{-\infty}^x t \phi(t) dt - \frac{1}{3} \frac{1}{\sqrt{n(n-1)}} M_3' \int_{-\infty}^x t^3 \phi(t) dt + O(n^{-2}).$$

Integral $\int_{-\infty}^x t \phi(t) dt$ izračunava se na sledeći način:

$$\int_{-\infty}^x t \phi(t) dt = \int_{-\infty}^x t \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \left. \begin{array}{l} \text{smena} \\ -\frac{t^2}{2} = u \\ -tdt = du \end{array} \right\} = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\frac{x^2}{2}} e^u du = -\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = -\phi(x).$$

Prema analognom postupku izračunava se i sledeći integral:

$$\begin{aligned} \int_{-\infty}^x t^3 \phi(t) dt &= \int_{-\infty}^x t^3 \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \left. \begin{array}{l} \text{smena} \\ -\frac{t^2}{2} = u \\ -tdt = du \end{array} \right\} = 2 \int_{-\infty}^{-\frac{x^2}{2}} u \frac{1}{\sqrt{2\pi}} e^u du = \frac{2}{\sqrt{2\pi}} \int_{-\infty}^{-\frac{x^2}{2}} u e^u du = \\ &= \frac{2}{\sqrt{2\pi}} (u e^u \Big|_{-\infty}^{-\frac{x^2}{2}} - \int_{-\infty}^{-\frac{x^2}{2}} e^u du) = \frac{2}{\sqrt{2\pi}} \left(-\frac{x^2}{2} e^{-\frac{x^2}{2}} - e^{-u} \Big|_{-\infty}^{-\frac{x^2}{2}} \right) = \frac{2}{\sqrt{2\pi}} \left(-\frac{x^2}{2} e^{-\frac{x^2}{2}} - e^{-\frac{x^2}{2}} \right) = \frac{1}{\sqrt{2\pi}} (-x^2 - 2) e^{-\frac{x^2}{2}} \\ &= (-x^2 - 2) \phi(x). \end{aligned}$$

Sledi da je

$$P(W_n \leq x) = \Phi(x) + \frac{1}{\sqrt{n(n-1)}} \frac{M_3'}{6} (2x^2 + 1) \phi(x) + O(n^{-2}). \quad (2.43)$$

Pošto važi $T_2 = W_n + O(n^{-2})$ sledi da je Edgeworth-ov razvoj raspodele t -statistike dat izrazom (2.33). \square

Kako bi se korigovala asimetrija raspodele, uvešće se kubna transformacija t -statistike, koja je monotona i ima eksplicitnu inverznu formulu. Nova transformacija, zasnovana na Edgeworth-ovom razvoju (2.33), definisana je sledećom formulom²²:

$$T_2^* = T_2^*(U) = U + \frac{1}{2} U^2 + \frac{1}{12} U^3 + \frac{1}{6} \frac{1}{n(n-1)} \hat{M}_3', \quad (2.44)$$

gde je \hat{M}_3' ocena momenta M_3' , a $U = T_2 / \sqrt{n(n-1)}$.

Transformacija T_2^* zadovoljava sledeće uslove (Hoaglin, 1985):

²² Ova transformacija uvedena je na ekvivalentan način kao transformacija u radu Ivković et al. (2020).

- $T_{transf}(U) - \frac{1}{6} \frac{1}{n(n-1)} \hat{M}'_3$ je približno U za U blisko nuli,
- $T_{transf}(U) - \frac{1}{6} \frac{1}{n(n-1)} \hat{M}'_3$ je jednako nuli za U jednako nuli.

Njena inverzna funkcija je oblika:

$$T_2^{*-1}(t) = \left[8 + 12 \left(t - \frac{1}{6} \frac{1}{n(n-1)} \hat{M}'_3 \right) \right]^{-1/3} - 2, \quad (2.45)$$

i na osnovu nje se može konstruisati $(1-\alpha)100\%$ *bootstrap-t* interval poverenja za regresioni koeficijent β_2 :

$$I_{\beta_2} = \left[\mathbf{b}_2 - \sqrt{n(n-1)} T_2^{*-1} \left(\frac{\hat{t}^{(1-\alpha/2)}}{\sqrt{n(n-1)}} \right) S_{b_2}, \mathbf{b}_2 - \sqrt{n(n-1)} T_2^{*-1} \left(\frac{\hat{t}^{(\alpha/2)}}{\sqrt{n(n-1)}} \right) S_{b_2} \right]. \quad (2.46)$$

Interval poverenja za regresioni koeficijent β_2 zasnovan na transformaciji T_2^* u daljem radu označen je kao T_2^* interval.

2.4 Analiza pouzdanosti intervala poverenja za regresione koeficijente

U cilju ispitivanja pouzdanosti intervalnih ocena analiziraće se proporcije simuliranih intervala za regresione koeficijente u prostom linearnom i kvadratnom regresionom modelu, koji su uvedeni u odeljcima 2.2 i 2.3, u situacijama kada slučajna greška u regresionom modelu ne sledi normalnu raspodelu.

2.4.1 Analiza pouzdanosti intervala poverenja za regresioni koeficijent u prostom linearnom regresionom modelu

Predmet istraživanja ovog odeljka je poređenje proporcija simuliranih intervala (za regresioni koeficijent) koji sadrže nepoznati parametar u prostom linearnom regresionom modelu kada slučajna greška nije normalno raspodeljena. U svrhe istraživanja razmatrana su dva slučaja. Predmet analize u prvom slučaju biće eksponencijalna raspodela, dok će u drugom slučaju biti analizirana Weibull-ova raspodela. Predmet razmatranja može biti bilo koja asimetrična raspodela.

U svakom slučaju sprovedeće se ista procedura. Nominalni nivo pouzdanosti biće 95%. Poći će se od prostog linearnog regresionog modela sa regresionim koeficijentima $\beta_0 = 150$ i $\beta_1 = 20$. Za svaku kombinaciju parametara, izabraće se po 1000 uzoraka veličine 10, 20, 50 i 80, na osnovu kojih će se formirati t -intervali. U cilju konstruisanja *bootstrap-t* intervala i intervala poverenja zasnovanih na transformisanim statistikama kombinovanim sa *bootstrap* metodom, izabraće se 1000 *bootstrap* uzoraka iz svakog uzorka.

Analiza je sprovedena u programskom jeziku R, koristeći originalne kodove. U nastavku slede rezultati do kojih se došlo.

U prvom slučaju ispitivane su proporcije simuliranih intervala ukoliko je raspodela slučajne greške eksponencijalna²³ sa parametrom λ koji uzima vrednosti 0,2 i 2. Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_1 prikazane su u Tabeli 2.1. Uočava se da su za male uzorke (veliĉine 10), u zavisnosti od parametra λ eksponencijalne raspodele, proporcije simuliranih *bootstrap-t* intervala bile najbliže 0,95, dok su za uzorke veliĉine 20 najbolji izbor *bootstrap-t* i T_3 interval. Za uzorke veliĉine 50 najbolji rezultati su dobijeni primenom intervala T_2 i T_3 , dok su za uzorke veliĉine 80 najbolje rezultate dali intervali T_1 i T_2 .

Tabela 2.1: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent kada slučajna greška ima eksponencijalnu raspodelu

λ	0,2				2			
n	10	20	50	80	10	20	50	80
t	0,924	0,915	0,901	0,849	0,937	0,923	0,856	0,908
<i>Bootstrap-t</i>	0,948	0,912	0,924	0,925	0,952	0,941	0,927	0,938
T_1	0,790	0,917	0,930	0,942	0,829	0,918	0,932	0,944
T_2	0,940	0,930	0,914	0,945	0,958	0,960	0,934	0,967
T_3	0,624	0,940	0,943	0,957	0,584	0,935	0,967	0,963

Izvor: Preuzeto iz Ivković et al. (2020)

Predmet razmatranja u drugom slučaju su bile proporcije simuliranih intervala pod pretpostavkom da je raspodela slučajne greške Weibull-ova²⁴ sa parametrom a koji uzima vrednosti 1,5 i 3, i parametrom $b=1$. U Tabeli 2.2 predstavljene su proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_1 . Za uzorke veliĉine 10, u zavisnosti od parametra a Weibull-ove raspodele, proporcije simuliranih *bootstrap-t* i T_2 intervala bile su najbliže 0,95, dok su za uzorke veliĉine 20 najbolji rezultati ostvareni primenom T_2 i T_3 intervala. Kada je reĉ o velikim uzorcima (veliĉine 50) najbolji rezultati su dobijeni primenom *bootstrap-t*, T_1 i T_2 intervala, dok su za uzorke veliĉine 80, proporcije simuliranih *bootstrap-t* i T_3 intervala bile najbliže 0,95.

²³ Funkcija gustine eksponencijalne raspodele nalazi se u odeljku 2.1.2.

²⁴ Funkcija gustine Weibull-ove raspodele nalazi se u odeljku 2.1.1.

Tabela 2.2: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent kada slučajna greška ima Weibull-ovu raspodelu

a	1,5				3			
b	1				1			
n	10	20	50	80	10	20	50	80
t	0,939	0,923	0,890	0,871	0,919	0,900	0,914	0,881
Bootstrap-t	0,942	0,921	0,921	0,955	0,956	0,929	0,949	0,941
T_1	0,815	0,935	0,940	0,866	0,804	0,901	0,949	0,963
T_2	0,950	0,920	0,942	0,962	0,959	0,942	0,939	0,961
T_3	0,563	0,943	0,966	0,959	0,574	0,932	0,953	0,956

Izvor: Preuzeto iz Ivković et al. (2020)

Na osnovu prethodnih rezultata izveden je generalni zaključak za eksponencijalnu i Weibull-ovu raspodelu. Kada je reč o malim uzorcima, proporcije simuliranih *bootstrap-t* intervala i *bootstrap-t* intervala sa transformacijama T_2 i T_3 bile su najbliže 0,95. Prilikom razmatranja velikih uzoraka, utvrđeno je da su, u većini slučajeva, proporcije simuliranih *bootstrap-t* intervala sa transformacijama T_2 i T_3 bile najbliže nominalnom nivou pouzdanosti.

2.4.2 Analiza pouzdanosti intervala poverenja za regresione koeficijente u kvadratnom regresionom modelu

U situacijama kada slučajna greška sledi asimetričnu raspodelu analiziraće se proporcije simuliranih intervala za regresione koeficijente u kvadratnom regresionom modelu. Analiziraće se tri slučaja. U prvom slučaju predmet interesovanja biće gama raspodela, zatim će se razmatrati Weibull-ova raspodela, dok će u trećem slučaju biti ispitivana eksponencijalna raspodela. Potrebno je naglasiti da predmet razmatranja može biti bilo koja asimetrična raspodela.

U svakom slučaju sprovedeće se ista procedura. Razmatraće se nominalni nivoi pouzdanosti od 95% i 90%. Poći će se od kvadratnog regresionog modela sa regresionim koeficijentima $\beta_0 = 100$, $\beta_1 = 10$ i $\beta_2 = 2$. Za svaku kombinaciju parametara, izabraće se po 1000 uzoraka veličine 10, 20, 50 i 80, na osnovu kojih će se konstruisani t -intervali. U cilju formiranja intervala poverenja zasnovanih na metodama ponovljenih uzoraka i intervala zasnovanih na transformisanoj statistici (ovi intervali konstruišu se samo za regresioni koeficijent β_2), generisaće se 1000 *bootstrap* uzoraka iz svakog uzorka.

Za sve razmatrane intervale napisani su originalni kodovi u programskom jeziku R. U nastavku slede rezultati do kojih se došlo.

U prvom slučaju analizirane su proporcije simuliranih intervala ukoliko je raspodela slučajne greške gama²⁵ sa parametrom $a=2$ i parametrom b koji uzima vrednosti 0,5; 1; 2,5 i 5. Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_1 prikazane su u Tabeli 2.3. Analiza rezultata pokazala je sledeće: za male uzorke (veličine 10), proporcije simuliranih BC i

²⁵ Funkcija gustine gama raspodele nalazi se u odeljku 2.1.3.

percentil intervala bile su najbliže 0,95, dok su za uzorke veličine 20 najbolji rezultati dobijeni primenom *jackknife* i percentil intervala; za velike uzorke (veličine 50), ustanovljeno je da su u zavisnosti od parametra b gama raspodele, najbolji rezultati dobijeni primenom percentil, *bootstrap-t* i BC intervala, a za uzorke veličine 80 kao najbolji izbor su se pokazali percentil, BC, *jackknife* i *bootstrap-t* interval.

Tabela 2.3: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_1 kada slučajna greška ima gama raspodelu

a	b	n	t	<i>Bootstrap-t</i>	BC	BCa	Percentil	<i>Jackknife</i>
2	0,5	10	0,931	0,972	0,962	0,976	0,954	0,963
		20	0,908	0,941	0,934	0,934	0,932	0,950
		50	0,919	0,943	0,842	0,923	0,916	0,963
		80	0,985	0,937	0,925	0,939	0,943	0,957
2	1	10	0,919	0,977	0,947	0,970	0,965	0,954
		20	0,915	0,938	0,919	0,920	0,956	0,944
		50	0,924	0,951	0,940	0,928	0,952	0,912
		80	0,982	0,966	0,940	0,927	0,958	0,943
2	2,5	10	0,921	0,972	0,957	0,967	0,966	0,960
		20	0,972	0,930	0,933	0,942	0,933	0,953
		50	0,848	0,938	0,955	0,940	0,941	0,913
		80	0,859	0,966	0,950	0,929	0,942	0,810
2	5	10	0,925	0,982	0,956	0,967	0,961	0,967
		20	0,920	0,933	0,959	0,932	0,935	0,948
		50	0,925	0,940	0,940	0,933	0,946	0,967
		80	0,911	0,949	0,959	0,856	0,935	0,946

Izvor: Izlaz iz programskog jezika R

Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_2 prikazane su u Tabeli 2.4. Izvedeni su sledeći zaključci: kada je reč o uzorcima veličine 10, intervali BC, T_2^* i *jackknife* su dali najbolje rezultate; za uzorke veličine 20 kao najbolji su se pokazali intervali T_2^* , BC i *bootstrap-t*; kod uzoraka veličine 50, u zavisnosti od parametra b , proporcije simuliranih T_2^* , *bootstrap-t*, percentil i BC intervala su bile najbolje, dok su za uzorke veličine 80, intervali T_2^* i percentil obezbedili najbolje rezultate.

Tabela 2.4: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_2 kada slučajna greška ima gama raspodelu

a	b	n	t	<i>Bootstrap-t</i>	BC	BCa	Percentil	<i>Jackknife</i>	T_2^*
2	0,5	10	0,933	0,976	0,966	0,980	0,963	0,959	0,960
		20	0,908	0,942	0,938	0,931	0,938	0,974	0,951
		50	0,906	0,947	0,889	0,922	0,906	0,961	0,939
		80	0,986	0,952	0,932	0,927	0,944	0,968	0,950
2	1	10	0,926	0,983	0,961	0,969	0,968	0,959	0,955
		20	0,913	0,944	0,923	0,939	0,957	0,969	0,957
		50	0,927	0,953	0,937	0,913	0,957	0,914	0,948
		80	0,979	0,953	0,939	0,918	0,949	0,929	0,955
2	2,5	10	0,925	0,968	0,960	0,970	0,966	0,964	0,966
		20	0,972	0,929	0,940	0,943	0,938	0,957	0,951
		50	0,841	0,943	0,954	0,942	0,949	0,917	0,933
		80	0,857	0,967	0,955	0,934	0,938	0,861	0,954
2	5	10	0,929	0,984	0,959	0,966	0,969	0,969	0,966
		20	0,907	0,948	0,951	0,944	0,947	0,955	0,941
		50	0,921	0,911	0,949	0,942	0,952	0,970	0,930
		80	0,925	0,956	0,960	0,853	0,936	0,938	0,950

Izvor: Izlaz iz programskog jezika R

Proporcije simuliranih 90% dvostranih intervala poverenja za regresioni koeficijent β_1 predstavljene su u Tabeli 2.5. Izvedeni su sledeći zaključci: za uzorke veličine 10, u zavisnosti od parametra b gama raspodele, percentil, *jackknife* i BC intervali su dali najbolje rezultate, dok su se za uzorke veličine 20 kao najbolji pokazali *bootstrap-t*, percentil, *jackknife* i BCa intervali; kada je reč o uzorcima veličine 50, proporcije simuliranih percentil i *jackknife* intervala su bile najbolje, a za uzorke veličine 80 su se kao najbolji izbor pokazali intervali percentil, *jackknife* i BC.

Tabela 2.5: Proporcije simuliranih 90% dvostranih intervala za regresioni koeficijent β_1 kada slučajna greška ima gama raspodelu

a	b	n	t	<i>Bootstrap-t</i>	BC	BCa	Percentil	<i>Jackknife</i>
2	0,5	10	0,925	0,945	0,918	0,930	0,915	0,905
		20	0,863	0,883	0,884	0,859	0,900	0,900
		50	0,879	0,893	0,875	0,879	0,889	0,906
		80	0,953	0,871	0,899	0,881	0,903	0,817
2	1	10	0,926	0,918	0,930	0,934	0,912	0,922
		20	0,865	0,900	0,876	0,872	0,896	0,919
		50	0,810	0,913	0,858	0,872	0,907	0,893
		80	0,946	0,888	0,903	0,881	0,854	0,894
2	2,5	10	0,841	0,927	0,906	0,927	0,932	0,917
		20	0,859	0,905	0,849	0,898	0,886	0,929
		50	0,941	0,878	0,888	0,876	0,910	0,896
		80	0,935	0,922	0,870	0,917	0,911	0,918
2	5	10	0,925	0,931	0,918	0,928	0,906	0,905
		20	0,872	0,897	0,880	0,877	0,877	0,909
		50	0,765	0,889	0,858	0,863	0,894	0,912
		80	0,936	0,922	0,921	0,834	0,888	0,901

Izvor: Izlaz iz programskog jezika R

Proporcije simuliranih 90% dvostranih intervala poverenja za regresioni koeficijent β_2 predstavljene su u Tabeli 2.6. Kod uzoraka veličine 10, percentil, T_2^* i *jackknife* intervali su obezbedili najbolje rezultate, dok su za uzorke veličine 20, u zavisnosti od parametra b , proporcije simuliranih percentil, *bootstrap-t* i BCa intervala bile najbliže 0,90; kada je reč o uzorcima veličine 50, najbolji izbor su bili T_2^* , *jackknife*, percentil i *bootstrap-t* intervali, dok su za uzorke veličine 80 najbolji rezultati ostvareni primenom T_2^* i *jackknife* intervala.

Tabela 2.6: Proporcije simuliranih 90% dvostranih intervala za regresioni koeficijent β_2 kada slučajna greška ima gama raspodelu

a	b	n	t	<i>Bootstrap-t</i>	BC	BCa	Percentil	<i>Jackknife</i>	T_2^*
2	0,5	10	0,929	0,947	0,933	0,929	0,918	0,904	0,937
		20	0,864	0,886	0,880	0,869	0,901	0,913	0,888
		50	0,865	0,880	0,886	0,886	0,851	0,934	0,904
		80	0,945	0,891	0,893	0,887	0,885	0,852	0,903
2	1	10	0,926	0,927	0,935	0,940	0,916	0,927	0,930
		20	0,860	0,905	0,877	0,870	0,897	0,919	0,890
		50	0,807	0,908	0,844	0,879	0,902	0,924	0,910
		80	0,947	0,884	0,908	0,880	0,890	0,909	0,906
2	2,5	10	0,857	0,941	0,911	0,935	0,936	0,925	0,897
		20	0,861	0,897	0,854	0,901	0,885	0,920	0,884
		50	0,933	0,872	0,885	0,884	0,916	0,900	0,890
		80	0,925	0,922	0,873	0,920	0,881	0,914	0,915
2	5	10	0,931	0,948	0,925	0,941	0,904	0,924	0,923
		20	0,867	0,899	0,875	0,879	0,891	0,929	0,891
		50	0,780	0,903	0,865	0,867	0,909	0,928	0,894
		80	0,926	0,924	0,916	0,837	0,867	0,892	0,894

Izvor: Izlaz iz programskog jezika R

Predmet proučavanja u drugom slučaju su bile proporcije simuliranih intervala pod pretpostavkom da je raspodela slučajne greške Weibull-ova²⁶ sa parametrom a koji uzima vrednosti 0,5; 1; 2 i 5, i parametrom $b=1$. U Tabeli 2.7 predstavljene su proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_1 . Izvedeni su sledeći zaključci: za uzorke veličine 10, percentil, *jackknife*, BC i BCa intervali su obezbedili najbolje rezultate, dok su za uzorke veličine 20, u zavisnosti od parametra a Weibull-ove raspodele, proporcije simuliranih percentil, *jackknife* i *bootstrap-t* intervala bile najbliže 0,95; u slučaju uzoraka veličine 50, percentil, *jackknife*, BC i BCa intervali su obezbedili najbolje rezultate, a za uzorke veličine 80 kao najbolji su se pokazali intervali *jackknife*, *bootstrap-t* i BCa.

²⁶ Funkcija gustine Weibull-ove raspodele nalazi se u odeljku 2.1.1.

Tabela 2.7: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_1 kada slučajna greška ima Weibull-ovu raspodelu

a	b	n	t	<i>Bootstrap-t</i>	BC	BCa	Percentil	<i>Jackknife</i>
0,5	1	10	0,914	0,975	0,961	0,949	0,966	0,989
		20	0,927	0,955	0,923	0,897	0,940	0,986
		50	0,992	0,962	0,896	0,857	0,946	0,973
		80	0,980	0,915	0,923	0,830	0,943	0,956
1	1	10	0,927	0,978	0,968	0,961	0,947	0,979
		20	0,913	0,943	0,928	0,930	0,931	0,959
		50	0,913	0,941	0,944	0,935	0,935	0,947
		80	0,913	0,951	0,961	0,927	0,940	0,965
2	1	10	0,932	0,966	0,962	0,973	0,970	0,947
		20	0,920	0,929	0,938	0,941	0,934	0,958
		50	0,910	0,932	0,878	0,930	0,944	0,943
		80	0,977	0,940	0,963	0,949	0,881	0,933
5	1	10	0,930	0,978	0,958	0,972	0,960	0,968
		20	0,922	0,965	0,939	0,941	0,947	0,956
		50	0,910	0,956	0,953	0,947	0,946	0,960
		80	0,905	0,955	0,946	0,960	0,937	0,951

Izvor: Izlaz iz programskog jezika R

U Tabeli 2.8 predstavljene su proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_2 . Analiza uzoraka veličine 10 pokazala je da su T_2^* , *jackknife* i BCa intervali obezbedili najbolje rezultate, dok su za uzorke veličine 20 najbolji izbor bili *bootstrap-t*, T_2^* i BCa intervali; za uzorke veličine 50, u zavisnosti od parametra a , proporcije simuliranih T_2^* i percentil intervala su bile najbliže 0,95, a za uzorke veličine 80 najbolji rezultati ostvareni su primenom T_2^* , *bootstrap-t* i BC intervala.

Tabela 2.8: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_2 kada slučajna greška ima Weibull-ovu raspodelu

a	b	n	t	<i>Bootstrap-t</i>	BC	BCa	Percentil	<i>Jackknife</i>	T_2^*
0,5	1	10	0,919	0,978	0,959	0,947	0,973	0,970	0,954
		20	0,928	0,950	0,924	0,898	0,928	0,984	0,951
		50	0,991	0,962	0,904	0,862	0,947	0,977	0,948
		80	0,980	0,932	0,929	0,881	0,923	0,973	0,942
1	1	10	0,927	0,980	0,978	0,969	0,973	0,977	0,963
		20	0,928	0,949	0,940	0,923	0,932	0,970	0,962
		50	0,923	0,922	0,944	0,934	0,937	0,963	0,945
		80	0,923	0,949	0,952	0,930	0,946	0,969	0,944
2	1	10	0,928	0,973	0,973	0,981	0,973	0,967	0,966
		20	0,921	0,936	0,941	0,948	0,935	0,964	0,933
		50	0,923	0,932	0,879	0,927	0,943	0,942	0,952
		80	0,979	0,927	0,961	0,954	0,834	0,967	0,953
5	1	10	0,932	0,983	0,973	0,971	0,966	0,963	0,975
		20	0,934	0,969	0,939	0,941	0,946	0,959	0,947
		50	0,913	0,954	0,953	0,954	0,948	0,957	0,939
		80	0,902	0,942	0,948	0,955	0,939	0,957	0,958

Izvor: Izlaz iz programskog jezika R

U Tabeli 2.9 su prikazane proporcije simuliranih 90% dvostranih intervala za regresioni koeficijent β_1 . Razmatranjem rezultata uočeno je da su za uzorke veličine 10, u zavisnosti od parametra a Weibull-ove raspodele, proporcije simuliranih percentil, *jackknife*, BC i BCa intervala bile najbliže 0,90, dok su za uzorke veličine 20 najbolji izbor bili percentil, *jackknife*, BC i *bootstrap-t* intervali. Kada je reč o uzorcima veličine 50 *jackknife* i *bootstrap-t* intervali obezbedili su najbolje rezultate, dok su za uzorke veličine 80 najbolji bili percentil, *jackknife*, BC, BCa i *bootstrap-t* intervali.

Tabela 2.9: Proporcije simuliranih 90% dvostranih intervala za regresioni koeficijent β_1 kada slučajna greška ima Weibull-ovu raspodelu

a	b	n	t	<i>Bootstrap-t</i>	BC	BCa	Percentil	<i>Jackknife</i>
0,5	1	10	0,938	0,927	0,913	0,905	0,908	0,932
		20	0,958	0,870	0,853	0,799	0,890	0,925
		50	0,945	0,725	0,830	0,815	0,885	0,909
		80	0,951	0,861	0,805	0,827	0,900	0,896
1	1	10	0,930	0,944	0,910	0,919	0,908	0,922
		20	0,865	0,893	0,874	0,858	0,879	0,912
		50	0,858	0,896	0,859	0,878	0,873	0,913
		80	0,964	0,736	0,891	0,878	0,880	0,902
2	1	10	0,922	0,910	0,905	0,927	0,925	0,918
		20	0,849	0,893	0,884	0,887	0,883	0,892
		50	0,862	0,914	0,894	0,906	0,876	0,899
		80	0,751	0,906	0,923	0,906	0,877	0,867
5	1	10	0,925	0,938	0,928	0,931	0,914	0,902
		20	0,867	0,863	0,893	0,916	0,889	0,893
		50	0,926	0,897	0,895	0,871	0,864	0,927
		80	0,942	0,922	0,904	0,812	0,852	0,916

Izvor: Izlaz iz programskog jezika R

U Tabeli 2.10 su prikazane proporcije simuliranih 90% dvostranih intervala za regresioni koeficijent β_2 . Za uzorke veličine 10 proporcije simuliranih percentil, BCa, *bootstrap-t* i *jackknife* intervala su bile najbolje, dok je analizom uzoraka veličine 20 utvrđeno da su proporcije simuliranih T_2^* , percentil i BC intervala bile najbliže 0,90. Za uzorke veličine 50, u zavisnosti od parametra a , kao najbolji su se pokazali intervali T_2^* , BC, BCa, *jackknife* i *bootstrap-t*, dok su najbolje rezultate za uzorke veličine 80 obezbedili T_2^* , percentil i *bootstrap-t* intervali.

Tabela 2.10: Proporcije simuliranih 90% dvostranih intervala za regresioni koeficijent β_2 kada slučajna greška ima Weibull-ovu raspodelu

a	b	n	t	<i>Bootstrap-t</i>	BC	BCa	Percentil	<i>Jackknife</i>	T_2^*
0,5	1	10	0,931	0,922	0,915	0,910	0,905	0,937	0,925
		20	0,954	0,886	0,863	0,822	0,898	0,931	0,897
		50	0,954	0,730	0,836	0,808	0,878	0,906	0,865
		80	0,946	0,868	0,719	0,827	0,896	0,876	0,855
1	1	10	0,925	0,943	0,933	0,917	0,926	0,926	0,921
		20	0,872	0,890	0,880	0,867	0,875	0,916	0,893
		50	0,842	0,903	0,871	0,874	0,857	0,919	0,852
		80	0,940	0,790	0,892	0,857	0,879	0,919	0,904
2	1	10	0,925	0,914	0,921	0,929	0,927	0,921	0,936
		20	0,857	0,915	0,899	0,893	0,878	0,897	0,906
		50	0,874	0,911	0,898	0,898	0,873	0,909	0,904
		80	0,725	0,903	0,925	0,914	0,876	0,832	0,917
5	1	10	0,923	0,942	0,938	0,936	0,937	0,909	0,937
		20	0,876	0,882	0,896	0,911	0,894	0,911	0,901
		50	0,929	0,916	0,896	0,884	0,876	0,919	0,900
		80	0,943	0,927	0,913	0,832	0,861	0,917	0,908

Izvor: Izlaz iz programskog jezika R

U trećem slučaju ispitivane su proporcije simuliranih intervala ukoliko je raspodela slučajne greške eksponencijalna²⁷ sa parametrom λ koji uzima vrednosti 2,5; 0,625; 0,4 i 0,25. Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_1 predstavljene su u Tabeli 2.11. Analizom rezultata uočeno je da su za uzorke veličine 10 *jackknife*, percentil, BC i BCa intervali dali najbolje rezultate, a da su za uzorke veličine 20 najbolji izbor bili *bootstrap-t* i *jackknife* intervali. Kada je reč o uzorcima veličine 50, u zavisnosti od parametra λ eksponencijalne raspodele, proporcije simuliranih *bootstrap-t* i percentil intervala su bile najbliže 0,95, dok su za uzorke veličine 80 najbolje rezultate obezbedili intervali *bootstrap-t* i BC.

²⁷ Funkcija gustine eksponencijalne raspodele nalazi se u odeljku 2.1.2.

Tabela 2.11: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_1 kada slučajna greška ima eksponencijalnu raspodelu

λ	n	t	<i>Bootstrap-t</i>	BC	BCa	Percentil	<i>Jackknife</i>
2,5	10	0,915	0,976	0,957	0,957	0,967	0,967
	20	0,903	0,953	0,901	0,938	0,934	0,963
	50	0,838	0,913	0,904	0,903	0,950	0,959
	80	0,971	0,946	0,948	0,935	0,870	0,929
0,625	10	0,909	0,970	0,971	0,966	0,946	0,959
	20	0,926	0,945	0,929	0,925	0,929	0,958
	50	0,926	0,950	0,899	0,902	0,922	0,948
	80	0,978	0,955	0,952	0,928	0,946	0,943
0,4	10	0,925	0,977	0,974	0,967	0,962	0,956
	20	0,922	0,942	0,907	0,921	0,938	0,952
	50	0,891	0,940	0,900	0,921	0,947	0,906
	80	0,979	0,945	0,892	0,922	0,900	0,942
0,25	10	0,919	0,982	0,954	0,969	0,969	0,955
	20	0,917	0,943	0,934	0,932	0,936	0,959
	50	0,908	0,950	0,933	0,936	0,943	0,959
	80	0,810	0,953	0,942	0,927	0,846	0,969

Izvor: Izlaz iz programskog jezika R

Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_2 predstavljene su u Tabeli 2.12. Izvedeni su sledeći zaključci: za uzorke veličine 10, *jackknife*, T_2^* i BC intervali su obezbedili najbolje rezultate, a za uzorke veličine 20 najbolji izbor su bili T_2^* i *bootstrap-t* intervali. Analiza uzoraka veličine 50 pokazala je da su u zavisnosti od parametra λ , proporcije simuliranih T_2^* , *bootstrap-t* i percentil intervala bile najbolje, dok su za uzorke veličine 80 najbolji izbor bili intervali T_2^* , percentil i BCa.

Tabela 2.12: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_2 kada slučajna greška ima eksponencijalnu raspodelu

λ	n	t	<i>Bootstrap-t</i>	BC	BCa	Percentil	<i>Jackknife</i>	T_2^*
2,5	10	0,922	0,977	0,953	0,961	0,976	0,966	0,977
	20	0,887	0,954	0,919	0,936	0,936	0,963	0,951
	50	0,842	0,927	0,920	0,917	0,930	0,963	0,948
	80	0,978	0,942	0,954	0,950	0,877	0,964	0,967
0,625	10	0,918	0,969	0,976	0,965	0,968	0,974	0,959
	20	0,924	0,950	0,945	0,927	0,927	0,955	0,975
	50	0,890	0,949	0,910	0,896	0,912	0,968	0,954
	80	0,975	0,954	0,958	0,940	0,950	0,947	0,923
0,4	10	0,929	0,974	0,976	0,974	0,964	0,956	0,965
	20	0,925	0,945	0,921	0,927	0,943	0,961	0,953
	50	0,902	0,938	0,917	0,918	0,950	0,906	0,928
	80	0,978	0,952	0,886	0,928	0,919	0,954	0,949
0,25	10	0,925	0,986	0,964	0,975	0,970	0,960	0,966
	20	0,917	0,950	0,940	0,937	0,939	0,965	0,938
	50	0,911	0,941	0,942	0,943	0,941	0,969	0,944
	80	0,821	0,963	0,957	0,922	0,829	0,962	0,956

Izvor: Izlaz iz programskog jezika R

Proporcije simuliranih 90% dvostranih intervala za regresioni koeficijent β_1 prikazane su u Tabeli 2.13. Za uzorke veličine 10, percentil interval je dao najbolje rezultate, dok su se za uzorke veličine 20, u zavisnosti od parametra λ eksponencijalne raspodele, kao najbolji izbor pokazali *jackknife*, *bootstrap-t* i percentil intervali. Primećuje se da su za uzorke veličine 50 proporcije simuliranih *jackknife*, percentil i BCa intervala bile najbolje, a za uzorke veličine 80 najbolji rezultati su ostvareni primenom *jackknife*, percentil i BC intervala.

Tabela 2.13: Proporcije simuliranih 90% dvostranih intervala za regresioni koeficijent β_1 kada slučajna greška ima eksponencijalnu raspodelu

λ	n	t	<i>Bootstrap-t</i>	BC	BCa	Percentil	<i>Jackknife</i>
2,5	10	0,876	0,926	0,932	0,914	0,901	0,920
	20	0,926	0,898	0,886	0,883	0,887	0,878
	50	0,814	0,851	0,846	0,847	0,871	0,913
	80	0,946	0,892	0,774	0,872	0,797	0,906
0,625	10	0,920	0,946	0,931	0,927	0,903	0,918
	20	0,928	0,888	0,883	0,867	0,894	0,897
	50	0,855	0,889	0,885	0,900	0,901	0,899
	80	0,785	0,922	0,922	0,876	0,906	0,918
0,4	10	0,925	0,950	0,916	0,928	0,913	0,924
	20	0,932	0,878	0,864	0,860	0,888	0,912
	50	0,930	0,859	0,890	0,871	0,898	0,907
	80	0,832	0,906	0,892	0,823	0,898	0,912
0,25	10	0,871	0,938	0,927	0,932	0,908	0,916
	20	0,876	0,894	0,883	0,857	0,889	0,926
	50	0,825	0,884	0,892	0,868	0,848	0,897
	80	0,943	0,901	0,900	0,885	0,897	0,908

Izvor: Izlaz iz programskog jezika R

Proporcije simuliranih 90% dvostranih intervala za regresioni koeficijent β_2 prikazane su u Tabeli 2.14. Analiza uzoraka veličine 10 je pokazala da su percentil i *jackknife* intervali obezbedili najbolje rezultate, dok su se za uzorke veličine 20 kao najbolji izbor pokazali T_2^* , *bootstrap-t* i percentil intervali. Kod uzoraka veličine 50 je uočeno da su u zavisnosti od parametra λ , proporcije simuliranih T_2^* i percentil intervala bile najbolje, dok se za uzorke veličine 80 interval T_2^* pokazao kao najbolji.

Tabela 2.14: Proporcije simuliranih 90% dvostranih intervala za regresioni koeficijent β_2 kada slučajna greška ima eksponencijalnu raspodelu

λ	n	t	<i>Bootstrap-t</i>	BC	BCa	Percentil	<i>Jackknife</i>	T_2^*
2,5	10	0,870	0,926	0,930	0,930	0,907	0,912	0,933
	20	0,930	0,902	0,888	0,892	0,894	0,892	0,885
	50	0,848	0,864	0,844	0,819	0,880	0,909	0,892
	80	0,948	0,870	0,758	0,879	0,771	0,919	0,908
0,625	10	0,922	0,938	0,939	0,928	0,917	0,919	0,921
	20	0,934	0,896	0,894	0,861	0,898	0,905	0,875
	50	0,818	0,890	0,887	0,893	0,892	0,909	0,895
	80	0,768	0,910	0,923	0,887	0,915	0,918	0,909
0,4	10	0,922	0,945	0,929	0,928	0,920	0,920	0,934
	20	0,930	0,875	0,870	0,858	0,893	0,908	0,894
	50	0,926	0,856	0,883	0,880	0,893	0,909	0,913
	80	0,847	0,896	0,888	0,788	0,894	0,927	0,902
0,25	10	0,876	0,938	0,930	0,934	0,910	0,922	0,922
	20	0,876	0,897	0,891	0,873	0,880	0,927	0,898
	50	0,832	0,883	0,878	0,874	0,850	0,889	0,890
	80	0,937	0,890	0,891	0,866	0,894	0,922	0,901

Izvor: Izlaz iz programskog jezika R

Uzimajući u obzir prethodno predstavljene rezultate, kada je reč o gama, Weibull-ovoj i eksponencijalnoj raspodeli, izvedeni su sledeći zaključci. Kada je nominalni nivo pouzdanosti 95% za parametar β_1 utvrđeno je sledeće: u većini slučajeva su za male uzorke proporcije simuliranih BC i *jackknife* intervala bile najbliže 0,95, dok su za velike uzorke najbolji rezultati dobijeni korišćenjem percentil i *bootstrap-t* intervala. Analiza rezultata za parametar β_2 pokazala je da su u većini slučajeva za male uzorke proporcije simuliranih *bootstrap-t* intervala sa transformacijom T_2^* i *bootstrap-t* intervala bile najpribližnije nominalnom nivou pouzdanosti, dok se za velike uzorke kao najbolji pokazao *bootstrap-t* interval sa transformacijom T_2^* . Ako je nominalni nivo pouzdanosti 90%, za parametar β_1 je utvrđeno da su u većini slučajeva za male uzorke proporcije simuliranih percentil i *bootstrap-t* intervala bile najbliže 0,90, dok su za velike uzorke najbolji rezultati dobijeni korišćenjem *jackknife* i percentil intervala. Kada je reč o parametru β_2 , u većini slučajeva su za male uzorke proporcije simuliranih percentil i *bootstrap-t* intervala sa transformacijom T_2^* bile najbolje. Za velike uzorke se *bootstrap-t* interval sa transformacijom T_2^* pokazao kao najbolji izbor.

Zaključuje se da je prilikom intervalnog ocenjivanja parametra β_1 najbolje koristiti percentile, *bootstrap-t* i *jackknife* intervale, dok se za parametar β_2 preporučuje primena *bootstrap-t* intervala sa transformacijom T_2^* i percentil intervala.

2.5 Analiza i primena u korporativnim finansijama

Na početku ovog odeljka biće dat uvod u ekonomsku disciplinu korporativne finansije, a zatim će se detaljno opisati promenljive na osnovu kojih će se izvršiti analiza pouzdanosti intervala poverenja za regresione koeficijente u prostom linearnom i kvadratnom regresionom modelu. Reč je o indikatoru verovatnoće bankrotstva i količniku ukupog duga. Na kraju će se izvršiti analiza proporcija simuliranih intervala za regresione koeficijente.

2.5.1 Korporativne finansije

Razvoj korporativnih finansija vezuje se za početak 20. veka. Značajan doprinos nastanku ove discipline imale su institucije tržišta kapitala i usavršavanje finansijskih instrumenata (Aragon, 1989; Ross et al., 2003; Brealey et al., 2011; Damodaran, 2015; Besley & Brigham, 2018; Todorović & Ivanišević, 2018).

Korporativne finansije, kao ekonomska disciplina, trebalo bi da daju odgovor na sledeća pitanja: koje dugoročne investicije bi trebalo preduzeti; kako obezbediti dugoročno finansiranje da bi se platile investicije i kako upravljati svakodnevnim finansijskim aktivnostima poput prikupljanja novca od kupaca i plaćanja dobavljačima (Ross et al., 2003).

U periodu razvoja korporativnih finansija glavni cilj kompanije bio je da nabavi kapital, zbog čega se smatra da je ovakav pristup finansijama tradicionalan. Predmet interesovanja su bili pravni i finansijski odnosi između kompanije i finansijskih institucija, kao i instrumenti pomoću kojih se kapital nabavlja. Ključne zamerke koje su upućivane tradicionalnom pristupu korporativnim finansijama su u tome što je problem finansija u određenoj meri trebalo da se razmatra sa stanovišta donosioca investicionih odluka u kompaniji, kratkoročni izvori finansiranja se nisu razmatrali, velika pažnja se posvećivala problemima finansiranja akcionarskih kompanija u odnosu na druge oblike kompanije, i drugo. Tradicionalni pristup korporativnim finansijama se nije bavio temama kao što su: cena kapitala, utvrđivanje postojanja razlike u ceni kapitala u slučaju primene različitih metoda finansiranja, u kojoj meri očekivani prinosi zadovoljavaju finansijske standarde, itd. U periodu posle Drugog svetskog rata raste tržišna konkurencija, usavršavaju se finansijsko tržište i finansijski instrumenti, a penzioni fondovi i osiguravajuće kompanije su smatrani glavnim investitorima. Odgovornost menadžera dobija na značaju, zbog čega se u cilju donošenja dobrih finansijskih odluka javlja potreba za upotrebom naučnih metoda. Razvoj savremenog pristupa korporativnim finansijama počeo je krajem 50-ih godina prošlog veka, kada su korporativne finansije počele da se zasnivaju na kvantitativnim metodama. Upotreba ovih metoda omogućila je menadžerima da se usmere ka optimalnoj alokaciji kapitala kako bi se ostvarili dugoročni ciljevi kompanije. Savremeni pristup korporativnim finansijama podrazumeva da se pored nabavke kapitala razmatra i vrsta sredstava koju bi kompanija trebalo da nabavi, kao i način na koji bi potrebna sredstva trebalo da budu finansirana (Todorović & Ivanišević, 2018).

Ova disciplina obuhvata sledeće tri oblasti: investiranje, finansiranje i dividende. Investiranje podrazumeva donošenje odluka o tome na koji način kompanija ulaže svoje resurse.

Kompanije imaju ograničene resurse, zbog čega je jedno od najvažnijih pitanja koje se razmatra u korporativnoj finansijskoj teoriji donošenje odluke o raspodeli resursa. Investicione odluke obuhvataju one odluke koje se odnose na stvaranje prihoda i dobiti (na primer, uvođenje nove linije proizvoda), zatim odluke koje se odnose na štednju novca (na primer, izgradnja efikasnijeg sistema distribucije), kao i odluke o tome da li i koliko kredita odobriti kupcima. U investicione odluke ubrajaju se i odluke koje se odnose na upravljanje zalihama, kao i odluke u vezi sa ulaskom na određena tržišta. Resurse bi trebalo ulagati u projekte koji obezbeđuju prinose veće od minimalno prihvatljive zahtevane stope prinosa. Zahtevana stopa prinosa bi trebalo da bude veća za rizičnije projekte, kao i da odražava udeo sopstvenog kapitala i duga u ukupnim finansijskim sredstvima, dok bi prinosi na projekte trebalo biti mereni na osnovu ostvarenih novčanih tokova. Finansiranje podrazumeva upravljanje finansijskim sredstvima koja se koriste za investiranje, odnosno određivanje odnosa sopstvenog kapitala i duga koji maksimiziraju vrednost izvršenog ulaganja. Svaka investicija se finansira iz sopstvenog kapitala ili pozajmljenog novca (duga). Kod privatnih kompanija, štednja vlasnika predstavlja kapital, a dug se sastoji od bankarskih pozajmica. Što se tiče javnih kompanija, sopstveni kapital potiče iz emisije akcija i akumuliranja dobitka, dok je dug u obliku obveznica. Kada je reč o dividendama, potrebno je utvrditi koji deo dobitka bi trebalo ponovo uložiti, a koji deo vratiti vlasnicima (Damodaran, 2015).

Postoji nekoliko ključnih faktora koji su doprineli razvoju korporativnih finansija (Todorović & Ivanišević, 2018, str. 3-4):

- Globalizacija – reč je o globalizaciji poslovanja koja zavisi od razvoja transporta, ponašanja kupaca, konstantnog rasta troškova istraživanja i razvoja i sposobnosti multinacionalnih kompanija da u kratkom roku organizuju proizvodnju u zemljama u kojima će cena koštanja biti znatno niža.
- Razvoj informacione i komunikacione tehnologije - plaćanje, evidencija prodaje, kao i nabavka zaliha obavljaju se koristeći računare, a zahvaljujući upotrebi interneta došlo je do razvoja IT platformi za e-trgovinu, e-poslovanje, itd.
- Stalni razvoj finansijske prakse i teorije – razvijaju se usluge poput konsultantskih usluga, procene vrednosti kompanije, standardizacije finansijskih analiza. Usavršavaju se metode koje se koriste za merenje dostignuća kompanije, a sledeći koncepti sve više dobijaju na značaju: koncept stvaranja vrednosti za akcionare koji podrazumeva dodavanje vrednosti za akcionare – SVA (skraćeno od engl. Shareholder Value Added), koncept upotrebe merila dodate tržišne vrednosti – MVA (skraćeno od engl. Market Value Added) i koncept dodate ekonomske vrednosti – EVA (skraćeno od engl. Economic Value Added).
- Upravljanje rizikom – dolazi do porasta emitovanja i prometa finansijskih derivata, kao što su opcije i fjučersi (značajni su zbog mogućnosti ostvarenja visokih prinosa i kao instrument zaštite od različitih rizika).

Predmet korporativnih finansija je upravljanje finansijama u kompaniji. Ova naučna disciplina je u bliskoj vezi sa disciplinama kao što su analiza hartija od vrednosti, finansijska tržišta, bankarstvo i osiguranje. Na primer, finansijski menadžeri bi trebalo da znaju pod kojim uslovima može da se nabavi kapital i na koji način finansijska tržišta vrednuju hartije od vrednosti (Todorović & Ivanišević, 2018). Finansijski menadžeri posreduju između kompanije i investitora. Sa jedne strane pomažu kompaniji donoseći dobre poslovne odluke, a sa druge strane saraduju sa bankama i finansijskim tržištima (Brealey et al., 2011).

Discipline koje su od posebnog značaja za finansijske menadžere su računovodstvo, informacione tehnologije i strateški menadžment. Računovodstvo omogućava korišćenje finansijskih izveštaja na osnovu kojih se sprovode finansijske analize, razmatra finansijska pozicija kompanije i planiraju budući novčani tokovi. Takođe, podrazumeva se da finansijski menadžeri budu upoznati sa informacionim tehnologijama čijom se upotrebom dobijaju relevantne informacije. Na kraju, neophodno je razumevanje šansi i opasnosti iz okruženja sa kojima se kompanija suočava. U tom smislu, strateški menadžment pomaže da se izvrši dugoročno finansijsko planiranje, da se sagledaju potencijalni investicioni projekti, i drugo (Todorović & Ivanišević, 2018).

2.5.2 Indikator verovatnoće bankrotstva

Najveći problem koji se može desiti investitorima je bankrot kompanije. Postoje različiti pristupi koji se koriste da bi se izvršila procena finansijske snage kompanije. Korišćenje finansijskih indikatora često može biti od pomoći kako bi se ustanovila dugoročna održivost kompanije. Kreditna sposobnost kompanije može se proceniti na osnovu finansijskih indikatora profitabilnosti, likvidnosti i solventnosti. Problem se javlja kada se dobijaju kontradiktorni rezultati. Na primer, indikatori profitabilnosti mogu ukazivati na dobro stanje kompanije, dok indikatori likvidnosti pokazuju pogoršanje kratkoročne sigurnosti (Kuč, 2017, str. 62).

U cilju rešavanja navedenog problema, Edward Altman (Altman, 1968) je razvio Z-skor (engl. Z-Score) model, koji se koristi za predviđanje bankrotstva. Altman (1968) je primenio multivarijacionu diskriminacionu analizu na podacima kompanija iz proizvodnih delatnosti čijim akcijama se otvoreno trgovalo na berzi. Koristio je stratifikovani slučajni uzorak koji je obuhvatao 66 kompanija (kompanije koje su bankrotirale u periodu od 1946. do 1965. su činile prvu polovinu uzorka, dok su drugu polovinu činile kompanije koje nisu bankrotirale u istom periodu).

Prilikom formiranja modela utvrđeno je da pet promenljivih na najbolji način objašnjava bankrotstvo kompanije. Z-skor diskriminaciona funkcija može se predstaviti u sledećem obliku (Altman, 1968, str. 594):

$$Z = 0,012X_1 + 0,014X_2 + 0,033X_3 + 0,006X_4 + 0,999X_5 ,$$

gde se navedene promenljive mogu izračunati na sledeći način:

$$X_1 = \frac{\text{neto obrtni kapital}}{\text{ukupna imovina}} ,$$

$$X_2 = \frac{\text{neraspodeljeni dobitak}}{\text{ukupna imovina}} ,$$

$$X_3 = \frac{\text{dobitak pre kamata i poreza}}{\text{ukupna imovina}} ,$$

$$X_4 = \frac{\text{tržišna vrednost sopstvenog kapitala}}{\text{knjižovodstvena vrednost ukupnih obaveza}} ,$$

$$X_5 = \frac{\text{prihodi od prodaje}}{\text{ukupna imovina}}.$$

Altman je utvrdio sledeća pravila za interpretaciju Z-skora odnosno indikatora verovatnoće bankrotstva kompanije: vrednost Z-skora ispod 1,8 ukazuje da je velika verovatnoća bankrota kompanije; za kompaniju čiji je Z-skor iznad 3 se smatra da neće bankrotirati. Vrednost Z-skora između 1,8 i 3 ukazuje da je status kompanije nedefinisan (velika je verovatnoća da se napravi greška u predviđanju).

Nedostatak Z-skor modela je u tome što ga nije moguće primeniti na preduzeća koja nisu u formi otvorenog akcionarskog društva. Takođe, ovaj model nema veliku preciznost ukoliko se primenjuje na neproizvodna preduzeća.

Sa ciljem da se otklone navedeni nedostaci, kreirane su dve verzije originalnog modela. Prva verzija je namenjena za predviđanje bankrotstva preduzeća kojima se ne trguje na berzi (engl. Z'-Score Private Firm Model). Razlika u odnosu na prvobitni model je u tome što je promenljiva X_4 modifikovana i računa se na sledeći način:

$$X_4 = \frac{\text{knjigovodstvena vrednost sopstvenog kapitala}}{\text{knjigovodstvena vrednost ukupnih obaveza}}.$$

Ova modifikacija je uticala na sve promenljive, tako da je novi model određen sledećom formulom (Altman, 1993):

$$Z' = 0,717X_1 + 0,847X_2 + 3,107X_3 + 0,420X_4 + 0,998X_5.$$

Altman et al. (1995) su kreirali drugu verziju kako bi se model primenio na neproizvodna preduzeća i preduzeća van područja na kojima je razvijeno tržište kapitala. Reč je o modelu koji se koristi da bi se procenio kreditni rizik preduzeća koja se nalaze na tržištima u razvoju, odnosno EM-skor modelu (engl. Emerging Markets Scoring Model). Razlika u odnosu na originalni model je u tome što je promenljiva X_4 modifikovana i određuje se na identičan način kao u Z'-skor modelu. Promenljiva X_5 je zamenjena konstantom da bi se eliminisao uticaj delatnosti. Model se može predstaviti na sledeći način:

$$Z'' = 6,56X_1 + 3,26X_2 + 6,72X_3 + 1,05X_4 + 3,25.$$

Kako bi se odredila verovatnoća bankrotstva kompanije na osnovu EM-skora potrebno je imati uvid u ekvivalentan kreditni rejting američkih korporativnih obveznica. U Tabeli 2.15 prikazani su kriterijumi za određivanje ekvivalentnog kreditnog rejtinga. Može se uočiti da postoje tri zone rizika. Ako je EM-skor manji od 4,15 kompanija je u crvenoj zoni (ekvivalentan kreditni rejting je u intervalu od B- do D). Ovo je signal da postoji visoka verovatnoća bankrotstva kompanije. Ukoliko je EM-skor veći od 5,85, kompanija se nalazi u zoni komfora (ekvivalentan kreditni rejting je u intervalu od BBB do AAA). U tom slučaju je kreditni rizik na niskom nivou. Siva zona se nalazi se u intervalu od 4,15 do 5,85 (ekvivalentan kreditni rejting je u intervalu od BBB- do B).

Tabela 2.15: EM-skor i kreditni rejting kompanije

EM-skor kriterijum	Kreditni rejting
>8,15	AAA
7,60-8,15	AA+
7,30-7,60	AA
7,00-7,30	AA-
6,85-7,00	A+
6,65-6,85	A
6,40-6,65	A-
6,25-6,40	BBB+
5,85-6,25	BBB
5,65-5,85	BBB-
5,25-5,65	BB+
4,95-5,25	BB
4,75-4,95	BB-
4,50-4,75	B+
4,15-4,50	B
3,75-4,15	B-
3,20-3,75	CCC+
2,50-3,20	CCC
1,75-2,50	CCC-
<1,75	D

Izvor: Preuzeto iz Altman (2005, str. 314)

2.5.3 Količnik ukupnog duga

Količnik ukupnog duga (engl. total debt ratio) ili ukupan finansijski leveridž (engl. total financial leverage) određuje se kao odnos ukupnih obaveza i ukupne imovine i koristi za procenu nivoa duga kompanije. Pokazatelj je dugoročne sigurnosti poverilaca, s obzirom da daje informaciju koliko imovine ostaje na raspolaganju u slučaju likvidacije (Rajan & Zingales, 1995, str. 1428).

Leveridž, kao poslovni termin, odnosi se na dug ili pozajmljivanje sredstava da bi kompanija finansirala kupovinu opreme, zaliha i druge imovine. Ukoliko vlasnici kompanija koriste dug da bi finansirali kupovinu imovine, raste rizik da kompanija bankrotira. Ako je kompanija viđena kao visoko zadužena (tj. ima visok količnik ukupnog duga), znači da ima previše kredita i da svoj rast finansira dugom. Zaključuje se da će sa porastom količnika ukupnog duga Z-skor biti niži, odakle sledi da je veća verovatnoća da će kompanija bankrotirati.

Prilikom računanja količnika ukupnog duga moguće je koristiti tržišne ili knjigovodstvene vrednosti. Tržišna vrednost kapitala odražava očekivanja investitora u pogledu budućih prinosa, zbog čega tržišno iskazani indikatori govore o sposobnosti da se u budućnosti izmire obaveze. Knjigovodstvena vrednost određuje se kao razlika između vrednosti ukupne imovine i ukupnih obaveza. Na osnovu knjigovodstveno izraženih indikatora stiče se utisak o poslovanju kompanije u prošlosti (Hillier et al., 2012, str. 466).

U razvijenim zemljama tržišne vrednosti kompanije su uglavnom veće od knjigovodstvenih vrednosti, zbog čega su knjigovodstveno iskazani indikatori strukture finansiranja precenjeni. Suprotno, ako dođe do pada kreditnog rejtinga, tržišna vrednost duga može biti niža od nominalne vrednosti. Indikator leveridža koji balansira odnos između tržišnih i knjigovodstvenih vrednosti je količnik knjigovodstvene vrednosti duga i tržišne vrednosti sopstvenog kapitala. Na primer, ako je

tržišna vrednost veća od knjigovodstvene vrednosti, tržište smatra da je položaj kompanije takav da opravdava privlačenje dodatnog kapitala po povoljnim uslovima (White et al., 2003, str. 131).

2.5.4 Analiza pouzdanosti intervala poverenja za regresioni koeficijent u prostom linearnom regresionom modelu na realnim podacima

Predmet razmatranja ovog odeljka biće proporcije simuliranih intervala za regresioni koeficijent u prostom linearnom regresionom modelu zasnovane na podacima o indikatoru verovatnoće bankrotstva i količniku ukupnog duga i širine intervala poverenja zasnovane na podacima o zbirnim finansijskim indikatorima. Na kraju odeljka će se razmatrati proporcije simuliranih intervala za regresioni koeficijent u prostom linearnom regresionom modelu zasnovane na podacima o prosečnoj ceni leta i udaljenosti između gradova.

Primer 1 U ovom primeru analiziraće se proporcije simuliranih intervala za regresioni koeficijent u prostom linearnom regresionom modelu (predmet razmatranja biće t , $bootstrap-t$, T_1 , T_2 i T_3 intervali) bazirane na podacima o indikatoru verovatnoće bankrotstva²⁸ i količniku ukupnog duga u 157 kompanija u Srbiji u 2014. godini (Ivković et al., 2020). Ovi podaci su javno dostupni u finansijskim izveštajima preduzeća. Istraživanje je sprovedeno na najvećim preduzećima u Srbiji prema kriterijumu visine poslovnih prihoda, odnosno na preduzećima koja ostvaruju najveći nivo poslovne aktivnosti. Razmatrana su preduzeća čiji su poslovni prihodi u 2013. godini bili iznad 5 milijardi dinara.

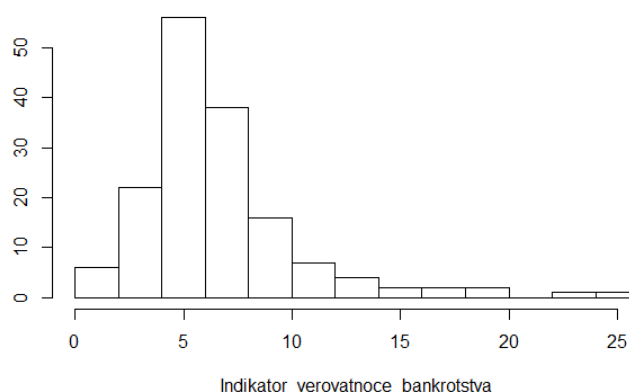
Deskriptivne mere za analizirane podatke prikazane su u Tabeli 2.16.

Tabela 2.16: Deskriptivne mere za indikator verovatnoće bankrotstva i količnik ukupnog duga

Promenljiva	Broj kompanija	Sredina	Standardna devijacija
Indikator verovatnoće bankrotstva	157	6,65	3,82
Količnik ukupnog duga	157	0,55	0,25

Izvor: Preuzeto iz Ivković et al. (2020)

Slikom 2.9 predstavljen je histogram raspodele indikatora verovatnoće bankrotstva.

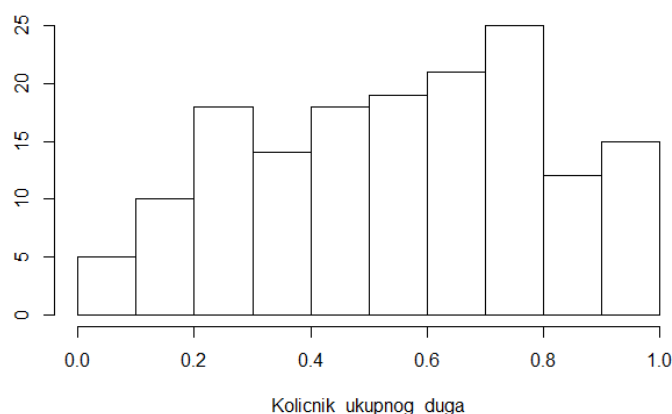


Slika 2.9: Histogram raspodele indikatora verovatnoće bankrotstva

Izvor: Preuzeto iz Ivković et al. (2020)

²⁸ Podaci o indikatoru verovatnoće bankrotstva i količniku ukupnog duga nalaze se u PRILOZIMA 1 i 2.

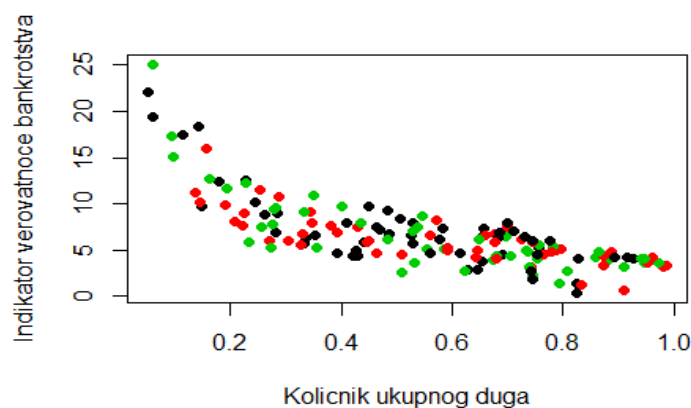
Histogram raspodele količnika ukupnog duga predstavljen je Slikom 2.10.



Slika 2.10: Histogram raspodele količnika ukupnog duga

Izvor: Preuzeto iz Ivković et al. (2020)

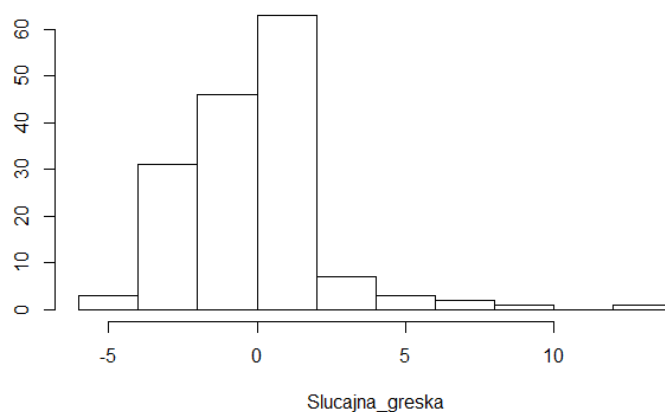
Na Slici 2.11 prikazan je dijagram raspršenosti dobijen na osnovu podataka o količniku ukupnog duga i indikatoru verovatnoće bankrotstva. Koristeći podatke o indikatoru verovatnoće bankrotstva koji predstavlja zavisnu promenljivu i količniku ukupnog duga koji predstavlja objašnjavajuću promenljivu dobijen je prost linearni regresioni model osnovnog skupa sa regresionim koeficijentima $\beta_0 = 12,98$ i $\beta_1 = -11,43$. Koeficijent determinacije ovog modela iznosi 56,06%, što znači da je 56,06% varijacija promenljive indikator verovatnoće bankrotstva objašnjeno promenljivom količnik ukupnog duga. Standardna devijacija slučajne greške je 2,533, što predstavlja odstupanja stvarnih vrednosti promenljive indikator verovatnoće bankrotstva od njenih prosečnih vrednosti na regresionoj pravnoj osnovnog skupa.



Slika 2.11: Dijagram raspršenosti na osnovu podataka o količniku ukupnog duga i indikatoru verovatnoće bankrotstva

Izvor: Izlaz iz programskog jezika R

Histogram raspodele slučajne greške, zasnovan na dobijenom modelu, predstavljen je na Slici 2.12. Uočava se da slučajna greška nije normalno raspodeljena. Takođe, upotrebljen je Shapiro-Wilk test normalnosti koji je potvrdio da slučajna greška nije normalno raspodeljena (H_0 : Slučajna greška je normalno raspodeljena, a dobijena p -vrednost je približno 0).



Slika 2.12: Histogram raspodele slučajne greške

Izvor: Preuzeto iz Ivković et al. (2020)

Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent zasnovane na podacima o indikatoru verovatnoće bankrotstva i količniku ukupnog duga prikazane su u Tabeli 2.17.

Tabela 2.17: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent kada indikator verovatnoće bankrotstva zavisi od količnika ukupnog duga

n	10	20	50	80
t	0,835	0,807	0,879	0,920
Bootstrap-t	0,853	0,857	0,929	0,972
T_1	0,610	0,580	0,613	0,623
T_2	0,842	0,850	0,936	0,973
T_3	0,495	0,810	0,939	0,975

Izvor: Preuzeto iz Ivković et al. (2020)

Uočeno je da su za male uzorke (veliĉine 10 i 20) proporcije simuliranih *bootstrap-t* intervala bile najbolje meĊu razmatranim intervalima, ali ni one nisu blizu nominalnog nivoa pouzdanosti. Za velike uzorke (veliĉine 50 i 80), T_3 i *bootstrap-t* intervali su obezbedili najbolje rezultate.

Primer 2 Predmet razmatranja u ovom hipotetiĉkom primeru biće intervali poverenja za regresione koeficijente u prostim linearnim regresionim modelima i njihove širine, zasnovani na podacima o zbirnim finansijskim indikatorima u 105 opština u Srbiji za 2017. godinu.

Projektni timovi nekoliko konsultantskih kompanija su u periodu maj-jun 2018. godine sproveli istraživanje u kom su prikupljeni podaci za brojne indikatore za 105 opština u Srbiji u periodu od 2008. do 2017. godine. Korišćeni su različiti izvori podataka, poput javno dostupnih izvora (Republički Zavod za Statistiku) i finansijskih izveštaja.

Prikupljeni su podaci o sledećih osam grupa indikatora: indikatori povezanosti i mobilnosti; indikatori razvoja lokalnog tržišta; poljoprivredna produktivnost; pristup

domaćinstava indikatorima blagostanja i obrazovanja; indikatori razvoja malih kompanija; indikatori razvoja srednjih i velikih kompanija; indikatori demografskih kretanja i indikatori opštih ekonomskih uslova.

U ovom primeru²⁹ biće analizirani sledeći zbirni finansijski indikatori, koji pripadaju indikatorima razvoja malih, srednjih i velikih preduzeća, i indikatorima opštih ekonomskih uslova, a odnose se na 2017. godinu: broj zaposlenih³⁰, neto dobitak u hiljadama dinara, ukupna imovina u hiljadama dinara i troškovi sirovina u hiljadama dinara.

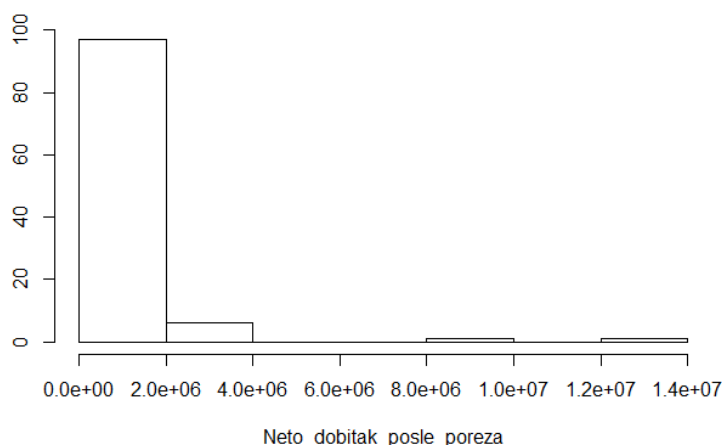
Deskriptivne mere za analizirane podatke prikazane su u Tabeli 2.18.

Tabela 2.18: Deskriptivne mere za zbirne finansijske indikatore

Promenljiva	Sredina	Standardna Devijacija	Koeficijent Asimetrije
Neto dobitak	625235,61	1639685,18	5,69
Ukupna imovina	10969610,22	26580245,93	5,42
Troškovi sirovina	1525438,80	2407229,72	4,21
Broj zaposlenih	1522,60	2714,49	4,26

Izvor: Izlaz iz programskog jezika R

Sledećim Slikama 2.13-2.16 predstavljeni su histogrami raspodele navedenih promenljivih. Na Slici 2.13 predstavljen je histogram raspodele neto dobitka.



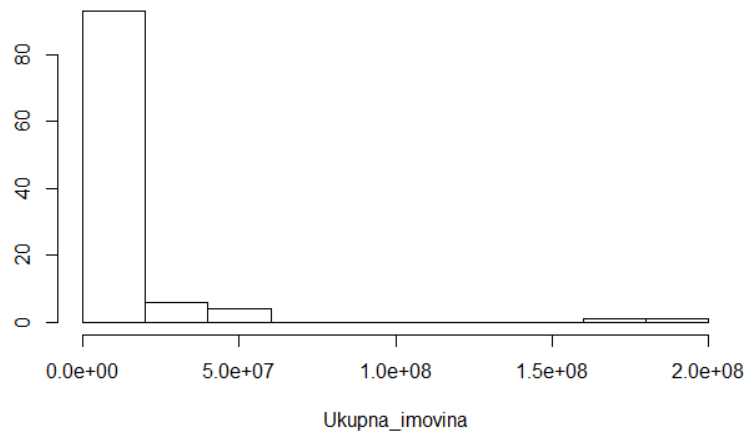
Slika 2.13: Histogram raspodele neto dobitka

Izvor: Izlaz iz programskog jezika R

Slikom 2.14 predstavljen je histogram raspodele ukupne imovine.

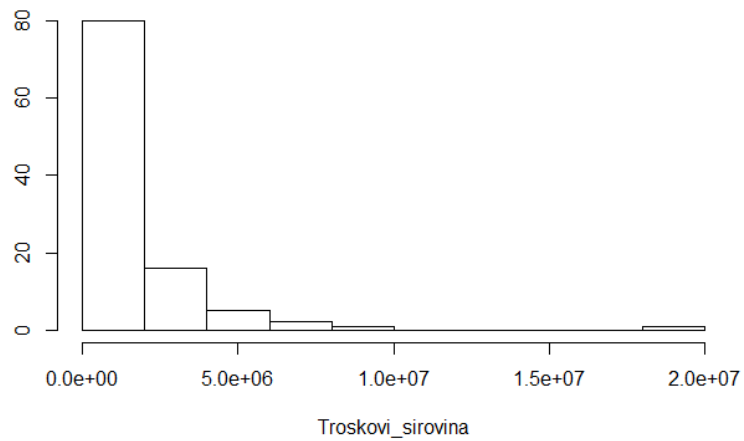
²⁹ Promenljiva ukupna imovina se koristi u regresionom modelu u odeljku 2.5.5.

³⁰ Podaci o neto dobitku, ukupnoj imovini, troškovima sirovina i broju zaposlenih nalaze se u PRILOZIMA 3, 4, 5 i 6.



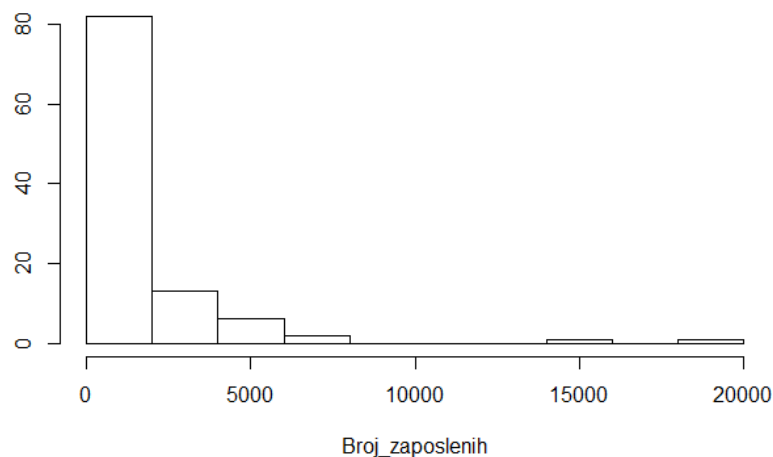
Slika 2.14: Histogram raspođe ukupne imovine
Izvor: Izlaz iz programskog jezika R

Histogram raspođe troškova sirovina predstavljen je Slikom 2.15.



Slika 2.15: Histogram raspođe troškova sirovina
Izvor: Izlaz iz programskog jezika R

Slika 2.16 predstavlja histogram raspođe broja zaposlenih.

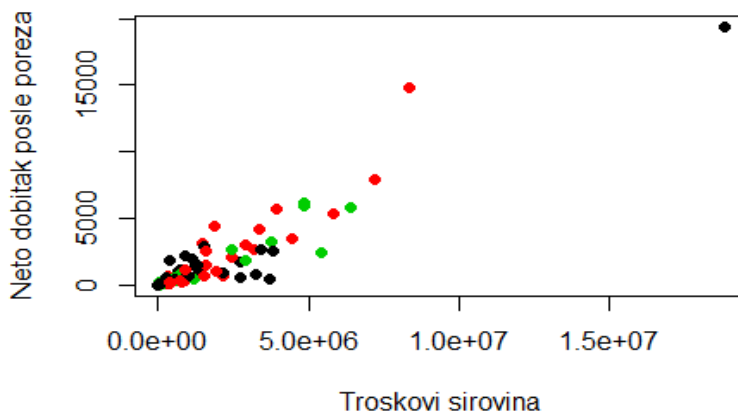


Slika 2.16: Histogram raspodele broja zaposlenih

Izvor: Izlaz iz programskog jezika R

Poći će se od pretpostavke da je neto dobitak zavisna promenljiva, a da su troškovi sirovina i broj zaposlenih objašnjavajuće promenljive. U ovom primeru ispitiće se širine intervala poverenja.

Na Slici 2.17 prikazan je dijagram raspršenosti dobijen na osnovu podataka o neto dobitku i troškovima sirovina. Koristeći navedene podatke dobijena je regresiona prava uzorka koja glasi: $\hat{y} = -211800 + 0,5487x$. Koeficijent determinacije iznosi 64,89%, odnosno 64,89% varijacija neto dobitka objašnjeno je troškovima sirovina. Standardna greška regresije koja pokazuje za koliko u proseku stvarne vrednosti neto dobitka odstupaju od njegovih ocenjenih vrednosti na regresionoj pravoj uzorka iznosi 976200 hiljada dinara.



Slika 2.17: Dijagram raspršenosti na osnovu podataka o troškovima sirovina i neto dobitku

Izvor: Izlaz iz programskog jezika R

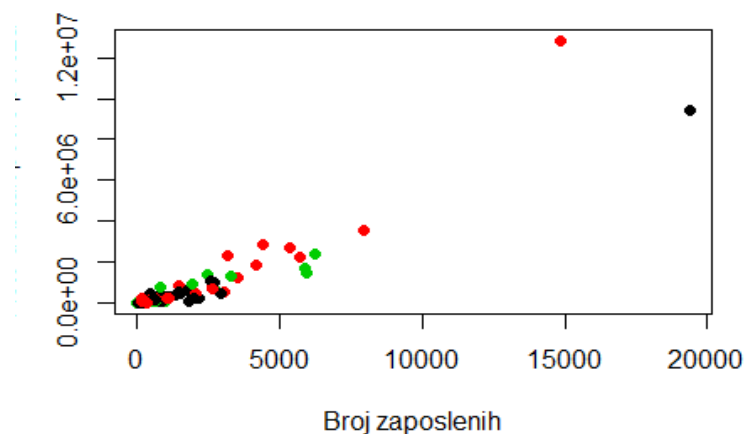
U Tabeli 2.19 date su širine 95% intervala poverenja za regresioni koeficijent bazirane na podacima o neto dobitku i troškovima sirovina.

Tabela 2.19: Širine 95% dvostranih intervala poverenja za regresioni koeficijent kada neto dobitak zavisi od troškova sirovina

	Interval	Širina intervala
t	(0,4698;0,6276)	0,1578
<i>Bootstrap-t</i>	(0,3883;0,9650)	0,5767
T_1	(0,4493;0,5487)	0,0994
T_2	(0,3863;1,0567)	0,6704
T_3	(0,3989;0,5487)	0,1498

Izvor: Izlaz iz programskog jezika R

Na Slici 2.18 prikazan je dijagram raspršenosti dobijen na osnovu podataka o neto dobitku i broju zaposlenih. Koristeći navedene podatke dobijena je regresiona prava uzorka koja glasi: $\hat{y} = -229155,39 + 561,14x$. Koeficijent determinacije iznosi 86,30% i pokazuje da je 86,30% varijacija neto dobitka objašnjeno brojem zaposlenih. Standardna greška regresije koja pokazuje za koliko u proseku stvarne vrednosti neto dobitka odstupaju od njegovih ocenjenih vrednosti na regresionoj pravoj uzorka iznosi 609900 hiljada dinara.



Slika 2.18: Dijagram raspršenosti na osnovu podataka o broju zaposlenih i neto dobitku

Izvor: Izlaz iz programskog jezika R

Širine 95% intervala poverenja za regresioni koeficijent zasnovane na podacima o neto dobitku i broju zaposlenih prikazane u Tabeli 2.20.

Tabela 2.20: Širine 95% dvostranih intervala poverenja za regresioni koeficijent kada neto dobitak zavisi od broja zaposlenih

	Interval	Širina intervala
t	(517,44;604,84)	87,40
Bootstrap-t	(407,56;846,83)	439,27
T_1	(486,29;561,14)	74,85
T_2	(426,71;966,92)	540,21
T_3	(442,05;561,14)	119,09

Izvor: Izlaz iz programskog jezika R

U oba slučaja najuži interval poverenja dobijen je primenom *bootstrap* metoda u kombinaciji sa transformacijom T_1 .

*Primer 3*³¹ Na osnovu podataka koji se odnose na prosečnu cenu leta i udaljenost između gradova analiziraće se proporcije simuliranih intervala za regresioni koeficijent u prostom linearnom regresionom modelu koristeći podatke o 1000 letova za četvrti kvartal 2002. godine³² (Ivković et al., 2020).

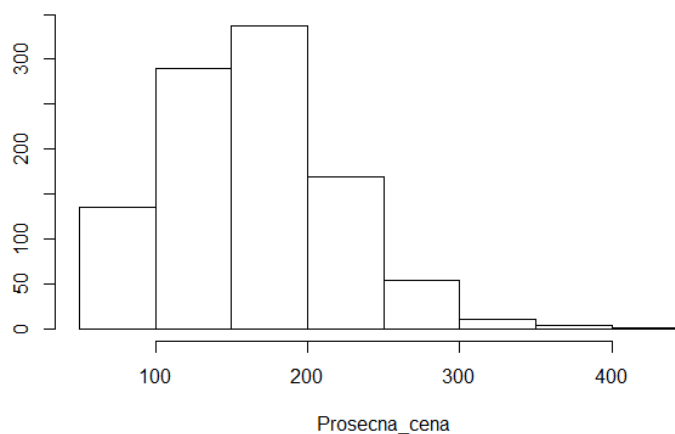
Deskriptivne mere za ispitivane podatke predstavljene su u Tabeli 2.21.

Tabela 2.21: Deskriptivne mere za prosečnu cenu i udaljenost

Promenljiva	Broj letova	Sredina	Standardna devijacija
Prosečna cena	1000	163,38	55,34
Udaljenost	1000	1056,97	642,88

Izvor: Preuzeto iz Ivković et al. (2020)

Histogram raspodele prosečne cene prikazan je na Slici 2.19.



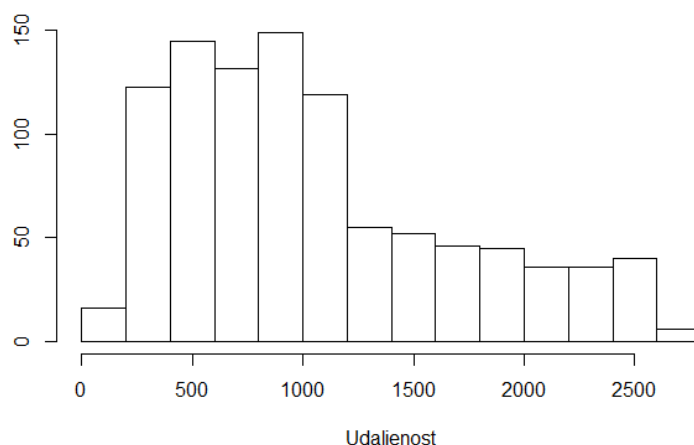
Slika 2.19: Histogram raspodele prosečne cene

Izvor: Preuzeto iz Ivković et al. (2020)

³¹ *Primer 3* je hipotetički primer.

³² Podaci su preuzeti sa sajta: <http://www.stat.ufl.edu/~winner/datasets.html>, iz baze podataka: Airfare i Demand 2002Q4 i biće navedeni u PRILOZIMA 7 i 8.

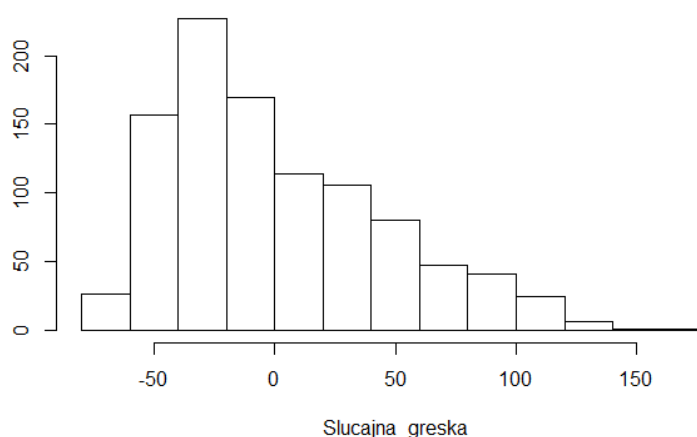
Slikom 2.20 predstavljen je histogram raspodele udaljenosti.



Slika 2.20: Histogram raspodele udaljenosti
Izvor: Preuzeto iz Ivković et al. (2020)

Koristeći navedene podatke o prosečnoj ceni leta koja predstavlja zavisnu promenljivu i udaljenosti koja predstavlja objašnjavajuću promenljivu dobijen je prost linearni regresioni model osnovnog skupa sa regresionim koeficijentima $\beta_0 = 109,95$ i $\beta_1 = 0,05$. Koeficijent determinacije ovog modela iznosi 34,48%, što znači da je 34,48% varijacija promenljive prosečna cena objašnjeno promenljivom udaljenost. Standardna devijacija slučajne greške je 44,795, što predstavlja odstupanja stvarnih vrednosti promenljive prosečna cena od njenih prosečnih vrednosti na regresionoj pravoj osnovnog skupa.

Na Slici 2.21 predstavljen je histogram raspodele slučajne greške, na osnovu koga se uočava da slučajna greška nije normalno raspodeljena. Takođe, sproveden je Shapiro-Wilk test normalnosti koji je pokazao isti rezultat kao histogram (H_0 : Slučajna greška je normalno raspodeljena, a dobijena p -vrednost je približno 0).



Slika 2.21: Histogram raspodele slučajne greške
Izvor: Preuzeto iz Ivković et al. (2020)

U Tabeli 2.22 predstavljene su proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent zasnovane na podacima o prosečnoj ceni i udaljenosti.

Tabela 2.22: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent kada prosečna cena zavisi od udaljenosti

n	10	20	50	80
t	0,898	0,908	0,902	0,924
Bootstrap-t	0,923	0,929	0,932	0,941
T_1	0,795	0,818	0,840	0,900
T_2	0,901	0,890	0,918	0,944
T_3	0,561	0,881	0,942	0,950

Izvor: Preuzeto iz Ivković et al. (2020)

Na osnovu Tabele 2.22 uočava se da su za male uzorke (veličine 10 i 20), proporcije simuliranih *bootstrap-t* intervala bile najbolje (iznad 0,920), dok je kod velikih uzoraka (veličine 50 i 80) T_3 interval obezbedio najbolje rezultate (iznad 0,940).

Uzimajući u obzir rezultate predstavljene u odeljcima 2.4.1 i 2.5.4 potvrđuje se hipoteza

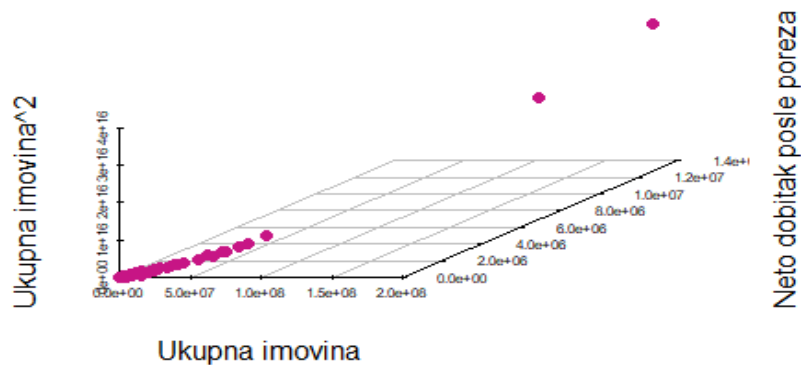
H_1 : Proporcije simuliranih intervala poverenja za regresioni koeficijent u prostom linearnom regresionom modelu koji su zasnovani na novoj transformaciji t statistike, T_3 , za velike uzorke, bliže su nominalnom nivou pouzdanosti u odnosu na proporcije ostalih proučavanih intervala.

2.5.5 Analiza pouzdanosti intervala poverenja za regresione koeficijente u kvadratnom regresionom modelu na realnim podacima

Primer 1 Predmet istraživanja u ovom primeru biće proporcije simuliranih intervala za regresione koeficijente u kvadratnom regresionom modelu zasnovane na podacima o neto dobitku i ukupnoj imovini u 105 opština u Srbiji u 2017. godini.³³

Na Slici 2.22 prikazan je dijagram raspršenosti dobijen na osnovu podataka o neto dobitku i ukupnoj imovini. Koristeći podatke o neto dobitku koji predstavlja zavisnu promenljivu, ukupnoj imovini i kvadriranim vrednostima ukupne imovine koje predstavljaju objašnjavajuće promenljive dobijen je kvadratni regresioni model osnovnog skupa sa regresionim koeficijentima $\beta_0 = -4106$, $\beta_1 = 5,404 \cdot 10^{-2}$ i $\beta_2 = 4,453 \cdot 10^{-11}$. Koeficijent determinacije ovog modela je 98,97%, odnosno 98,97% varijacija neto dobitka objašnjeno je ukupnom imovinom i kvadriranim vrednostima ukupne imovine. Standardna devijacija slučajne greške je 165582,61 hiljada dinara, što predstavlja odstupanja stvarnih vrednosti neto dobitka od njegovih prosečnih vrednosti na regresionoj ravni osnovnog skupa.

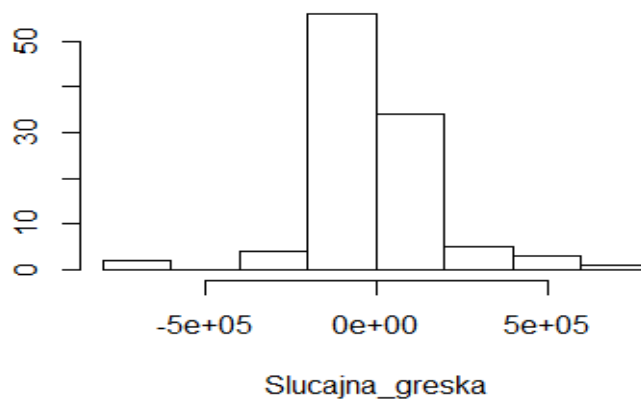
³³ Podaci su korišćeni u *Primeru 2* u odeljku 2.5.4 i biće navedeni u PRILOZIMA 3 i 4.



Slika 2.22: Dijagram raspršenosti na osnovu podataka o neto dobitku i ukupnoj imovini

Izvor: Izlaz iz programskog jezika R

Slika 2.23 predstavlja histogram raspodele slučajne greške, zasnovan na prethodnom modelu. Na osnovu histograma je uočeno da slučajna greška nije normalno raspodeljena. Shapiro-Wilk test normalnosti takođe je potvrdio da slučajna greška nije normalno raspodeljena (H_0 : Slučajna greška je normalno raspodeljena, a p -vrednost je približno 0).



Slika 2.23: Histogram raspodele slučajne greške

Izvor: Izlaz iz programskog jezika R

Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_1 zasnovane na podacima o neto dobitku i ukupnoj imovini prikazane su u Tabeli 2.23. Analiza rezultata pokazala je sledeće: za uzorke veličine 10, proporcija simuliranih percentil intervala bila je najbliža nominalnom nivou pouzdanosti, dok je kod uzoraka veličine 20, *jackknife* interval dao najbolje rezultate; za velike uzorke (veće od 50), najbolji rezultati su dobijeni primenom *bootstrap-t* intervala.

Tabela 2.23: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_1 kada neto dobitak zavisi od ukupne imovine

n	t	<i>Bootstrap-t</i>	BC	BCa	Percentil	<i>Jackknife</i>
10	0,767	0,825	0,935	0,933	0,959	0,932
20	0,734	0,876	0,976	0,967	0,982	0,943
50	0,793	0,957	0,984	0,970	0,994	0,975

Izvor: Izlaz iz programskog jezika R

Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_2 zasnovane na podacima o neto dobitku i ukupnoj imovini prikazane su u Tabeli 2.24. Za male uzorke (veličine 10 i 20), *jackknife* interval je obezbedio najbolje rezultate, dok su za velike uzorke proporcije simuliranih T_2^* intervala bile najbolje.

Tabela 2.24: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_2 kada neto dobitak zavisi od ukupne imovine

n	t	<i>Bootstrap-t</i>	BC	BCa	Percentil	<i>Jackknife</i>	T_2^*
10	0,711	0,814	0,976	0,977	0,977	0,956	0,794
20	0,703	0,848	0,993	0,990	0,991	0,953	0,847
50	0,703	0,915	0,997	0,996	0,995	0,983	0,922

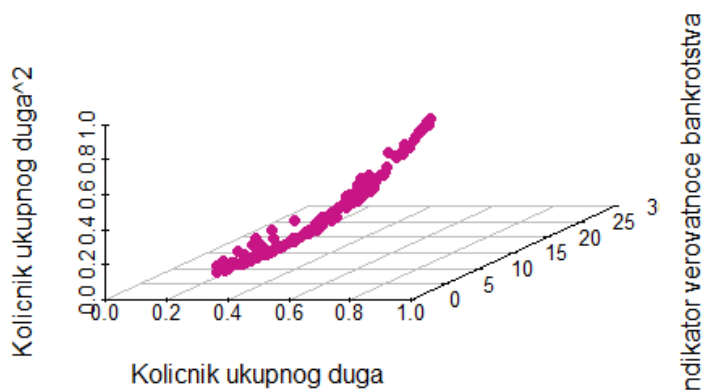
Izvor: Izlaz iz programskog jezika R

Primer 2³⁴ Proporcije simuliranih intervala za regresione koeficijente u kvadratnom regresionom modelu bazirane na podacima o indikatoru verovatnoće bankrotstva i količniku ukupnog duga u 157 kompanija u Srbiji u 2014. godini biće predstavljene u ovom primeru.

Na Slici 2.24 prikazan je dijagram raspršenosti dobijen na osnovu podataka o indikatoru verovatnoće bankrotstva i količniku ukupnog duga. Koristeći navedene podatke o indikatoru verovatnoće bankrotstva koji predstavlja zavisnu promenljivu, količniku ukupnog duga i kvadriranim vrednostima količnika ukupnog duga koji predstavljaju objašnjavajuće promenljive dobijen je kvadratni regresioni model osnovnog skupa sa sledećim regresionim koeficijentima $\beta_0 = 17,89$, $\beta_1 = -34,90$ i $\beta_2 = 21,90$. Koeficijent determinacije ovog modela je 68,3%, odnosno 68,3% varijacija promenljive indikator verovatnoće bankrotstva objašnjeno je količnikom ukupnog duga i kvadriranim vrednostima količnika ukupnog duga. Standardna devijacija slučajne greške je 2,151 i pokazuje koliko je odstupanje stvarnih vrednosti promenljive indikator verovatnoće bankrotstva od njenih prosečnih vrednosti na regresionoj ravni osnovnog skupa. S obzirom da je koeficijent determinacije kvadratnog regresionog modela veći od koeficijenta determinacije prostog linearnog regresionog modela, a standardna devijacija slučajne greške manja u poređenju sa prostim

³⁴ *Primer 2* je hipotetički primer.

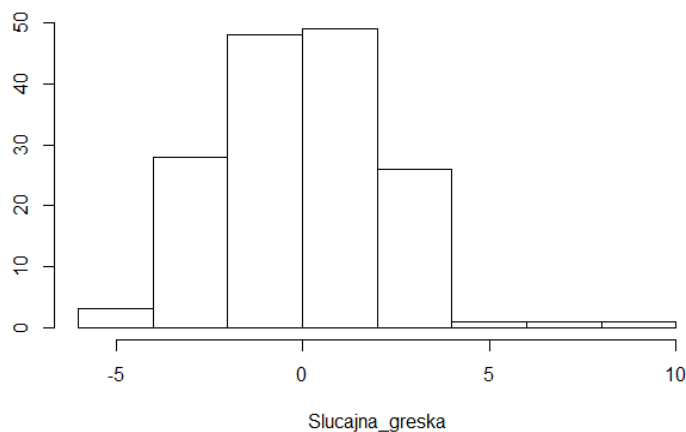
linearnim regresionim modelom, zaključeno je da analiziranim podacima više odgovara kvadratni regresioni model.



Slika 2.24: Dijagram raspršenosti na osnovu podataka o indikatoru verovatnoće bankrotstva i količniku ukupnog duga

Izvor: Izlaz iz programskog jezika R

Histogram raspodele slučajne greške, zasnovan na prethodnom modelu, prikazan je na Slici 2.25. Na osnovu histograma je uočeno da slučajna greška nije normalno raspodeljena. Takođe, Shapiro-Wilk test normalnosti je potvrdio da slučajna greška nije normalno raspodeljena (H_0 : Slučajna greška je normalno raspodeljena, a p -vrednost je približno 0,01).



Slika 2.25: Histogram raspodele slučajne greške

Izvor: Izlaz iz programskog jezika R

Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_1 zasnovane na podacima o indikatoru verovatnoće bankrotstva i količniku ukupnog duga prikazane su u Tabeli 2.25. Izvedeni su sledeći zaključci: za male uzorke (veličine 10), proporcija simuliranih BCa intervala bila je najbliža nominalnom nivou pouzdanosti, dok je kod uzoraka veličine 20, *jackknife* interval obezbedio najbolje rezultate. Utvrđeno je da su za velike uzorke (veličine 50 i 80), proporcije simuliranih *jackknife* i percentil intervala bile najbolje.

Tabela 2.25: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_1 kada indikator verovatnoće bankrotstva zavisi od količnika ukupnog duga

n	t	<i>Bootstrap-t</i>	BC	BCa	Percentil	<i>Jackknife</i>
10	0,837	0,916	0,940	0,941	0,935	0,912
20	0,782	0,838	0,847	0,862	0,799	0,873
50	0,850	0,919	0,910	0,932	0,906	0,941
80	0,892	0,968	0,955	0,973	0,946	0,981

Izvor: Izlaz iz programskog jezika R

Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_2 zasnovane na podacima o indikatoru verovatnoće bankrotstva i količniku ukupnog duga prikazane su u Tabeli 2.26. Analiza uzoraka veličine 10 pokazala je da je BCa interval obezbedio najbolje rezultate, dok je za uzorke veličine 20 proporcija simuliranih *jackknife* intervala bila najbliža 0,95. Primećeno je da su za uzorke veličine 50 najbolji rezultati ostvareni primenom intervala T_2^* , a za uzorke veličine 80 korišćenjem percentil intervala.

Tabela 2.26: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_2 kada indikator verovatnoće bankrotstva zavisi od količnika ukupnog duga

n	t	<i>Bootstrap-t</i>	BC	BCa	Percentil	<i>Jackknife</i>	T_2^*
10	0,879	0,932	0,955	0,951	0,943	0,928	0,923
20	0,832	0,860	0,861	0,882	0,817	0,890	0,855
50	0,883	0,925	0,918	0,938	0,912	0,942	0,943
80	0,922	0,967	0,962	0,975	0,952	0,980	0,964

Izvor: Izlaz iz programskog jezika R

Uzimajući u obzir rezultate predstavljene u odeljcima 2.4.2 i 2.5.5 potvrđuje se hipoteza

H_2 : Proporcije simuliranih intervala poverenja za regresioni koeficijent β_2 u kvadratnom regresionom modelu koji su zasnovani na novoj transformaciji t statistike, T_2^* , za velike uzorke, bliže su nominalnom nivou pouzdanosti u odnosu na proporcije ostalih proučavanih intervala.

2.6 R kodovi za izabrane intervale

R je besplatan programski jezik koji omogućava korišćenje niza statističkih i grafičkih tehnika. Radi na velikom broju UNIX platformi. Takođe, Windows i MacOS podržavaju upotrebu ovog programskog jezika.

2.6.1 R kodovi za intervale poverenja za regresioni koeficijent u prostom linearnom regresionom modelu

U sledećim tabelama predstavljeni su originalni kodovi za t , $bootstrap-t$, T_1 i T_2 intervale poverenja za regresioni koeficijent u prostom linearnom regresionom modelu.

Tabela 2.27: Kodovi za t interval i $bootstrap-t$ interval za regresioni koeficijent u prostom linearnom regresionom modelu

t interval poverenja	$Bootstrap-t$ interval poverenja
<pre>x<-1:500 y<-150+20*x+rnorm(500,0,2^0.5) brsim<-1000 n<-10 b1<-rep(NA,brsim) sb1<-rep(NA,brsim) l95<-rep(NA,brsim) u95<-rep(NA,brsim) for (i in 1:brsim){ uzorak<-sample(1:500,n) x1<-x[uzorak] y1<-y[uzorak] b1[i]<-lm(y1~x1)\$coefficients[2] sb1[i]<-summary(lm(y1~x1))\$coefficients[2,2] l95[i]<-b1[i]-qt(0.975,n-2)*sb1[i] u95[i]<-b1[i]+qt(0.975,n-2)*sb1[i] } mean((l95<=20)&(20<=u95))</pre>	<pre>x<-1:500 y<-150+20*x+rnorm(500,0,1) brsim<-1000 B <- 1000 n<-5 b1<-rep(NA,brsim) sb1<-rep(NA,brsim) boot_b1<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) boot_sb1<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) for(j in 1:brsim){ uzorak1 <- sample(1:500,n) x1<-x[uzorak1] y1<-y[uzorak1] ocena_b1<-lm(y1~x1) ocena_sb1<-summary(lm(y1~x1)) b1[j]<-lm(y1~x1)\$coefficients[2] sb1[j]<-summary(lm(y1~x1))\$coefficients[2,2] for (i in 1:B){ uzorak2 <-sample(uzorak1,replace=T) x2<- x[uzorak2] y2 <- y[uzorak2] ocena_boot_b1<-lm(y2~x2) boot_b1[j,i] <- ocena_boot_b1\$coefficients[2] ocena_boot_sb1<-summary(lm(y2~x2)) boot_sb1[j,i] <- ocena_boot_sb1\$coefficients[2,2] }} boot_t=(boot_b1-b1)/boot_sb1 t.low=apply(boot_t, 1, quantile, probs = c(0.025), na.rm = TRUE) t.high=apply(boot_t, 1, quantile, probs = c(0.975), na.rm = TRUE) l95<-b1 - t.high*sb1 u95<-b1 - t.low*sb1 mean((l95<=20)&(20<=u95))</pre>

Tabela 2.28: Kodovi za T_1 interval i T_2 interval za regresioni koeficijent u prostom linearnom regresionom modelu

T_1 interval poverenja	T_2 interval poverenja
<pre> x<-1:500 y<-150+20*x+rnorm(500,0,1) brsim<-1000 B <- 100 n<-10 b1<-rep(NA,brsim) sb1<-rep(NA,brsim) x_ind<-matrix(rep(NA,brsim*n),nrow=brsim,ncol=n) y_ind<-matrix(rep(NA,brsim*n),nrow=brsim,ncol=n) xprim<-matrix(rep(NA,brsim*n),nrow=brsim,ncol=n) boot_b1<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) boot_sb1<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) for(j in 1:brsim){ uzorak1 <- sample(1:500,n) x1<-x[uzorak1] y1<-y[uzorak1] sredina_x1=mean(x1) sredina_y1=mean(y1) V=(1/n)*(sum((x1-sredina_x1)^2*(y1-sredina_y1)^2))/((sum((x1- sredina_x1)^2)^2)-((2*20)/(n*n))*(sum((x1-sredina_x1)*(y1- sredina_y1)))/(sum((x1-sredina_x1)^2)))+(20*20)/(n*n)) for(k in 1:n){ x_ind[j,k]<-x1[k] y_ind[j,k]<-y1[k] xprim[j,k]<-(((x_ind[j,k]-sredina_x1)*(y_ind[j,k]- sredina_y1))/(sum((x1-sredina_x1)^2)-(20/n))/(V^0.5)) ocena_b1<-lm(y1~x1) ocena_sb1<-summary(lm(y1~x1)) b1[j]<-lm(y1~x1)\$coefficients[2] sb1[j]<-summary(lm(y1~x1))\$coefficients[2,2] for (i in 1:B){ uzorak2 <-sample(uzorak1,replace=T) x2<- x[uzorak2] y2 <- y[uzorak2] ocena_boot_b1<-lm(y2~x2) boot_b1[j,i] <- ocena_boot_b1\$coefficients[2] ocena_boot_sb1<-summary(lm(y2~x2)) boot_sb1[j,i] <- ocena_boot_sb1\$coefficients[2,2] }} xprim3=xprim^3 m3prim_kapa<-apply(xprim3,1,mean) boot_t=(boot_b1-b1)/boot_sb1 t.low=apply(boot_t, 1, quantile, probs = c(0.025), na.rm = TRUE) t.high=apply(boot_t, 1, quantile, probs = c(0.975), na.rm = TRUE) t.low1=t.low/(n^0.5) t.high1=t.high/(n^0.5) tlow_inv=(((1/3)*m3prim_kapa)^(-1))*((1+m3prim_kapa*(t.low1- (1/(6*n))*m3prim_kapa))^(1/3))-((1/3)*m3prim_kapa)^(-1)) thigh_inv=(((1/3)*m3prim_kapa)^(-1))*((1+m3prim_kapa*(t.high1- (1/(6*n))*m3prim_kapa))^(1/3))-((1/3)*m3prim_kapa)^(-1)) tlow_inv[is.na(tlow_inv)] <- 0 thigh_inv[is.na(thigh_inv)] <- 0 I95<-b1 - (n^0.5)*thigh_inv*sb1 u95<-b1 - (n^0.5)*tlow_inv*sb1 mean((I95<=20)&(20<=u95)) </pre>	<pre> x<-1:10 y<-150+20*x+rnorm(10,0,1) brsim<-5 B <- 4 n<-5 b1<-rep(NA,brsim) sb1<-rep(NA,brsim) x_ind<-matrix(rep(NA,brsim*n),nrow=brsim,ncol=n) y_ind<-matrix(rep(NA,brsim*n),nrow=brsim,ncol=n) xprim<-matrix(rep(NA,brsim*n),nrow=brsim,ncol=n) boot_b1<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) boot_sb1<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) for(j in 1:brsim){ uzorak1 <- sample(1:10,n) x1<-x[uzorak1] y1<-y[uzorak1] sredina_x1=mean(x1) sredina_y1=mean(y1) V=(1/n)*(sum((x1-sredina_x1)^2*(y1- sredina_y1)^2))/((sum((x1-sredina_x1)^2)^2)- ((2*20)/(n*n))*(sum((x1-sredina_x1)*(y1- sredina_y1)))/(sum((x1-sredina_x1)^2)))+(20*20)/(n*n)) for(k in 1:n){ x_ind[j,k]<-x1[k] y_ind[j,k]<-y1[k] xprim[j,k]<-(((x_ind[j,k]-sredina_x1)*(y_ind[j,k]- sredina_y1))/(sum((x1-sredina_x1)^2)-(20/n))/(V^0.5)) ocena_b1<-lm(y1~x1) ocena_sb1<-summary(lm(y1~x1)) b1[j]<-lm(y1~x1)\$coefficients[2] sb1[j]<-summary(lm(y1~x1))\$coefficients[2,2] for (i in 1:B){ uzorak2 <-sample(uzorak1,replace=T) x2<- x[uzorak2] y2 <- y[uzorak2] ocena_boot_b1<-lm(y2~x2) boot_b1[j,i] <- ocena_boot_b1\$coefficients[2] ocena_boot_sb1<-summary(lm(y2~x2)) boot_sb1[j,i] <- ocena_boot_sb1\$coefficients[2,2] }} xprim3=xprim^3 m3prim_kapa<-apply(xprim3,1,mean) boot_t=(boot_b1-b1)/boot_sb1 t.low=apply(boot_t, 1, quantile, probs = c(0.025), na.rm = TRUE) t.high=apply(boot_t, 1, quantile, probs = c(0.975), na.rm = TRUE) t.low1=t.low/(n^0.5) t.high1=t.high/(n^0.5) tlow_inv=(((2/3)*(n^(-1/2))*m3prim_kapa)^(-1))*log(((2/3)*(n^(- 1/2))*m3prim_kapa)*(t.low1-(1/(6*n))*m3prim_kapa)+1) thigh_inv=(((2/3)*(n^(-1/2))*m3prim_kapa)^(-1))*log(((2/3)*(n^(- 1/2))*m3prim_kapa)*(t.high1-(1/(6*n))*m3prim_kapa)+1) tlow_inv[is.na(tlow_inv)] <- 0 thigh_inv[is.na(thigh_inv)] <- 0 I95<-b1 - (n^0.5)*thigh_inv*sb1 u95<-b1 - (n^0.5)*tlow_inv*sb1 mean((I95<=20)&(20<=u95)) </pre>

2.6.2 R kodovi za intervale poverenja za regresione koeficijente u kvadratnom regresionom modelu

U sledećim tabelama navedeni su originalni kodovi za percentil, *jackknife*, BCa i T_2^* intervale poverenja za regresioni koeficijent u kvadratnom regresionom modelu.

Tabela 2.29: Kodovi za percentil interval i *jackknife* interval za regresione koeficijente u kvadratnom regresionom modelu

Percentil interval poverenja	<i>Jackknife</i> interval poverenja
<pre>x<-1:500 x2<-x^2 y<-100+10*x+2*x2+rweibull(500,0.5,1) brsim<-1000 B <- 1000 n<-10 b1<-rep(NA,brsim) b2<-rep(NA,brsim) boot_b1<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) boot_b2<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) for(j in 1:brsim){ uzorak1 <- sample(1:500,n) x1<-x[uzorak1] x21<-x2[uzorak1] y1<-y[uzorak1] ocena_b<-lm(y1~x1+x21) b1[j]<-ocena_b\$coefficients[2] b2[j]<-ocena_b\$coefficients[3] for (i in 1:B){ uzorak2 <-sample(uzorak1,replace=T) x1b<- x[uzorak2] x21b <- x2[uzorak2] y1b <- y[uzorak2] ocena_boot_b<-lm(y1b~x1b+x21b) boot_b1[j,i] <- ocena_boot_b\$coefficients[2] boot_b2[j,i] <- ocena_boot_b\$coefficients[3] } } t.lowb1=apply(boot_b1, 1, quantile, probs = c(0.025), na.rm = TRUE) t.highb1=apply(boot_b1, 1, quantile, probs = c(0.975), na.rm = TRUE) l95_b1<-t.lowb1 u95_b1<-t.highb1 mean((l95_b1<=10)&(10<=u95_b1)) t.lowb2=apply(boot_b2, 1, quantile, probs = c(0.025), na.rm = TRUE) t.highb2=apply(boot_b2, 1, quantile, probs = c(0.975), na.rm = TRUE) l95_b2<-t.lowb2 u95_b2<-t.highb2 mean((l95_b2<=2)&(2<=u95_b2))</pre>	<pre>x<-1:500 x2<-x^2 y<-100+10*x+2*x2+rweibull(500,5,1) brsim<-1000 n<-80 b1<-rep(NA,brsim) b2<-rep(NA,brsim) partial_b1<-matrix(rep(NA,brsim*n),nrow=brsim,ncol=n) partial_b2<-matrix(rep(NA,brsim*n),nrow=brsim,ncol=n) pseudo_b1<-matrix(rep(NA,brsim*n),nrow=brsim,ncol=n) pseudo_b2<-matrix(rep(NA,brsim*n),nrow=brsim,ncol=n) for(j in 1:brsim){ uzorak1 <- sample(1:500,n) x1<-x[uzorak1] x21<-x2[uzorak1] y1<-y[uzorak1] ocena_b<-lm(y1~x1+x21) b1[j]<-ocena_b\$coefficients[2] b2[j]<-ocena_b\$coefficients[3] for(i in 1:n){ uzorak_jack<-sample(uzorak1[-i],n-1) x_jack<-x[uzorak_jack] x2_jack<-x2[uzorak_jack] y_jack<-y[uzorak_jack] partial_ocena_b<-lm(y_jack~x_jack+x2_jack) partial_b1[j,i]<-partial_ocena_b\$coefficients[2] partial_b2[j,i]<-partial_ocena_b\$coefficients[3] pseudo_b1[j,i]<-n*b1[j]-(n-1)*partial_b1[j,i] pseudo_b2[j,i]<-n*b2[j]-(n-1)*partial_b2[j,i] } } jack_b1<-apply(pseudo_b1,1,mean) jack_se_b1<-sqrt((apply(((pseudo_b1-jack_b1)^2),1,sum))/(n*(n-1))) l95_b1<-jack_b1-qt(0.975,n-3)*jack_se_b1 u95_b1<-jack_b1+qt(0.975,n-3)*jack_se_b1 mean((l95_b1<=10)&(10<=u95_b1)) jack_b2<-apply(pseudo_b2,1,mean) jack_se_b2<-sqrt((apply(((pseudo_b2-jack_b2)^2),1,sum))/(n*(n-1))) l95_b2<-jack_b2-qt(0.975,n-3)*jack_se_b2 u95_b2<-jack_b2+qt(0.975,n-3)*jack_se_b2 mean((l95_b2<=2)&(2<=u95_b2))</pre>

Tabela 2.30: Kodovi za BCa interval i T_2^* interval za regresione koeficijente u kvadratnom regresionom modelu

BCa interval poverenja	T_2^* interval poverenja
<pre> x<-1:500 x2<-x^2 y<-100+10*x+2*x2+rweibull(500,5,1) brsim<-1000 B<-1000 n<-80 alfa=0.05 b1<-rep(NA,brsim) b2<-rep(NA,brsim) b1_jack<-matrix(rep(NA,brsim*n),nrow=brsim,ncol=n) b2_jack<-matrix(rep(NA,brsim*n),nrow=brsim,ncol=n) b1_tacka<-rep(NA,brsim) b2_tacka<-rep(NA,brsim) boot_b1<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) boot_b2<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) for(j in 1:brsim){ uzorak1 <- sample(1:500,n) x1<-x[uzorak1] x21<-x2[uzorak1] y1<-y[uzorak1] ocena_b<-lm(y1~x1+x21) b1[j]<-ocena_b\$coefficients[2] b2[j]<-ocena_b\$coefficients[3] for(k in 1:n){ uzorak_jack<-sample(uzorak1[-k],n-1) x_jack<-x[uzorak_jack] x2_jack<-x2[uzorak_jack] y_jack<-y[uzorak_jack] ocena_b_jack<-lm(y_jack~x_jack+x2_jack) b1_jack[j,k]=ocena_b_jack\$coefficients[2] b2_jack[j,k]=ocena_b_jack\$coefficients[3] } b1_tacka=apply(b1_jack,1,mean) b2_tacka=apply(b2_jack,1,mean) for (i in 1:B){ uzorak2 <-sample(uzorak1,replace=T) x1b <- x[uzorak2] x21b <- x2[uzorak2] y1b <- y[uzorak2] ocena_boot_b<-lm(y1b~x1b+x21b) boot_b1[j,i] <- ocena_boot_b\$coefficients[2] boot_b2[j,i] <- ocena_boot_b\$coefficients[3] } } a_kapa_b1=(apply((b1_tacka- b1_jack)^3,1,sum))/(6*(apply((b1_tacka- b1_jack)^2,1,sum))^(3/2)) a_kapa_b2=(apply((b2_tacka- b2_jack)^3,1,sum))/(6*(apply((b2_tacka- b2_jack)^2,1,sum))^(3/2)) br_manji_b1=rowSums(boot_b1<b1, na.rm=TRUE) cumbr_b1=br_manji_b1/B z0_kapa_b1=qnorm(cumbr_b1) br_manji_b2=rowSums(boot_b2<b2, na.rm=TRUE) cumbr_b2=br_manji_b2/B z0_kapa_b2=qnorm(cumbr_b2) alfa1_b1=pnorm(z0_kapa_b1+((z0_kapa_b1+qnorm(alfa/2))/(1- a_kapa_b1*(z0_kapa_b1+qnorm(alfa/2)))))) alfa2_b1=pnorm(z0_kapa_b1+((z0_kapa_b1+qnorm(1-alfa/2))/(1- a_kapa_b1*(z0_kapa_b1+qnorm(1-alfa/2)))))) </pre>	<pre> x<-1:500 x2<-x^2 y<-100+10*x+2*x2+rgamma(500,2,2) brsim<-1050 B<-1000 n<-80 beta2=2 b2<-rep(NA,brsim) sb2<-rep(NA,brsim) boot_b2<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) boot_sb2<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) part_down<- rep(NA,brsim) part_downsq<- rep(NA,brsim) V_1 <- rep(NA,brsim) V_2 <- rep(NA,brsim) V <- rep(NA,brsim) xprim_up<- matrix(rep(NA,brsim*(n^2- n)),nrow=brsim,ncol=n^2-n) xprim<- matrix(rep(NA,brsim*(n^2-n)),nrow=brsim,ncol=n^2- n) for(j in 1:brsim){ print(paste("j = ", j)) uzorak1<-sample(x,n) x1<-x[uzorak1] x21<-x2[uzorak1] y1<-y[uzorak1] x1_bar=mean(x1) y1_bar=mean(y1) x1_barkv=x1_bar^2 ocena_b<-lm(y1~x1+x21) ocena_sb<-summary(lm(y1~x1+x21)) b2[j]<-ocena_b\$coefficients[3] sb2[j]<-ocena_sb\$coefficients[3,2] for (i in 1:B){ uzorak2 <-sample(uzorak1,replace=T) x1b <- x[uzorak2] x21b <- x2[uzorak2] y1b <- y[uzorak2] ocena_boot_b<-lm(y1b~x1b+x21b) boot_b2[j,i] <- ocena_boot_b\$coefficients[3] } ocena_boot_sb<- summary(lm(y1b~x1b+x21b)) b_sb2 <- try(ocena_boot_sb\$coefficients[3,2]) if (class(b_sb2) == "try-error") { print(paste("Greska: i = ", i, "--- j = ", j)) b2[j] <- NA sb2[j] <- NA } boot_b2[j,i] <- NA boot_sb2[j,i] <- NA } else { boot_sb2[j,i] <- b_sb2 } } comb<- expand.grid(1:n, 1:n) comb_ind<- comb[1] != comb[2] comb<- comb[comb_ind,] comb </pre>

Tabela 2.30: Kodovi za BCa interval i T_2^* interval za regresione koeficijente u kvadratnom regresionom modelu (nastavak)

<pre> alfa1_b1=pnorm(z0_kapa_b1+((z0_kapa_b1+qnorm(alfa/2))/(1- a_kapa_b1*(z0_kapa_b1+qnorm(alfa/2)))))) alfa2_b1=pnorm(z0_kapa_b1+((z0_kapa_b1+qnorm(1- alfa/2))/(1-a_kapa_b1*(z0_kapa_b1+qnorm(1-alfa/2)))))) alfa1_b2=pnorm(z0_kapa_b2+((z0_kapa_b2+qnorm(alfa/2))/(1- a_kapa_b2*(z0_kapa_b2+qnorm(alfa/2)))))) alfa2_b2=pnorm(z0_kapa_b2+((z0_kapa_b2+qnorm(1- alfa/2))/(1-a_kapa_b2*(z0_kapa_b2+qnorm(1-alfa/2)))))) low_b1=apply(boot_b1, 1, quantile, probs = c(alfa1_b1), na.rm = TRUE) lb1=diag(low_b1) high_b1=apply(boot_b1, 1, quantile, probs = c(alfa2_b1), na.rm = TRUE) hb1=diag(high_b1) mean((lb1<=10)&(10<=hb1)) low_b2=apply(boot_b2, 1, quantile, probs = c(alfa1_b2), na.rm = TRUE) lb2=diag(low_b2) high_b2=apply(boot_b2, 1, quantile, probs = c(alfa2_b2), na.rm = TRUE) hb2=diag(high_b2) mean((lb2<=2)&(2<=hb2)) </pre>	<pre> part_down<- ((sum ((x1 - x1_bar)^2)) * (sum ((x21 - x1_barkv)^2)) - (sum ((x1 - x1_bar) * (x21 - x1_barkv)))^2) part_downsq<- part_down^2 V_1<-0 V_2<-0 for (s_comb in 1:nrow(comb)) { s <- comb[s_comb, 1] r <- comb[s_comb, 2] pV_1 <- ((x1[s] - x1_bar)^2 * (y1[s] - y1_bar)^2 * (x1[r] - x1_bar)^4 * (x1[s] - x1[r])^2) pV_2 <- ((x1[s] - x1_bar) * (y1[s] - y1_bar) * (x1[r] - x1_bar)^2 * (x1[s] - x1[r])) V_1<-V_1+pV_1 V_2<-V_2+pV_2 V <- (1/((n^2)-n))*(V_1/part_downsq) - ((2*beta2)/((n^2)*((n- 1)^2)))*(V_2/part_down) + ((beta2^2)/((n^2)*((n-1)^2))) } for (s_comb in 1:nrow(comb)) { s <- comb[s_comb, 1] r <- comb[s_comb, 2] xprim_up [j,s_comb] <- (x1[s] - x1_bar) * (y1[s] - y1_bar) * (x1[r] - x1_bar)^2 * (x1[s] - x1[r]) xprim [j,s_comb] <- (xprim_up [j,s_comb] / part_down) - (beta2/(n^2 - n))/ (V^0.5) } b2_na_ind <- which(is.na(b2)) b2 <- as.numeric(na.omit(b2)) sb2 <- as.numeric(na.omit(sb2)) boot_b2 <- na.omit(boot_b2) boot_sb2 <- na.omit(boot_sb2) ind<- sample(seq(b2), 1000, replace = FALSE) b2 <- b2[ind] sb2 <- sb2[ind] boot_b2 <- boot_b2[ind,] boot_sb2 <- boot_sb2[ind,] if (length(b2_na_ind) > 0) { xprim_up<- xprim_up[-b2_na_ind,] xprim<- xprim[-b2_na_ind,] } xprim_up<- xprim_up[ind,] xprim<- xprim[ind,] xprim3=xprim^3 m3<-apply(xprim3,1,mean) boot_t=(boot_b2-b2)/boot_sb2 t.low=apply(boot_t, 1, quantile, probs = c(0.025), na.rm = TRUE) t.high=apply(boot_t, 1, quantile, probs = c(0.975), na.rm = TRUE) tl=t.low/((n*(n-1))^0.5) th=t.high/((n*(n-1))^0.5) a1=8+12*(tl-m3/(6*n*(n-1))) a_low=sign(a1)*abs(a1)^(1/3) a2=8+12*(th-m3/(6*n*(n-1))) a_high=sign(a2)*abs(a2)^(1/3) tlow_inv=a_low-2 thigh_inv=a_high-2 tlow_inv[is.na(tlow_inv)] <- 0 thigh_inv[is.na(thigh_inv)] <- 0 l95 <- b2-((n*(n-1))^0.5)*thigh_inv*sb2 u95 <- b2-((n*(n-1))^0.5)*tlow_inv*sb2 mean((l95<=2)&(2<=u95)) </pre>
--	---

III INTERVALI POVERENJA ZA MERE DISPERZIJE

Treći deo disertacije odnosi se na intervale poverenja za mere disperzije: varijansu i koeficijent varijacije. Pitanje koje se postavlja je sledeće: kada će proporcije simuliranih intervala biti najbliže nominalnom nivou pouzdanosti kada osnovni skup iz kog se bira uzorak nema normalnu raspodelu, već sledi neku asimetričnu raspodelu. Na početku trećeg dela biće predstavljeni intervale poverenja za varijansu (apsolutna mera disperzije) i koeficijent varijacije (relativna mera disperzije). Nakon toga biće prikazani rezultati simulacionih studija, kao i rezultati dobijeni na realnim podacima. Na kraju će biti navedeni kodovi iz programskog jezika *R* za izabrane intervale.

3.1 Intervali poverenja za varijansu

Varijansa osnovnog skupa je najčešće korišćena apsolutna mera disperzije. Postojeći intervale poverenja za apsolutne mere disperzije osnovnog skupa konstruisani su korišćenjem χ^2 statistike. Ovi intervale su adekvatni ako osnovni skup sledi normalnu raspodelu. Međutim, cilj je utvrditi koje intervalne ocene ostvaruju pouzdanost najbližu nominalnoj ukoliko osnovni skup ne sledi normalnu raspodelu, već neku od asimetričnih raspodela. Postoje neke alternative intervalima poverenja zasnovanim na χ^2 statistici koje mogu biti korišćene u takvim situacijama.

Veoma mali broj autora se bavio intervalima poverenja koji su robusni na odstupanja od normalnosti ili na prisustvo ekstremnih vrednosti. Barham i Jeyaratnam (1999) su došli do zaključka da je interval poverenja za varijansu, koji je zasnovan na oceni varijabiliteta koja predstavlja linearnu kombinaciju poređanih statistika (engl. *L*-estimate of scale), robusan. Hummel et al. (2005) su predložili dva intervala poverenja za standardnu devijaciju osnovnog skupa. Reč je o intervalu zasnovanom na logaritamskoj transformaciji i intervalu sa korigovanim stepenima slobode. Čojbašić i Tomović (2007) su primenili metode ponovljenih uzoraka za konstruisanje intervala poverenja za varijansu osnovnog skupa, dok su Banik et al. (2014) poredili osobine intervala poverenja za standardnu devijaciju u slučaju simetričnih i asimetričnih raspodela. Burch (2014) je proučavao različite metode za ocenjivanje koeficijenta spljoštenosti, da bi zatim dobijene ocene koristio prilikom konstruisanja intervala poverenja za varijansu asimetričnih raspodela. Bonett (2006), Niwitpong i Kirdwichai (2008), Čojbašić i Lončar (2011) i Abu-Shawiesh et al. (2011) su se takođe bavili intervalima poverenja za standardnu devijaciju odnosno varijansu u slučajevima kada osnovni skup ne sledi normalnu raspodelu ili kada su prisutne ekstremne vrednosti.

U odeljku 3.1 biće predstavljeni intervale poverenja za varijansu osnovnog skupa.

3.1.1 χ^2 interval poverenja

Neka su X_1, \dots, X_n nezavisne slučajne promenljive iz normalne raspodele sa parametrima μ i σ , tj. $X_i \sim N(\mu, \sigma^2)$. Ako je \bar{X} sredina uzorka, a $S^2 = (1/(n-1)) \sum_{i=1}^n (X_i - \bar{X})^2$ varijansa uzorka, onda sledi da statistika $(n-1)S^2 / \sigma^2$ ima χ^2 raspodelu sa $n-1$ stepeni slobode. Interval poverenja za varijansu skupa pouzdanosti $(1-\alpha)100\%$, zasnovan na prethodnoj statistici, sledećeg je oblika:

$$I_{\sigma^2} = \left[\frac{(n-1)S^2}{\chi^2_{1-\alpha/2, (n-1)}}, \frac{(n-1)S^2}{\chi^2_{\alpha/2, (n-1)}} \right], \quad (3.1)$$

gde $\chi^2_{1-\alpha/2}$ i $\chi^2_{\alpha/2}$ predstavljaju $1-\alpha/2$ i $\alpha/2$ percentile χ^2 raspodele sa $n-1$ stepeni slobode. Interval zasnovan na χ^2 statistici (3.1) osetljiv je na minimalna narušavanja pretpostavke o normalnosti. U nastavku će biti razmatrani alternativni intervali poverenja, koji se mogu koristiti kada je narušena pretpostavka o normalnosti.

3.1.2 Bonett-ov interval poverenja

Neka su X_1, \dots, X_n neprekidne, nezavisne i identično raspodeljene slučajne promenljive sa $E(X_i) = \mu$, $Var(X_i) = \sigma^2$ i konačnim četvrtim momentom. Bonett (2006) je predložio sledeću ocenu koeficijenta spljoštenosti, γ_4 :

$$\hat{\gamma}_4 = n \frac{\sum_{i=1}^n (X_i - m)^4}{\left(\sum_{i=1}^n (X_i - \bar{X})^2 \right)^2},$$

gde je sa m označena Bonett-ova odsečena sredina koja eliminiše procenat vrednosti koji se računa po formuli $1/\lceil 2(n-4)^{1/2} \rceil$. Interval poverenja za varijansu skupa (čiji je nivo pouzdanosti $(1-\alpha)100\%$) zapisuje se u sledećoj formi (Bonett, 2006, str. 777):

$$I_{\sigma^2} = \left[\exp\left(\ln(cS^2) - Z_{1-\alpha/2} se\right), \exp\left(\ln(cS^2) + Z_{1-\alpha/2} se\right) \right], \quad (3.2)$$

gde je konstanta c oblika $c = n/(n - Z_{1-\alpha/2})$, S^2 je varijansa uzorka, $Z_{1-\alpha/2}$ je $1-\alpha/2$ percentil standardizovane normalne raspodele, a standardna greška od $\ln(S^2)$ je oblika $se = c \left[\left\{ \hat{\gamma}_4 (n-3) / n \right\} / (n-1) \right]^{1/2}$.

3.1.3 Steve-ov interval poverenja zasnovan na normalnoj aproksimaciji u slučaju velikih uzoraka

Steve-ov interval poverenja za varijansu skupa pouzdanosti $(1-\alpha)100\%$ zasnovan na normalnoj aproksimaciji u slučaju velikih uzoraka je sledećeg oblika (Steve, 1990):

$$I_{\sigma^2} = \left[\frac{S^2}{1 - Z_{\alpha/2} \sqrt{\frac{\hat{\gamma}_3 - 1}{n}}}, \frac{S^2}{1 + Z_{\alpha/2} \sqrt{\frac{\hat{\gamma}_3 - 1}{n}}} \right], \quad (3.3)$$

gde je S^2 varijansa uzorka, $Z_{\alpha/2}$ je $\alpha/2$ percentil standardizovane normalne raspodele, a

$\hat{\gamma}_3 = n \frac{\sum_{i=1}^n (X_i - \bar{X})^4}{\left(\sum_{i=1}^n (X_i - \bar{X})^2\right)^2}$ je ocena koeficijenta spljoštenosti.

3.1.4 Interval poverenja zasnovan na logaritamskoj transformaciji

Raspodela varijanse uzorka, S^2 je znatno asimetrična za malo n . U cilju korigovanja asimetrije, Hummel et al. (2005) su modifikovali interval (3.3) i predložili $(1-\alpha)100\%$ interval poverenja za varijansu osnovnog skupa:

$$I_{\sigma^2} = \left[S^2 \exp\left(Z_{\alpha/2} \sqrt{\frac{\hat{\gamma}_3 - 1}{n}} \right), S^2 \exp\left(-Z_{\alpha/2} \sqrt{\frac{\hat{\gamma}_3 - 1}{n}} \right) \right], \quad (3.4)$$

gde $Z_{\alpha/2}$ predstavlja $\alpha/2$ percentil Z raspodele, a $\hat{\gamma}_3$ je ocena koeficijenta spljoštenosti. Ovaj interval će u simulacijama biti imenovan sa LOG interval.

3.1.5 Interval poverenja sa korigovanim stepenima slobode

U cilju konstruisanja intervala poverenja sa korigovanim stepenima slobode, Hummel et al. (2005) su korigovali broj stepeni slobode χ^2 raspodele, pri čemu su dobili sledeći $(1-\alpha)100\%$ interval poverenja za varijansu osnovnog skupa:

$$I_{\sigma^2} = \left[\frac{\hat{r}S^2}{\chi^2_{1-\alpha/2, \hat{r}}}, \frac{\hat{r}S^2}{\chi^2_{\alpha/2, \hat{r}}} \right], \quad (3.5)$$

gde je $\hat{r} = \frac{2n}{\hat{\gamma}_e + \left(\frac{2n}{n-1}\right)}$ ocena broja stepeni slobode, dok je $\hat{\gamma}_e$ ocena koeficijenta spljoštenosti, koja

se definiše kao $\hat{\gamma}_e = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \frac{\sum_{i=1}^n (X_i - \bar{X})^4}{S^4} - \frac{3(n-1)^2}{(n-2)(n-3)}$. Ukoliko je slučajan uzorak

izabran iz normalne raspodele, tada je $r = n-1$ i (3.5) se svodi na (3.1). Ovaj interval će u simulacijama biti označen sa ADF interval.

3.2 Intervali poverenja za koeficijent varijacije

Koeficijent varijacije predstavlja važnu relativnu meru disperzije, koja nema svoju jedinicu mere. Jedan od postojećih metoda za intervalno ocenjivanje koeficijenta varijacije osnovnog skupa zasnovan je na t -statistici i adekvatan je ukoliko osnovni skup sledi normalnu raspodelu. Pitanje koje se postavlja je koji intervali poverenja za koeficijent varijacije su odgovarajući kada osnovni skup sledi neku asimetričnu raspodelu.

Veliki broj autora bio je zainteresovan za proučavanje intervala poverenja za koeficijent varijacije osnovnog skupa (McKay, 1932; Miller, 1991; Curto & Pinto, 2009; Albatineh et al., 2014). McKay (1932) se bavio utvrđivanjem raspodele statistike CV , dok je Miller (1991) izveo asimptotski normalnu statistiku testa za testiranje hipoteza o koeficijentu varijacije normalne raspodele. Rezultati do kojih su došli se mogu iskoristiti za konstruisanje intervala poverenja za koeficijent varijacije. Curto i Pinto (2009) su izveli asimptotsku raspodelu statistike CV , na osnovu koje su konstruisali interval poverenja za koeficijent varijacije. Albatineh et al. (2014) su se bavili intervalnim ocenjivanjem koeficijenta varijacije koristeći rangirano uzorkovanje. Panichkitkosolkul (2009) je poredio osobine nekoliko intervala poverenja za koeficijent varijacije u slučaju kada je osnovni skup normalno raspodeljen. Međutim, retki su autori koji su analizirali proporcije simuliranih intervala za koeficijent varijacije kada osnovni skup sledi neku asimetričnu raspodelu (Albatineh et al., 2014).

U nastavku će biti razmatrani postojeći intervali poverenja za koeficijent varijacije osnovnog skupa. Zatim, biće analizirani intervali poverenja zasnovani na metodama ponovljenih uzoraka, rangiranom uzorkovanju, metodama ponovljenih uzoraka u kombinaciji sa rangiranim uzorkovanjem i parcijalnom uzorkovanju. Na kraju, biće predložene transformacije postojećih intervala poverenja koristeći odsečenu sredinu, koje su zasnovane na prostom slučajnom uzorkovanju i rangiranom uzorkovanju, kao i *bootstrap* transformacije postojećih intervala poverenja (Ivković & Rajić, 2019).

3.2.1 Postojeći intervali poverenja

Predmet analize ovog dela biće intervali poverenja za koeficijent varijacije osnovnog skupa koji se navode u literaturi (videti Miller, 1991; Curto & Pinto, 2009; McKay, 1932). Potrebno je napomenuti da će u daljem tekstu ocena koeficijenta varijacije osnovnog skupa biti označena sa CV .

3.2.1.1 Interval poverenja zasnovan na t statistici

Neka su X_1, \dots, X_n nezavisne i identično raspodeljene slučajne promenljive iz normalne raspodele sa parametrima μ i σ . Ocena standardne devijacije statistike CV označena je sa S_{CV} . Pod pretpostavkom da osnovni skup sledi normalnu raspodelu, statistika $t = \frac{CV - CV}{S_{CV}}$ sledi Student-ovu raspodelu sa $n-1$ stepeni slobode, gde CV predstavlja koeficijent varijacije osnovnog skupa. Interval poverenja za koeficijent varijacije pouzdanosti $(1-\alpha)100\%$ ima sledeću formu i u daljem tekstu će biti označen sa t interval (Abdi, 2010, str. 170):

$$I_{CV} = \left[CV - t_{1-\alpha/2, n-1} S_{CV}, CV + t_{1-\alpha/2, n-1} S_{CV} \right], \quad (3.6)$$

gde $t_{1-\alpha/2}$ predstavlja $1-\alpha/2$ percentil t raspodele sa $n-1$ stepeni slobode.

Interval poverenja zasnovan na t statistici je osetljiv na minimalna odstupanja od pretpostavke o normalnoj raspodeli osnovnog skupa.

3.2.1.2 Miller-ov interval poverenja

Miller (1991) je pokazao da raspodela statistike $CV = S / \bar{X}$ teži normalnoj raspodeli sa sredinom σ / μ i standardnom devijacijom $\frac{1}{(n-1)}(\sigma / \mu)^2 \left[0,5 + (\sigma / \mu)^2 \right]$, ako je raspodela osnovnog skupa normalna sa sredinom μ i standardnom devijacijom σ . U tom slučaju, $(1-\alpha)100\%$ interval poverenja za koeficijent varijacije osnovnog skupa može biti predstavljen na sledeći način (Miller, 1991, str. 3354):

$$I_{CV} = \left[CV - Z_{1-\alpha/2} \sqrt{\frac{1}{(n-1)} CV^2 (0,5 + CV^2)}, CV + Z_{1-\alpha/2} \sqrt{\frac{1}{(n-1)} CV^2 (0,5 + CV^2)} \right], \quad (3.7)$$

gde $Z_{1-\alpha/2}$ predstavlja $1-\alpha/2$ percentil standardizovane normalne raspodele.

3.2.1.3 Curto-ov i Pinto-ov interval poverenja

Neka su X_1, \dots, X_n nezavisne i identično raspodeljene slučajne promenljive. Curto-ov i Pinto-ov (engl. Curto & Pinto's) $(1-\alpha)100\%$ interval poverenja za koeficijent varijacije osnovnog skupa je sledećeg oblika (Curto & Pinto, 2009, str. 24):

$$I_{CV} = \left[CV - Z_{1-\alpha/2} \sqrt{(1/n) (CV^4 + 0,5 CV^2)}, CV + Z_{1-\alpha/2} \sqrt{(1/n) (CV^4 + 0,5 CV^2)} \right], \quad (3.8)$$

gde $Z_{1-\alpha/2}$ predstavlja $1-\alpha/2$ percentil standardizovane normalne raspodele. U nastavku će ovaj interval biti označen sa C&P interval.

3.2.1.4 McKay-ev interval poverenja

Za konstrukciju McKay-evog intervala poverenja koristi se χ^2 raspodela. Interval poverenja određen sa nivoom pouzdanosti $(1-\alpha)100\%$ za koeficijent varijacije osnovnog skupa predstavljen je sledećom formulom (McKay, 1932, str. 697):

$$I_{CV} = \left[CV \left[\left(\frac{\chi_{1-\alpha/2, n-1}^2}{n} - 1 \right) CV^2 + \frac{\chi_{1-\alpha/2, n-1}^2}{n-1} \right]^{-1/2}, CV \left[\left(\frac{\chi_{\alpha/2, n-1}^2}{n} - 1 \right) CV^2 + \frac{\chi_{\alpha/2, n-1}^2}{n-1} \right]^{-1/2} \right], \quad (3.9)$$

gde $\chi_{1-\alpha/2}^2$ i $\chi_{\alpha/2}^2$ predstavljaju $1-\alpha/2$ i $\alpha/2$ percentile χ^2 raspodele.

3.2.2 Intervali poverenja zasnovani na metodama ponovljenih uzoraka

U cilju konstruisanja intervala poverenja zasnovanih na metodama ponovljenih uzoraka, prvo je potrebno generisati B *bootstrap* uzoraka iz originalnog uzorka, a zatim u svakom od B *bootstrap* uzoraka izračunati vrednosti statistike CV , u oznaci CV^* .

Bootstrap-t interval poverenja

Za potrebe formiranja *bootstrap-t* intervala poverenja, u svakom od B *bootstrap* uzoraka potrebno je izračunati vrednost statistike $T^* = \frac{CV^* - CV}{S_{CV^*}}$, gde S_{CV^*} predstavlja ocenu standardne devijacije statistike CV^* . Interval poverenja određen sa nivoom pouzdanosti $(1-\alpha)100\%$ za koeficijent varijacije skupa predstavljen je sledećom formulom:

$$I_{CV} = \left[CV - \hat{t}_{1-\alpha/2} S_{CV}, CV - \hat{t}_{\alpha/2} S_{CV} \right], \quad (3.10)$$

gde $\hat{t}_{1-\alpha/2}$ i $\hat{t}_{\alpha/2}$ predstavljaju $1-\alpha/2$ i $\alpha/2$ percentile raspodele statistike T^* , a S_{CV} je ocena standardne devijacije statistike CV .

Percentil interval poverenja

Za konstrukciju percentil intervala poverenja za koeficijent varijacije, prvo je potrebno u neopadajućem poretku rangirati vrednosti statistike CV u *bootstrap* uzorcima. Zatim je potrebno odrediti $\alpha/2$ i $1-\alpha/2$ percentile raspodele *bootstrap* replika, koji su označeni sa $CV_{\alpha/2}^*$ i $CV_{1-\alpha/2}^*$. Interval poverenja određen sa nivoom pouzdanosti $(1-\alpha)100\%$ za koeficijent varijacije skupa može biti definisan u sledećoj formi:

$$I_{CV} = \left[CV_{\alpha/2}^*, CV_{1-\alpha/2}^* \right]. \quad (3.11)$$

BCa interval poverenja

BCa interval poverenja određen sa nivoom pouzdanosti $(1-\alpha)100\%$ za koeficijent varijacije može se zapisati u sledećem obliku:

$$I_{CV} = \left[CV_{\alpha_1}^*, CV_{\alpha_2}^* \right]. \quad (3.12)$$

Postupak konstrukcije ovog intervala i izračunavanje vrednosti α_1 i α_2 detaljno su opisani u odeljku 1.3.

BC interval poverenja

BC interval poverenja određen sa nivoom pouzdanosti $(1-\alpha)100\%$ definiše se na sledeći način:

$$I_{CV} = \left[CV_{\alpha_1}^*, CV_{\alpha_2}^* \right], \quad (3.13)$$

gde se vrednosti α_1 i α_2 za percentile $CV_{\alpha_1}^*$ i $CV_{\alpha_2}^*$ izračunavaju na identičan način kao u slučaju BCa intervala.

Jackknife interval poverenja

U cilju konstruisanja *jackknife* intervala poverenja u svakom *jackknife* uzorku je potrebno izračunati vrednost statistike CV , koja se označava sa $CV_{(i)}$. Interval poverenja za koeficijent varijacije pouzdanosti $(1-\alpha)100\%$ može se zapisati na sledeći način:

$$I_{CV} = \left[\tilde{CV} - t_{1-\alpha/2, n-1} S_{jack}, \tilde{CV} + t_{1-\alpha/2, n-1} S_{jack} \right], \quad (3.14)$$

gde su:

$$\tilde{CV}_i = nCV - (n-1)CV_{(i)}, \quad \tilde{CV} = \sum_{i=1}^n \tilde{CV}_i / n, \quad S_{jack} = \left\{ \sum_{i=1}^n \left(\tilde{CV}_i - \tilde{CV} \right)^2 / [(n-1)n] \right\}^{1/2} \quad \text{i} \quad t_{1-\alpha/2} \text{ je}$$

$1-\alpha/2$ percentil t raspodele sa $n-1$ stepeni slobode.

3.2.3 Intervali poverenja zasnovani na rangiranom uzorkovanju

Moguće je konstruisati intervale poverenja za koeficijent varijacije koji su zasnovani na proceduri balansiranog rangiranog uzorkovanja (Ivković & Rajić, 2019, str. 5-6) koja je detaljno opisana u odeljku 1.4.1. U svrhe istraživanja biće razmatrani Miller-ov, Curto-ov i Pinto-ov i McKay-ev interval poverenja zasnovani na rangiranom uzorkovanju (u tabelama koje slede u odeljku 3.3 ovi intervali biće označeni sa RSS Miller-ov interval, RSS C&P interval i RSS McKay-ev interval).

3.2.4 Intervali poverenja zasnovani na metodama ponovljenih uzoraka u kombinaciji sa rangiranim uzorkovanjem

U svrhe istraživanja moguće je konstruisati intervale poverenja za koeficijent varijacije koji su zasnovani na proceduri *bootstrap* rangirano uzorkovanje po redovima (Ivković & Rajić, 2019, str. 6) koja je detaljno opisana u odeljku 1.4.2. Mogu se razmatrati percentil, BC i BCa intervali poverenja zasnovani na *bootstrap* rangiranom uzorkovanju po redovima (u tabelama koje slede u odeljku 3.3 ovi intervali biće označeni sa RSS percentil interval, RSS BC interval i RSS BCa interval).

3.2.5 Intervali poverenja zasnovani na parcijalnom rangiranom uzorkovanju

Intervali poverenja za koeficijent varijacije mogu se zasnivati i na proceduri parcijalnog rangiranog uzorkovanja (Ivković & Rajić, 2019, str. 6) koja je detaljno opisana u odeljku 1.4.3. Mogu se razmatrati Miller-ov, Curto-ov i Pinto-ov i McKay-ev interval poverenja zasnovani na parcijalnom rangiranom uzorkovanju (u tabelama koje slede u odeljku 3.3 ovi intervali biće označeni sa PRSS Miller-ov interval, PRSS C&P interval i PRSS McKay-ev interval).

3.2.6 Transformacije intervala poverenja

Moguće je vršiti i transformacije postojećih intervala poverenja kako bi se utvrdilo da li su proporcije simuliranih transformisanih intervala bliske nominalnom nivou pouzdanosti. U svrhe istraživanja biće predložene transformacije zasnovane na odsečenoj sredini i *bootstrap* transformacija postojećih intervala poverenja. Prvo, koristiće se odsečena sredina umesto aritmetičke sredine da bi se izračunala varijansa uzorka, jer odsečena sredina predstavlja bolji izbor ukoliko osnovni skup sledi asimetričnu raspodelu (Wilcox, 2005, str. 2066). Drugo, prilikom konstruisanja postojećih intervala poverenja koristiće se percentili *bootstrap* raspodele umesto tradicionalnih kritičnih vrednosti (Ivković & Rajić, 2019, str. 6-7).

Transformacije postojećih intervala zasnovane na odsečenoj sredini

Transformacije postojećih intervala poverenja zasnovane na odsečenoj sredini mogu se vršiti koristeći prosto slučajno uzorkovanje i rangirano uzorkovanje. To podrazumeva da će se varijansa uzorka računati na osnovu odsečene sredine umesto aritmetičke sredine. Ovaj postupak će biti sproveden na dva načina. Prvo, koristiće se Bonett-ova odsečena sredina koja eliminiše procenat vrednosti koji se računa po sledećoj formuli $1/\lceil 2(n-4)^{1/2} \rceil$ (Bonett, 2006, str. 776), tako da Bonett-ova odsečena sredina konvergira aritmetičkoj sredini sa porastom n . Nakon toga, biće korišćena odsečena sredina koja eliminiše 10%³⁵ vrednosti.

Miller-ov, Curto-ov i Pinto-ov i McKay-ev interval poverenja biće transformisani koristeći Bonett-ovu odsečenu sredinu i odsečenu sredinu koja eliminiše 10% vrednosti. Formule za transformisan Miller-ov, Curto-ov i Pinto-ov i McKay-ev interval poverenja koji su zasnovani na Bonett-ovoj odsečenoj sredini ekvivalentne su formulama (3.7), (3.8) i (3.9), gde je \bar{X} Bonett-ova odsečena sredina uzorka (u tabelama koje slede u odeljku 3.3 navedeni intervali biće označeni sa Miller-ov trim interval, C&P trim interval i McKay-ev trim interval).

Transformacije Miller-ovog, Curto-ovog i Pinto-ovog i McKay-evog intervala poverenja zasnovane na odsečenoj sredini koja eliminiše 10% vrednosti biće sprovedene na sličan način (u tabelama koje slede u odeljku 3.3 ovi intervali biće označeni sa Miller-ov trim 10% interval, C&P trim 10% interval i McKay-ev trim 10% interval, respektivno).

Takođe, biće konstruisani Miller-ov, Curto-ov i Pinto-ov i McKay-ev interval poverenja zasnovani na rangiranom uzorkovanju koristeći Bonett-ovu odsečenu sredinu i odsečenu sredinu

³⁵Prilikom istraživanja razmatrane su odsečene sredine kada se eliminiše 5% i 10% vrednosti. Kako nije postojala značajna razlika između dobijenih rezultata, predstavljani su samo rezultati za odsečenu sredinu koja eliminiše 10% vrednosti.

koje eliminiše 10% vrednosti (u tabelama koje slede u odeljku 3.3 navedeni intervali biće označeni sa RSS Miller-ov trim interval, RSS Miller-ov trim 10% interval, RSS C&P trim interval, RSS C&P trim 10% interval, RSS McKay-ev trim interval i RSS McKay-ev trim 10% interval).

Bootstrap transformacije postojećih intervala

Kako bi se proporcije simuliranih intervala približile nominalnom nivou pouzdanosti, moguće je za konstrukciju Miller-ovog i Curto-ovog i Pinto-ovog intervala koristiti percentile koji se koriste prilikom konstrukcije *bootstrap-t* intervala. Na taj način biće konstruisani Miller-ov *bootstrap-t* interval poverenja (označen sa Miller-ov *bootstrap-t* interval) i Curto-ov i Pinto-ov *bootstrap-t* interval poverenja (označen sa C&P *bootstrap-t*).

3.3 Analiza pouzdanosti intervala poverenja za mere disperzije

Prilikom analize pouzdanosti biće ispitane proporcije simuliranih intervala za mere disperzije osnovnog skupa, koji su uvedeni u odeljcima 3.1 i 3.2, u slučajevima kada osnovni skup ne sledi normalnu raspodelu i biće predstavljeni rezultati primene predloženih metoda na simuliranim podacima.

U odeljcima 3.3.1, 3.3.2 i 3.4.3, svi predstavljeni intervali poverenja biće razmatrani koristeći procedure prostog slučajnog uzorkovanja i balansiranog rangiranog uzorkovanja koja je detaljno opisana u odeljku 1.4.1. Intervali poverenja dobijeni procedurom balansiranog rangiranog uzorkovanja biće označeni na sledeći način: RSS χ^2 , RSS Bonett-ov, RSS Steve-ov, RSS LOG i RSS ADF. U cilju izračunavanja ocena sredine i varijanse osnovnog skupa kod rangiranog uzorkovanja korišćene su formule (1.11) i (1.12), koje su navedene u odeljku 1.4.1.

3.3.1 Analiza pouzdanosti intervala poverenja za varijansu u slučaju gama, log-normalne i eksponencijalne raspodele

Pod pretpostavkom da osnovni skup ne sledi normalnu raspodelu, biće analizirane proporcije simuliranih intervala za varijansu osnovnog skupa koristeći prosto slučajno uzorkovanje i rangirano uzorkovanje. Potrebno je pronaći odgovor na pitanje: kada su proporcije simuliranih intervala za varijansu najbliže nominalnom nivou pouzdanosti.

Prvo, u svrhe istraživanja razmatraće se tri slučaja sa: gama, log-normalnom i eksponencijalnom raspodelom. Važno je napomenuti da predmet razmatranja može biti bilo koja asimetrična raspodela.

Nominalni nivo pouzdanosti u svim simulacijama biće 95%. Za svaku kombinaciju parametara, biće izabrano po 1000 uzoraka veličine 15, 20, 50 i 80. Svi razmatrani intervali poverenja su implementirani primenom programskog jezika *R*. Slede rezultati do kojih se došlo.

U prvom slučaju razmatrana je gama³⁶ raspodela sa parametrom $a=2$ i parametrom b koji uzima vrednosti 0,5; 1,6 i 3,2. Proporcije simuliranih 95% dvostranih intervala za varijansu gama

³⁶ Funkcija gustine gama raspodele nalazi se u odeljku 2.1.3.

raspodele predstavljene su u Tabeli 3.1. Kada je reč o malim uzorcima (veličine 15 i 20) najbolji rezultati su dobijeni primenom RSS Bonett-ovog intervala. Za uzorke veličine 50, u zavisnosti od parametra b gama raspodele, najbolje su bile proporcije simuliranih Bonett-ovih i RSS LOG intervala, dok su za uzorke veličine 80 najbolji rezultati dobijeni primenom Bonett-ovog i RSS ADF intervala.

Tabela 3.1: Proporcije simuliranih 95% dvostranih intervala za varijansu gama raspodele

a	2				2				2			
b	0,5				1,6				3,2			
n	15	20	50	80	15	20	50	80	15	20	50	80
χ^2	0,831	0,838	0,784	0,803	0,824	0,854	0,805	0,847	0,829	0,807	0,765	0,758
RSS χ^2	0,865	0,881	0,818	0,861	0,887	0,875	0,898	0,852	0,871	0,868	0,856	0,840
Bonett-ov	0,937	0,909	0,957	0,945	0,911	0,924	0,933	0,958	0,910	0,923	0,936	0,952
RSS Bonett-ov	0,939	0,960	0,966	0,958	0,936	0,961	0,972	0,976	0,932	0,952	0,982	0,965
Steve-ov	0,714	0,769	0,902	0,913	0,750	0,726	0,900	0,918	0,723	0,750	0,901	0,907
RSS Steve-ov	0,748	0,773	0,932	0,936	0,757	0,747	0,923	0,939	0,737	0,832	0,923	0,941
LOG	0,797	0,788	0,903	0,923	0,767	0,795	0,885	0,915	0,797	0,840	0,885	0,905
RSS LOG	0,821	0,864	0,923	0,943	0,784	0,830	0,939	0,932	0,809	0,845	0,939	0,938
ADF	0,841	0,857	0,907	0,926	0,835	0,829	0,908	0,921	0,883	0,853	0,919	0,922
RSS ADF	0,895	0,907	0,935	0,952	0,883	0,911	0,934	0,943	0,893	0,905	0,932	0,936

Izvor: Preuzeto iz Ivković i Rajić (2018)

Predmet analize u drugom slučaju je log-normalna³⁷ raspodela sa parametrom $\mu = 2$ i parametrom σ koji uzima vrednosti 0,25; 0,5 i 0,6. U Tabeli 3.2 prikazane su proporcije simuliranih 95% dvostranih intervala za varijansu log-normalne raspodele. Za male uzorke, u zavisnosti od parametra σ log-normalne raspodele, proporcije simuliranih RSS χ^2 i RSS Bonett-ovih intervala bile su najbliže nominalnom nivou pouzdanosti. Analizom velikih uzoraka utvrđeno je da su za uzorke veličine 50 najbolji rezultati dobijeni korišćenjem RSS Bonett-ovog i Steve-ovog intervala, dok su se za uzorke veličine 80 najbolje pokazali Bonett-ov, RSS Bonett-ov, Steve-ov i RSS ADF interval.

³⁷ Funkcija gustine log-normalne raspodele nalazi se u odeljku 2.1.4.

Tabela 3.2: Proporcije simuliranih 95% dvostranih intervala za varijansu log-normalne raspodele

μ	2				2				2			
σ	0,25				0,5				0,6			
n	15	20	50	80	15	20	50	80	15	20	50	80
χ^2	0,911	0,920	0,909	0,912	0,817	0,758	0,669	0,762	0,734	0,709	0,655	0,556
RSS χ^2	0,946	0,937	0,918	0,938	0,846	0,850	0,828	0,766	0,753	0,737	0,671	0,682
Bonett-ov	0,965	0,964	0,972	0,969	0,922	0,930	0,937	0,949	0,915	0,900	0,919	0,939
RSS Bonett-ov	0,977	0,980	0,974	0,980	0,942	0,949	0,953	0,961	0,924	0,933	0,940	0,941
Steve-ov	0,788	0,833	0,947	0,945	0,658	0,679	0,911	0,920	0,610	0,579	0,842	0,866
RSS Steve-ov	0,839	0,883	0,954	0,964	0,690	0,729	0,927	0,936	0,636	0,612	0,884	0,888
LOG	0,807	0,864	0,923	0,929	0,747	0,768	0,872	0,905	0,676	0,749	0,791	0,862
RSS LOG	0,852	0,874	0,936	0,943	0,759	0,818	0,910	0,932	0,753	0,760	0,884	0,886
ADF	0,904	0,915	0,927	0,942	0,769	0,800	0,904	0,920	0,777	0,785	0,867	0,884
RSS ADF	0,924	0,935	0,963	0,955	0,838	0,850	0,914	0,925	0,802	0,840	0,875	0,897

Izvor: Preuzeto iz Ivković i Rajić (2018)

U trećem slučaju analizirana je eksponencijalna³⁸ raspodela sa parametrom λ koji uzima vrednosti 0,5; 1,5 i 2,2. Proporcije simuliranih 95% dvostranih intervala za varijansu eksponencijalne raspodele prikazane su u Tabeli 3.3. Kada je reč o malim uzorcima i o uzorcima veličine 50 najbolji rezultati su dobijeni pomoću RSS Bonett-ovog intervala. Za uzorke veličine 80, u zavisnosti od parametra λ eksponencijalne raspodele, proporcije simuliranih Bonett-ovih i RSS Bonett-ovih intervala su bile najbolje.

Tabela 3.3: Proporcije simuliranih 95% dvostranih intervala za varijansu eksponencijalne raspodele

λ	0,5				1,5				2,2			
n	15	20	50	80	15	20	50	80	15	20	50	80
χ^2	0,757	0,753	0,720	0,722	0,753	0,769	0,683	0,658	0,746	0,694	0,687	0,653
RSS χ^2	0,799	0,843	0,762	0,751	0,763	0,786	0,751	0,744	0,755	0,785	0,734	0,739
Bonett-ov	0,904	0,911	0,938	0,943	0,905	0,931	0,918	0,942	0,904	0,914	0,931	0,933
RSS Bonett-ov	0,930	0,943	0,958	0,960	0,935	0,953	0,960	0,963	0,923	0,938	0,945	0,966
Steve-ov	0,614	0,586	0,813	0,908	0,510	0,631	0,822	0,916	0,625	0,647	0,883	0,908
RSS Steve-ov	0,637	0,722	0,907	0,930	0,654	0,763	0,901	0,934	0,652	0,721	0,897	0,922
LOG	0,763	0,753	0,848	0,903	0,746	0,770	0,888	0,914	0,720	0,704	0,875	0,894
RSS LOG	0,769	0,832	0,891	0,917	0,795	0,836	0,902	0,920	0,726	0,808	0,908	0,917
ADF	0,781	0,844	0,904	0,907	0,764	0,797	0,900	0,918	0,780	0,778	0,893	0,906
RSS ADF	0,797	0,858	0,910	0,925	0,807	0,836	0,925	0,933	0,839	0,840	0,906	0,922

Izvor: Preuzeto iz Ivković i Rajić (2018)

Na osnovu navedenih rezultata zaključeno je da su za gama, log-normalnu i eksponencijalnu raspodelu, u većini slučajeva, za sve veličine uzoraka, najbolje proporcije simuliranih intervala za varijansu dobijene primenom rangiranog uzorkovanja.

³⁸ Funkcija gustine eksponencijalne raspodele nalazi se u odeljku 2.1.2.

3.3.2 Analiza pouzdanosti intervala poverenja za varijansu u slučaju *skew*-normalne, *skew*-Student-ove, *skew*-Laplace-ove i *skew* generalizovane *t* raspodele

U ovom delu razmatraće se sledeća četiri slučaja sa: *skew*-normalnom raspodelom, *skew*-Student-ovom raspodelom, *skew*-Laplace-ovom raspodelom i *skew* generalizovanom *t* raspodelom. Važno je napomenuti da može biti korišćena bilo koja druga asimetrična raspodela.

Nominalni nivo pouzdanosti u svim simulacijama biće 95%. Za svaku kombinaciju parametara biće generisano po 1000 uzoraka veličine 15, 20, 50 i 80. Svi predstavljeni intervali poverenja su implementirani koristeći programski jezik *R*.

U prvom slučaju razmatrana je *skew*-normalna³⁹ raspodela sa parametrima $\beta = 3$, $\gamma = 2$ i parametrom α koji uzima vrednosti 1; 2,5 i 5. U Tabeli 3.4 predstavljene su proporcije simuliranih 95% dvostranih intervala za varijansu *skew*-normalne raspodele. Za uzorke veličine 15 najbolji rezultati dobijeni su primenom χ^2 i RSS χ^2 intervala, dok su se za uzorke veličine 20 kao najbolji intervali pokazali Bonett-ov, RSS χ^2 i RSS ADF interval. Kada je reč o uzorcima veličine 50, u zavisnosti od parametra α *skew*-normalne raspodele, proporcije simuliranih RSS χ^2 , RSS Steve-ovih i RSS ADF intervala su bile najbolje, dok su za uzorke veličine 80 najbolji rezultati dobijeni primenom χ^2 , RSS LOG i RSS ADF intervala.

Tabela 3.4: Proporcije simuliranih 95% dvostranih intervala za varijansu *skew*-normalne raspodele

β	3				3				3			
γ	2				2				2			
α	1				2,5				5			
n	15	20	50	80	15	20	50	80	15	20	50	80
χ^2	0,936	0,968	0,912	0,949	0,939	0,918	0,895	0,942	0,935	0,904	0,898	0,868
RSS χ^2	0,964	0,971	0,963	0,959	0,946	0,946	0,945	0,944	0,926	0,953	0,935	0,897
Bonett-ov	0,977	0,973	0,980	0,986	0,978	0,963	0,980	0,965	0,970	0,950	0,961	0,969
RSS Bonett-ov	0,977	0,981	0,987	0,992	0,968	0,985	0,993	0,983	0,978	0,987	0,987	0,982
Steve-ov	0,845	0,896	0,901	0,937	0,840	0,870	0,905	0,927	0,812	0,875	0,899	0,935
RSS Steve-ov	0,892	0,892	0,958	0,952	0,875	0,892	0,941	0,941	0,852	0,889	0,926	0,936
LOG	0,882	0,894	0,900	0,926	0,858	0,858	0,914	0,930	0,831	0,867	0,904	0,933
RSS LOG	0,883	0,918	0,939	0,944	0,845	0,899	0,943	0,951	0,848	0,881	0,947	0,947
ADF	0,921	0,921	0,928	0,905	0,896	0,911	0,933	0,934	0,860	0,902	0,905	0,920
RSS ADF	0,917	0,937	0,964	0,949	0,916	0,928	0,963	0,947	0,910	0,934	0,949	0,955

Izvor: Preuzeto iz Rajić i Ivković (2018)

Predmet analize u drugom slučaju bila je *skew*-Student-ova⁴⁰ raspodela sa parametrima $\beta = 3$, $\gamma = 2$, parametrom α koji uzima vrednosti 1; 2,5 i 5 i brojem stepeni slobode $\nu = 10$. U Tabeli 3.5 predstavljene su proporcije simuliranih 95% dvostranih intervala za varijansu *skew*-Student-ove raspodele. Za male uzorke, u zavisnosti od parametra α *skew*-Student-ove raspodele,

³⁹ Funkcija gustine *skew*-normalne raspodele nalazi se u odeljku 2.1.5.

⁴⁰ Funkcija gustine *skew*-Student-ove raspodele nalazi se u odeljku 2.1.6.

najbolji rezultati su dobijeni primenom Bonett-ovog i RSS Bonett-ovog intervala. Kada je reč o uzorcima veličine 50 proporcije simuliranih Bonett-ovih i RSS ADF intervala su bile najbliže 0,95, dok su za uzorke veličine 80 najbolji rezultati dobijeni primenom Bonett-ovog i RSS Steve-ovog intervala.

Tabela 3.5: Proporcije simuliranih 95% dvostranih intervala za varijansu *skew*-Student-ove raspodele

β	3				3				3			
γ	2				2				2			
α	1				2,5				5			
ν	10				10				10			
n	15	20	50	80	15	20	50	80	15	20	50	80
χ^2	0,882	0,895	0,879	0,838	0,876	0,887	0,835	0,840	0,885	0,828	0,878	0,831
RSS χ^2	0,897	0,902	0,916	0,937	0,901	0,901	0,861	0,832	0,831	0,917	0,856	0,785
Bonett-ov	0,947	0,965	0,967	0,964	0,953	0,938	0,961	0,930	0,909	0,934	0,957	0,946
RSS Bonett-ov	0,953	0,965	0,980	0,975	0,970	0,962	0,969	0,980	0,939	0,959	0,974	0,962
Steve-ov	0,830	0,871	0,886	0,937	0,778	0,807	0,874	0,921	0,760	0,791	0,880	0,900
RSS Steve-ov	0,822	0,886	0,916	0,949	0,812	0,849	0,881	0,920	0,724	0,827	0,903	0,912
LOG	0,854	0,854	0,924	0,927	0,810	0,839	0,880	0,908	0,743	0,799	0,857	0,906
RSS LOG	0,832	0,908	0,942	0,934	0,796	0,864	0,933	0,919	0,837	0,840	0,908	0,918
ADF	0,901	0,897	0,906	0,947	0,890	0,858	0,889	0,898	0,838	0,848	0,906	0,925
RSS ADF	0,902	0,930	0,946	0,931	0,887	0,898	0,915	0,928	0,870	0,901	0,941	0,913

Izvor: Preuzeto iz Rajić i Ivković (2018)

U trećem slučaju ispitivana je *skew*-Laplace-ova⁴¹ raspodela sa parametrima $\mu = 3$, $\beta = 1$ i parametrom α koji uzima vrednosti 1, 2 i 4. Proporcije simuliranih 95% dvostranih intervala za varijansu *skew*-Laplace-ove raspodele predstavljene su u Tabeli 3.6. Analiza uzoraka veličine 15 pokazala je da su najbolji rezultati ostvareni pomoću Bonett-ovog i RSS Bonett-ovog intervala, dok je RSS Bonett-ov interval bio najbolji za uzorke veličine 20. Za uzorke veličine 50, u zavisnosti od parametra α *skew*-Laplace-ove raspodele, proporcije simuliranih Bonett-ovih i RSS Bonett-ovih intervala su bile najbolje. Bonett-ov i RSS ADF interval su bili najbolji izbor za uzorke veličine 80.

⁴¹ Funkcija gustine *skew*-Laplace-ove raspodele nalazi se u odeljku 2.1.7.

Tabela 3.6: Proporcije simuliranih 95% dvostranih intervala za varijansu *skew*-Laplace-ove raspodele

μ	3				3				3			
α	1				2				4			
β	1				1				1			
n	15	20	50	80	15	20	50	80	15	20	50	80
χ^2	0,854	0,799	0,838	0,788	0,789	0,752	0,729	0,790	0,791	0,762	0,713	0,735
RSS χ^2	0,848	0,822	0,886	0,722	0,847	0,819	0,815	0,715	0,768	0,792	0,718	0,764
Bonett-ov	0,937	0,933	0,954	0,950	0,934	0,921	0,926	0,939	0,890	0,900	0,920	0,938
RSS Bonett-ov	0,938	0,961	0,966	0,967	0,928	0,938	0,959	0,960	0,910	0,937	0,929	0,905
Steve-ov	0,732	0,827	0,910	0,939	0,718	0,786	0,899	0,900	0,611	0,664	0,830	0,828
RSS Steve-ov	0,744	0,801	0,918	0,940	0,687	0,781	0,933	0,936	0,704	0,732	0,827	0,869
LOG	0,801	0,824	0,881	0,915	0,747	0,841	0,871	0,866	0,762	0,739	0,864	0,858
RSS LOG	0,834	0,862	0,919	0,933	0,787	0,840	0,924	0,911	0,736	0,797	0,868	0,864
ADF	0,882	0,914	0,901	0,922	0,861	0,859	0,895	0,904	0,852	0,829	0,853	0,859
RSS ADF	0,897	0,890	0,941	0,931	0,869	0,897	0,915	0,941	0,823	0,846	0,886	0,892

Izvor: Preuzeto iz Rajić i Ivković (2018)

U četvrtom slučaju proučavana je *skew* generalizovana t^{42} raspodela sa parametrima $\mu=3$, $\sigma=1,5$, $p=5$, $q=Inf$ i parametrom λ koji uzima vrednosti 0,2; 0,5 i 0,8. U Tabeli 3.7 predstavljene su proporcije simuliranih 95% dvostranih intervala za varijansu *skew* generalizovane t raspodele. Za uzorke veličine 15, u zavisnosti od parametra λ *skew* generalizovane t raspodele, proporcije simuliranih ADF i RSS ADF intervala bile su najbliže 0,95, dok su RSS Steve-ov, RSS LOG, ADF i RSS ADF intervali bili najbolji izbor za uzorke veličine 20. Utvrđeno je da su za uzorke veličine 50 najbolji rezultati ostvareni primenom Steve-ovog i ADF intervala, a za uzorke veličine 80 kao najbolji su se pokazali RSS LOG, ADF i RSS ADF intervali.

⁴² Funkcija gustine *skew* generalizovane t raspodele nalazi se u odeljku 2.1.6.

Tabela 3.7: Proporcije simuliranih 95% dvostranih intervala za varijansu *skew* generalizovane *t* raspodele

μ	3				3				3			
σ	1,5				1,5				1,5			
λ	0,2				0,5				0,8			
p	5				5				5			
q	Inf				Inf				Inf			
n	15	20	50	80	15	20	50	80	15	20	50	80
χ^2	0,990	0,994	0,993	0,990	0,991	0,992	0,994	0,994	0,988	0,983	0,986	0,991
RSS χ^2	0,994	0,996	0,999	0,996	0,992	1,000	0,996	0,997	0,995	0,996	0,997	0,991
Bonett-ov	0,985	0,991	0,992	0,995	0,985	0,992	0,990	0,987	0,981	0,984	0,998	0,991
RSS Bonett-ov	0,990	0,995	1,000	0,998	0,984	0,995	0,996	0,999	0,988	0,997	0,999	0,998
Steve-ov	0,890	0,897	0,949	0,964	0,911	0,924	0,944	0,938	0,892	0,918	0,944	0,916
RSS Steve-ov	0,918	0,919	0,970	0,973	0,923	0,942	0,979	0,964	0,883	0,936	0,978	0,970
LOG	0,906	0,911	0,940	0,937	0,898	0,927	0,931	0,943	0,889	0,914	0,934	0,940
RSS LOG	0,927	0,954	0,970	0,960	0,908	0,940	0,965	0,951	0,918	0,941	0,954	0,957
ADF	0,940	0,946	0,965	0,945	0,940	0,939	0,946	0,961	0,924	0,940	0,953	0,926
RSS ADF	0,945	0,958	0,954	0,970	0,933	0,967	0,966	0,970	0,931	0,950	0,974	0,943

Izvor: Preuzeto iz Rajić i Ivković (2018)

Posmatrajući predstavljene rezultate, može se izvesti sledeći zaključak koji se odnosi na *skew*-normalnu, *skew*-Student-ovu, *skew*-Laplace-ovu i *skew* generalizovanu *t* raspodelu. U većini slučajeva, za sve veličine uzoraka, proporcije simuliranih intervala za varijansu zasnovanih na rangiranom uzorkovanju su najbliže nominalnom nivou pouzdanosti. Ovo se poklapa sa zaključkom do koga je došao Wolfe (2004) koji je utvrdio da su ocene dobijene putem rangiranog uzorkovanja efikasnije.

3.3.3 Analiza pouzdanosti intervala poverenja za koeficijent varijacije

Mnogi istraživači su razmatrali različite intervale poverenja za koeficijent varijacije osnovnog skupa. Međutim studije nisu u potpunosti uporedive, jer su sprovedene pod različitim uslovima simulacije. Cilj ovog istraživanja je da se uporede proporcije svih razmatranih simuliranih intervala pod istim uslovima simulacije ako osnovni skup nije normalno raspodeljen.

Nominalni nivo pouzdanosti u svim simulacijama biće 95%. U svrhe istraživanja biće razmatrana četiri slučaja sa: gama raspodelom, Weibull-ovom raspodelom, log-normalnom raspodelom i *skew*-normalnom raspodelom. Važno je napomenuti da su ove raspodele proizvoljno izabrane, a da predmet analize može biti bilo koja druga asimetrična raspodela.

Procedura koja će se sprovesti je ista za svaki slučaj i podrazumeva da će za svaku razmatranu raspodelu biti analizirane iste vrednosti koeficijenta varijacije osnovnog skupa: 0,1; 0,2 i 0,4. Za svaku kombinaciju parametara, biće generisano po 1000 uzoraka veličine 10, 20, 40 i 80.

Za potrebe formiranja intervala poverenja zasnovanih na metodama ponovljenih uzoraka, biće generisano 1000 *bootstrap* uzoraka iz svakog uzorka. U cilju konstruisanja intervala poverenja zasnovanih na rangiranom uzorkovanju, biće definisani parametri k i m na sledeći način: 1) $k=2$ i m uzima vrednosti 5, 10, 20 i 40; 2) $k=5$ i m uzima vrednosti 2, 4, 8 i 16. Analiza intervala poverenja zasnovanih na parcijalnom rangiranom uzorkovanju zahtevaće da se pored parametra $k=5$ i parametra m koji uzima vrednosti 2, 4, 8 i 16 uvede i parametar r koji uzima vrednosti 1 i 2. Za sve razmatrane intervale poverenja napisani su originalni kodovi u programskom jeziku R .

U prvom slučaju analizirana je gama⁴³ raspodela. Kako bi se dobile vrednosti koeficijenta varijacije osnovnog skupa: 0,1; 0,2 i 0,4, pretpostavljaće se da je parametar $b=2$, a zatim se parametar a prilagođava uzimajući vrednosti 100, 25 i 6,25. U Tabeli 3.8 predstavljene su proporcije simuliranih 95% dvostranih intervala za koeficijent varijacije gama raspodele. Za uzorke veličine 10, u zavisnosti od parametra a gama raspodele, najbolji rezultati su dobijeni primenom RSS C&P i McKay-evog intervala koji su zasnovani na odsečenoj sredini i C&P *bootstrap-t* intervala. Kada je reč o uzorcima veličine 20 najbolji izbor su bili RSS Miller-ov i McKay-ev interval koji su zasnovani na odsečenoj sredini. Kod uzoraka veličine 40, proporcije simuliranih RSS Miller-ovih, C&P, RSS C&P i McKay-evih intervala koji su zasnovani na odsečenoj sredini su bile najbliže 0,95, dok su za uzorke veličine 80 najbolji izbor bili RSS Miller-ov i McKay-ev interval koji su zasnovani na odsečenoj sredini.

⁴³ Funkcija gustine gama raspodele nalazi se u odeljku 2.1.3.

Tabela 3.8: Proporcije simuliranih 95% dvostranih intervala za koeficijent varijacije gama raspodele

a	100				25				6,25				
b	2				2				2				
n	10	20	40	80	10	20	40	80	10	20	40	80	
t	0,936	0,929	0,910	0,906	0,930	0,909	0,917	0,925	0,897	0,907	0,930	0,914	
Percentil	0,768	0,815	0,835	0,877	0,753	0,838	0,859	0,899	0,782	0,876	0,851	0,890	
RSS	$k=2$	0,789	0,873	0,903	0,943	0,801	0,881	0,907	0,915	0,787	0,870	0,867	0,922
Percentil	$k=5$	0,695	0,849	0,914	0,923	0,737	0,818	0,881	0,911	0,686	0,843	0,858	0,897
Bootstrap-t		0,861	0,914	0,930	0,937	0,862	0,919	0,931	0,940	0,851	0,915	0,945	0,934
BC		0,813	0,883	0,899	0,929	0,818	0,896	0,908	0,946	0,849	0,875	0,910	0,948
RSS BC	$k=2$	0,861	0,888	0,905	0,915	0,802	0,855	0,881	0,939	0,832	0,868	0,920	0,926
	$k=5$	0,691	0,867	0,917	0,922	0,703	0,872	0,901	0,917	0,674	0,851	0,893	0,894
Bca		0,840	0,896	0,926	0,935	0,846	0,877	0,924	0,947	0,866	0,899	0,922	0,934
RSS Bca	$k=2$	0,825	0,890	0,939	0,904	0,826	0,890	0,900	0,922	0,839	0,903	0,897	0,914
	$k=5$	0,698	0,869	0,900	0,927	0,704	0,875	0,902	0,921	0,727	0,852	0,892	0,919
Jackknife		0,927	0,940	0,939	0,926	0,926	0,929	0,954	0,945	0,925	0,918	0,934	0,953
Miller-ov		0,883	0,906	0,956	0,954	0,906	0,908	0,957	0,971	0,907	0,915	0,942	0,959
Miller-ov trim		0,929	0,937	0,953	0,951	0,929	0,942	0,941	0,943	0,935	0,948	0,953	0,949
Miller-ov trim 10%		0,913	0,928	0,937	0,934	0,917	0,930	0,939	0,944	0,915	0,944	0,958	0,949
RSS Miller-ov	$k=2$	0,910	0,901	0,918	0,936	0,895	0,935	0,932	0,943	0,925	0,957	0,963	0,955
	$k=5$	0,948	0,961	0,958	0,976	0,944	0,957	0,959	0,963	0,962	0,989	0,988	0,977
RSS Miller-ov trim	$k=2$	0,905	0,929	0,949	0,937	0,915	0,937	0,951	0,953	0,919	0,933	0,939	0,970
	$k=5$	0,947	0,949	0,967	0,963	0,960	0,957	0,958	0,946	0,961	0,971	0,970	0,975
RSS Miller-ov trim 10%	$k=2$	0,910	0,937	0,923	0,949	0,920	0,915	0,958	0,950	0,925	0,950	0,926	0,963
	$k=5$	0,951	0,973	0,962	0,956	0,946	0,973	0,977	0,970	0,964	0,961	0,970	0,957
Miller-ov bootstrap-t		0,893	0,896	0,941	0,946	0,885	0,913	0,948	0,947	0,921	0,956	0,970	0,965
PRSS Miller-ov	$r=1$	0,709	0,704	0,428	0,347	0,735	0,673	0,575	0,374	0,742	0,627	0,593	0,316
	$r=2$	0,855	0,857	0,823	0,712	0,835	0,813	0,832	0,800	0,857	0,862	0,868	0,826

Izvor: Preuzeto iz Ivković i Rajić (2019)

Tabela 3.8: Proporcije 95% dvostranih simuliranih intervala za koeficijent varijacije gama raspodele (nastavak)

a	100				25				6,25				
b	2				2				2				
n	10	20	40	80	10	20	40	80	10	20	40	80	
C&P	0,907	0,937	0,906	0,952	0,912	0,903	0,924	0,957	0,889	0,956	0,919	0,973	
C&P trim	0,904	0,952	0,943	0,952	0,923	0,954	0,956	0,951	0,922	0,945	0,951	0,954	
C&P trim 10%	0,902	0,931	0,932	0,943	0,868	0,926	0,938	0,962	0,904	0,923	0,951	0,962	
RSS C&P	$k=2$	0,883	0,923	0,927	0,944	0,881	0,928	0,934	0,963	0,891	0,922	0,944	0,961
	$k=5$	0,947	0,942	0,976	0,967	0,969	0,956	0,974	0,932	0,944	0,960	0,979	0,982
RSS C&P trim	$k=2$	0,911	0,934	0,924	0,962	0,920	0,923	0,952	0,954	0,936	0,946	0,923	0,918
	$k=5$	0,966	0,959	0,951	0,978	0,951	0,959	0,957	0,971	0,963	0,966	0,969	0,967
RSS C&P trim 10%	$k=2$	0,906	0,937	0,927	0,909	0,933	0,941	0,922	0,955	0,930	0,918	0,960	0,951
	$k=5$	0,943	0,966	0,949	0,967	0,971	0,965	0,965	0,954	0,939	0,965	0,985	0,982
C&P bootstrap-t	0,891	0,888	0,936	0,945	0,898	0,904	0,934	0,949	0,947	0,949	0,955	0,967	
PRSS C&P	$r=1$	0,736	0,650	0,502	0,214	0,689	0,650	0,523	0,221	0,732	0,684	0,573	0,357
	$r=2$	0,835	0,803	0,792	0,718	0,809	0,859	0,825	0,745	0,827	0,843	0,856	0,733
McKay-ev	0,948	0,956	0,941	0,965	0,957	0,963	0,964	0,935	0,973	0,978	0,963	0,975	
McKay-ev trim	0,954	0,951	0,949	0,950	0,952	0,946	0,942	0,950	0,944	0,978	0,972	0,950	
McKay-ev trim 10%	0,950	0,949	0,930	0,961	0,951	0,953	0,967	0,949	0,966	0,967	0,958	0,950	
RSS McKay-ev	$k=2$	0,944	0,966	0,965	0,958	0,944	0,958	0,927	0,942	0,983	0,972	0,970	0,961
	$k=5$	0,983	0,975	0,968	0,982	0,979	0,974	0,973	0,982	0,993	0,981	0,979	0,971
RSS McKay-ev trim	$k=2$	0,947	0,947	0,959	0,947	0,945	0,962	0,934	0,951	0,973	0,954	0,958	0,964
	$k=5$	0,974	0,977	0,969	0,963	0,972	0,968	0,980	0,979	0,986	0,987	0,978	0,970
RSS McKay-ev trim 10%	$k=2$	0,952	0,945	0,954	0,967	0,958	0,957	0,959	0,949	0,967	0,974	0,965	0,963
	$k=5$	0,967	0,979	0,972	0,971	0,962	0,970	0,984	0,972	0,981	0,981	0,974	0,986
PRSS McKay-ev	$r=1$	0,902	0,794	0,642	0,343	0,930	0,855	0,723	0,390	0,923	0,891	0,696	0,382
McKay-ev	$r=2$	0,920	0,927	0,909	0,888	0,944	0,917	0,909	0,906	0,958	0,913	0,943	0,854

U drugom slučaju predmet proučavanja je bila Weibull-ova⁴⁴ raspodela sa parametrom a koji uzima vrednosti 12,1543; 5,80 i 2,70, i parametrom $b = 2$. U cilju izračunavanja odgovarajućih vrednosti koeficijenta varijacije osnovnog skupa izvršena je procedura opisana u prvom slučaju. Proporcije simuliranih 95% dvostranih intervala za koeficijent varijacije Weibull-ove raspodele prikazane su u Tabeli 3.9. Za uzorke veličine 10, u zavisnosti od parametra a Weibull-ove raspodele, RSS Miller-ov, Miller-ov *bootstrap-t*, RSS McKay-ev i PRSS McKay-ev interval obezbedili su najbolje rezultate, dok su se za uzorke veličine 20 kao najbolji pokazali RSS Miller-ov, C&P, RSS C&P i RSS McKay-ev interval koji su zasnovani na odsečenoj sredini. Kada je reč o uzorcima veličine 40, najbolji rezultati su dobijeni pomoću Miller-ovog, RSS Miller-ovog, McKay-evog i RSS McKay-evog intervala koji su zasnovani na odsečenoj sredini. Za uzorke veličine 80 najbolji izbor su bili Miller-ov, C&P, RSS C&P i RSS McKay-ev interval koji su zasnovani na odsečenoj sredini.

Tabela 3.9: Proporcije 95% dvostranih simuliranih intervala za koeficijent varijacije Weibull-ove raspodele

a		12,1543				5,80				2,70			
b		2				2				2			
n		10	20	40	80	10	20	40	80	10	20	40	80
t		0,857	0,878	0,903	0,871	0,888	0,902	0,916	0,909	0,920	0,929	0,912	0,928
Percentil		0,747	0,844	0,861	0,853	0,775	0,876	0,917	0,860	0,826	0,903	0,886	0,932
RSS Percentil	$k=2$	0,738	0,834	0,892	0,925	0,825	0,853	0,907	0,919	0,810	0,895	0,925	0,937
	$k=5$	0,654	0,790	0,872	0,904	0,702	0,858	0,889	0,917	0,742	0,825	0,894	0,921
<i>Bootstrap-t</i>		0,877	0,897	0,895	0,913	0,876	0,916	0,923	0,940	0,900	0,933	0,938	0,915
BC		0,774	0,859	0,889	0,929	0,817	0,870	0,923	0,941	0,808	0,892	0,926	0,934
RSS BC	$k=2$	0,803	0,858	0,881	0,905	0,808	0,872	0,913	0,920	0,819	0,902	0,934	0,938
	$k=5$	0,688	0,803	0,863	0,915	0,695	0,829	0,888	0,936	0,740	0,888	0,900	0,919
Bca		0,786	0,848	0,908	0,914	0,825	0,856	0,932	0,934	0,869	0,922	0,925	0,936
RSS Bca	$k=2$	0,799	0,863	0,889	0,919	0,842	0,869	0,911	0,912	0,849	0,905	0,940	0,931
	$k=5$	0,657	0,806	0,883	0,919	0,676	0,818	0,887	0,907	0,727	0,856	0,891	0,892
<i>Jackknife</i>		0,883	0,896	0,886	0,932	0,899	0,915	0,928	0,937	0,932	0,917	0,940	0,930
Miller-ov		0,908	0,855	0,902	0,891	0,897	0,926	0,957	0,933	0,918	0,930	0,960	0,965
Miller-ov trim		0,864	0,883	0,908	0,903	0,905	0,916	0,949	0,954	0,935	0,944	0,953	0,968
Miller-ov trim 10%		0,858	0,849	0,885	0,843	0,879	0,907	0,943	0,949	0,907	0,938	0,951	0,970
RSS Miller-ov	$k=2$	0,868	0,878	0,877	0,871	0,907	0,927	0,911	0,947	0,927	0,956	0,968	0,934
	$k=5$	0,925	0,929	0,914	0,935	0,949	0,956	0,970	0,957	0,963	0,962	0,980	0,962
RSS Miller-ov trim	$k=2$	0,876	0,902	0,903	0,926	0,913	0,931	0,929	0,940	0,927	0,938	0,977	0,967
	$k=5$	0,928	0,904	0,940	0,914	0,958	0,950	0,973	0,971	0,972	0,970	0,981	0,989
RSS Miller-ov trim 10%	$k=2$	0,861	0,913	0,917	0,888	0,895	0,905	0,949	0,946	0,926	0,959	0,941	0,964
	$k=5$	0,917	0,937	0,954	0,938	0,939	0,962	0,968	0,964	0,966	0,968	0,984	0,982
Miller-ov <i>bootstrap-t</i>		0,867	0,888	0,913	0,927	0,924	0,927	0,945	0,944	0,951	0,958	0,962	0,974
PRSS Miller-ov	$r=1$	0,684	0,563	0,329	0,211	0,720	0,692	0,468	0,315	0,791	0,735	0,554	0,476
	$r=2$	0,774	0,771	0,714	0,615	0,787	0,805	0,814	0,743	0,880	0,917	0,879	0,833

Izvor: Preuzeto iz Ivković i Rajić (2019)

⁴⁴ Funkcija gustine Weibull-ove raspodele nalazi se u odeljku 2.1.1.

Tabela 3.9: Proporcije 95% dvostranih simuliranih intervala za koeficijent varijacije Weibull-ove raspodele (nastavak)

<i>a</i>	12,1543				5,80				2,70				
<i>b</i>	2				2				2				
<i>n</i>	10	20	40	80	10	20	40	80	10	20	40	80	
C&P	0,818	0,882	0,919	0,879	0,883	0,916	0,935	0,927	0,922	0,925	0,946	0,970	
C&P trim	0,829	0,857	0,881	0,860	0,888	0,935	0,948	0,948	0,885	0,950	0,946	0,946	
C&P trim 10%	0,866	0,876	0,834	0,890	0,876	0,901	0,929	0,920	0,922	0,952	0,959	0,971	
RSS C&P	<i>k=2</i>	0,842	0,888	0,851	0,895	0,880	0,895	0,930	0,942	0,936	0,957	0,956	0,969
	<i>k=5</i>	0,927	0,923	0,919	0,943	0,944	0,943	0,966	0,986	0,959	0,974	0,986	0,995
RSS C&P trim	<i>k=2</i>	0,867	0,887	0,901	0,893	0,847	0,891	0,944	0,942	0,915	0,941	0,975	0,957
	<i>k=5</i>	0,875	0,903	0,941	0,946	0,933	0,968	0,941	0,965	0,977	0,985	0,984	0,985
RSS C&P trim 10%	<i>k=2</i>	0,861	0,906	0,923	0,872	0,880	0,891	0,930	0,945	0,935	0,949	0,957	0,946
	<i>k=5</i>	0,926	0,939	0,935	0,941	0,921	0,955	0,974	0,962	0,955	0,980	0,986	0,988
C&P bootstrap-t	0,857	0,873	0,924	0,924	0,890	0,909	0,931	0,944	0,942	0,954	0,967	0,979	
PRSS C&P	<i>r=1</i>	0,625	0,524	0,422	0,230	0,690	0,621	0,367	0,546	0,764	0,631	0,527	0,426
	<i>r=2</i>	0,730	0,760	0,679	0,687	0,745	0,793	0,793	0,741	0,837	0,806	0,795	0,845
McKay-ev	0,881	0,903	0,878	0,911	0,937	0,936	0,935	0,944	0,975	0,962	0,980	0,970	
McKay-ev trim	0,876	0,882	0,876	0,883	0,943	0,943	0,949	0,952	0,968	0,965	0,973	0,973	
McKay-ev trim 10%	0,901	0,909	0,900	0,903	0,921	0,923	0,929	0,925	0,979	0,972	0,983	0,979	
RSS McKay-ev	<i>k=2</i>	0,940	0,901	0,904	0,933	0,943	0,934	0,934	0,953	0,984	0,980	0,988	0,970
	<i>k=5</i>	0,939	0,927	0,958	0,941	0,973	0,975	0,944	0,974	0,995	0,997	0,996	0,996
RSS McKay-ev trim	<i>k=2</i>	0,886	0,893	0,901	0,871	0,926	0,945	0,937	0,935	0,973	0,985	0,982	0,980
	<i>k=5</i>	0,905	0,935	0,950	0,929	0,978	0,966	0,970	0,966	0,992	0,997	0,997	0,993
RSS McKay-ev trim 10%	<i>k=2</i>	0,903	0,916	0,924	0,886	0,923	0,944	0,943	0,949	0,983	0,993	0,978	0,982
	<i>k=5</i>	0,925	0,939	0,926	0,932	0,979	0,987	0,970	0,967	0,993	0,991	0,986	0,996
PRSS McKay-ev	<i>r=1</i>	0,878	0,711	0,550	0,387	0,910	0,816	0,630	0,347	0,951	0,918	0,833	0,566
	<i>r=2</i>	0,906	0,881	0,845	0,731	0,944	0,932	0,866	0,830	0,961	0,952	0,953	0,957

U trećem slučaju analizirana je log-normalna⁴⁵ raspodela sa parametrom $\mu=2$ i parametrom σ koji uzima vrednosti 0,0998; 0,1980 i 0,3853. Kako bi se dobile odgovarajuće vrednosti koeficijenta varijacije osnovnog skupa sprovedena je opisana procedura. Proporcije simuliranih 95% dvostranih intervala za koeficijent varijacije log-normalne raspodele date su u Tabeli 3.10. Kada je reč o uzorcima veličine 10, proporcije simuliranih RSS Miller-ovih i RSS McKay-evih intervala koji su zasnovani na odsečenoj sredini bile su najbliže 0,95, dok su za uzorke veličine 20 najbolji rezultati dobijeni primenom McKay-evog i RSS McKay-evog intervala koji su zasnovani na odsečenoj sredini. Za uzorke veličine 40, u zavisnosti od parametra σ log-normalne raspodele, najbolji rezultati ostvareni su primenom RSS Miller-ovog, RSS C&P i McKay-evog intervala koji su zasnovani na odsečenoj sredini. BC, C&P i RSS Miller-ov interval, kao i RSS C&P i RSS McKay-ev interval koji su zasnovani na odsečenoj sredini su bili najbolji izbor za uzorke veličine 80.

Tabela 3.10: Proporcije 95% dvostranih simuliranih intervala za koeficijent varijacije log-normalne raspodele

μ	2				2				2				
σ	0,0998				0,1980				0,3853				
n	10	20	40	80	10	20	40	80	10	20	40	80	
t	0,895	0,913	0,896	0,921	0,906	0,910	0,919	0,897	0,875	0,892	0,829	0,880	
Percentil	0,793	0,841	0,903	0,900	0,777	0,818	0,859	0,896	0,712	0,841	0,853	0,808	
RSS Percentil	$k=2$	0,746	0,847	0,860	0,920	0,750	0,866	0,905	0,908	0,724	0,817	0,826	0,888
	$k=5$	0,694	0,817	0,900	0,917	0,659	0,848	0,900	0,902	0,657	0,787	0,765	0,785
Bootstrap-t		0,853	0,883	0,925	0,926	0,858	0,878	0,904	0,937	0,860	0,824	0,876	0,887
BC		0,800	0,900	0,912	0,947	0,804	0,886	0,909	0,933	0,816	0,868	0,857	0,882
RSS BC	$k=2$	0,818	0,868	0,915	0,914	0,811	0,870	0,893	0,927	0,783	0,797	0,875	0,912
	$k=5$	0,705	0,823	0,902	0,918	0,706	0,844	0,874	0,914	0,662	0,795	0,849	0,876
BCa		0,813	0,876	0,897	0,929	0,807	0,880	0,919	0,924	0,799	0,846	0,865	0,911
RSS BCa	$k=2$	0,840	0,906	0,907	0,925	0,824	0,880	0,905	0,925	0,807	0,845	0,869	0,899
	$k=5$	0,700	0,851	0,907	0,922	0,687	0,881	0,863	0,905	0,678	0,828	0,824	0,902
Jackknife		0,909	0,934	0,943	0,933	0,887	0,923	0,925	0,946	0,876	0,879	0,918	0,912
Miller-ov		0,892	0,925	0,910	0,941	0,880	0,932	0,926	0,909	0,881	0,910	0,915	0,926
Miller-ov trim		0,918	0,941	0,934	0,968	0,915	0,938	0,918	0,919	0,916	0,904	0,920	0,941
Miller-ov trim 10%		0,898	0,922	0,882	0,902	0,893	0,920	0,935	0,939	0,869	0,910	0,899	0,918
RSS Miller-ov	$k=2$	0,925	0,930	0,943	0,910	0,921	0,936	0,927	0,947	0,882	0,907	0,916	0,939
	$k=5$	0,959	0,944	0,960	0,975	0,936	0,980	0,955	0,966	0,936	0,941	0,945	0,957
RSS Miller-ov trim	$k=2$	0,903	0,945	0,964	0,972	0,915	0,908	0,945	0,957	0,878	0,911	0,918	0,904
	$k=5$	0,956	0,954	0,972	0,955	0,942	0,944	0,954	0,963	0,954	0,953	0,950	0,925
RSS Miller-ov trim 10%	$k=2$	0,921	0,915	0,969	0,932	0,918	0,943	0,909	0,933	0,898	0,905	0,895	0,934
	$k=5$	0,954	0,965	0,977	0,956	0,950	0,971	0,967	0,974	0,945	0,946	0,940	0,969
Miller-ov bootstrap-t		0,893	0,936	0,940	0,932	0,877	0,898	0,931	0,941	0,882	0,910	0,952	0,958
PRSS Miller-ov	$r=1$	0,723	0,576	0,443	0,221	0,757	0,665	0,546	0,215	0,723	0,570	0,347	0,325
	$r=2$	0,842	0,837	0,853	0,788	0,808	0,830	0,850	0,656	0,826	0,819	0,722	0,723

Izvor: Preuzeto iz Ivković i Rajić (2019)

⁴⁵ Funkcija gustine log-normalne raspodele nalazi se u odeljku 2.1.4.

Tabela 3.10: Proporcije 95% dvostranih simuliranih intervala za koeficijent varijacije log-normalne raspodele (nastavak)

μ	2				2				2				
σ	0,0998				0,1980				0,3853				
n	10	20	4	80	10	20	40	80	10	20	40	80	
C&P	0,877	0,933	0,919	0,947	0,888	0,904	0,907	0,889	0,870	0,861	0,883	0,915	
C&P trim	0,909	0,940	0,935	0,960	0,909	0,939	0,930	0,936	0,906	0,933	0,933	0,939	
C&P trim 10%	0,886	0,898	0,937	0,931	0,878	0,937	0,948	0,926	0,869	0,915	0,924	0,925	
RSS C&P	$k=2$	0,902	0,935	0,913	0,941	0,875	0,913	0,916	0,907	0,888	0,895	0,926	0,919
	$k=5$	0,964	0,973	0,942	0,977	0,941	0,956	0,947	0,956	0,919	0,915	0,931	0,956
RSS C&P trim	$k=2$	0,880	0,914	0,917	0,941	0,894	0,924	0,938	0,909	0,872	0,933	0,948	0,909
	$k=5$	0,935	0,961	0,975	0,977	0,946	0,942	0,979	0,959	0,928	0,926	0,958	0,927
RSS C&P trim 10%	$k=2$	0,889	0,932	0,949	0,953	0,903	0,906	0,921	0,935	0,890	0,919	0,905	0,935
	$k=5$	0,959	0,957	0,979	0,972	0,939	0,959	0,956	0,954	0,920	0,940	0,927	0,944
C&P bootstrap-t	0,875	0,905	0,922	0,945	0,902	0,898	0,922	0,935	0,912	0,869	0,926	0,920	
PRSS C&P	$r=1$	0,712	0,676	0,571	0,341	0,683	0,596	0,414	0,344	0,653	0,617	0,489	0,322
	$r=2$	0,810	0,832	0,833	0,814	0,815	0,869	0,781	0,754	0,784	0,781	0,753	0,738
McKay-ev	0,941	0,943	0,957	0,964	0,943	0,931	0,963	0,934	0,969	0,938	0,957	0,906	
McKay-ev trim	0,941	0,955	0,954	0,946	0,956	0,945	0,949	0,936	0,955	0,946	0,939	0,934	
McKay-ev trim 10%	0,959	0,933	0,953	0,924	0,930	0,959	0,936	0,934	0,943	0,971	0,942	0,937	
RSS McKay-ev	$k=2$	0,944	0,956	0,940	0,931	0,956	0,943	0,947	0,967	0,944	0,966	0,907	0,936
	$k=5$	0,964	0,976	0,967	0,977	0,953	0,969	0,955	0,969	0,983	0,970	0,947	0,946
RSS McKay-ev trim	$k=2$	0,938	0,940	0,933	0,957	0,953	0,921	0,945	0,935	0,939	0,951	0,941	0,921
	$k=5$	0,947	0,968	0,979	0,978	0,983	0,977	0,956	0,956	0,961	0,947	0,953	0,943
RSS McKay-ev trim 10%	$k=2$	0,947	0,950	0,937	0,936	0,929	0,939	0,934	0,928	0,954	0,938	0,931	0,948
	$k=5$	0,972	0,976	0,967	0,977	0,952	0,965	0,981	0,965	0,977	0,966	0,954	0,923
PRSS McKay-ev	$r=1$	0,893	0,823	0,704	0,402	0,934	0,847	0,638	0,457	0,936	0,871	0,673	0,415
	$r=2$	0,938	0,923	0,909	0,834	0,946	0,923	0,878	0,865	0,942	0,911	0,865	0,858

Predmet ispitivanja u četvrtom slučaju bila je *skew-normalna*⁴⁶ raspodela sa parametrima $\beta = 5$, $\alpha = 2$ i parametrom γ koji uzima vrednosti 0,7947; 1,7928 i 4,8186. Kako bi se dobile odgovarajuće vrednosti koeficijenta varijacije osnovnog skupa izvršena je opisana procedura. U Tabeli 3.11 predstavljene su proporcije simuliranih 95% dvostranih intervala za koeficijent varijacije *skew-normalne* raspodele. Za uzorke veličine 10, u zavisnosti od parametra γ *skew-normalne* raspodele, primenom RSS C&P i PRSS McKay-evog intervala, kao i RSS C&P intervala zasnovanog na odsečenoj sredini ostvareni su najbolji rezultati. Kada je reč o uzorcima veličine 20, najbolji izbor su bili Miller-ov *bootstrap-t* i RSS McKay-ev interval, kao i McKay-ev interval zasnovan na odsečenoj sredini. Primenom RSS Miller-ovog intervala, kao i RSS Miller-ovog i RSS McKay-evog intervala koji su zasnovani na odsečenoj sredini za uzorke veličine 40 dobijeni su rezultati najbliži 0,95. Za uzorke veličine 80 koristeći C&P, RSS C&P, McKay-ev i RSS McKay-ev interval koji su zasnovani na odsečenoj sredini dobijeni su najbolji rezultati.

Tabela 3.11: Proporcije 95% dvostranih simuliranih intervala za koeficijent varijacije *skew-normalne* raspodele

β	5				5				5				
γ	0,7947				1,7928				4,8186				
α	2				2				2				
n	10	20	40	80	10	20	40	80	10	20	40	80	
t	0,923	0,913	0,918	0,915	0,917	0,913	0,921	0,905	0,920	0,913	0,906	0,931	
Percentil	0,808	0,810	0,862	0,921	0,782	0,838	0,890	0,909	0,800	0,876	0,882	0,917	
RSS	$k=2$	0,770	0,866	0,902	0,905	0,796	0,875	0,900	0,929	0,809	0,898	0,908	0,904
Percentil	$k=5$	0,685	0,832	0,894	0,926	0,707	0,845	0,898	0,875	0,709	0,870	0,915	0,928
Bootstrap-t		0,850	0,880	0,918	0,939	0,847	0,896	0,916	0,923	0,907	0,924	0,917	0,946
BC		0,789	0,885	0,920	0,919	0,795	0,878	0,928	0,920	0,820	0,870	0,908	0,941
RSS BC	$k=2$	0,801	0,901	0,906	0,935	0,815	0,857	0,913	0,934	0,807	0,890	0,920	0,911
	$k=5$	0,719	0,869	0,886	0,917	0,703	0,838	0,891	0,877	0,721	0,826	0,848	0,945
Bca		0,805	0,880	0,928	0,923	0,806	0,901	0,906	0,923	0,840	0,899	0,935	0,921
RSS BCa	$k=2$	0,838	0,891	0,908	0,929	0,822	0,888	0,909	0,915	0,841	0,889	0,911	0,923
	$k=5$	0,714	0,829	0,894	0,897	0,700	0,880	0,882	0,923	0,738	0,867	0,893	0,915
Jackknife		0,913	0,902	0,926	0,921	0,921	0,939	0,944	0,935	0,909	0,923	0,929	0,943
Miller-ov		0,888	0,921	0,920	0,914	0,912	0,926	0,943	0,930	0,920	0,930	0,942	0,934
Miller-ov trim		0,888	0,906	0,953	0,940	0,908	0,920	0,942	0,926	0,906	0,922	0,967	0,979
Miller-ov trim 10%		0,864	0,895	0,934	0,918	0,909	0,932	0,933	0,960	0,905	0,930	0,938	0,980
RSS Miller-ov	$k=2$	0,906	0,928	0,933	0,911	0,919	0,941	0,915	0,936	0,925	0,942	0,954	0,979
	$k=5$	0,929	0,953	0,929	0,970	0,943	0,954	0,956	0,938	0,952	0,968	0,981	0,990
RSS Miller-ov trim	$k=2$	0,922	0,915	0,925	0,937	0,903	0,943	0,915	0,911	0,917	0,923	0,927	0,974
	$k=5$	0,971	0,981	0,965	0,977	0,938	0,963	0,952	0,985	0,965	0,976	0,982	0,985
RSS Miller-ov trim 10%	$k=2$	0,906	0,914	0,932	0,903	0,884	0,953	0,950	0,934	0,937	0,957	0,939	0,982
	$k=5$	0,942	0,953	0,974	0,968	0,951	0,954	0,964	0,963	0,967	0,983	0,974	0,965
Miller-ov bootstrap-t		0,860	0,899	0,948	0,944	0,889	0,913	0,928	0,951	0,952	0,953	0,963	0,975
PRSS Miller-ov	$r=1$	0,730	0,582	0,541	0,282	0,708	0,699	0,423	0,243	0,732	0,700	0,544	0,324
	$r=2$	0,828	0,766	0,775	0,777	0,841	0,854	0,812	0,789	0,837	0,831	0,885	0,753

Izvor: Preuzeto iz Ivković i Rajić (2019)

⁴⁶ Funkcija gustine *skew-normalne* raspodele nalazi se u odeljku 2.1.5.

Tabela 3.11: Proporcije 95% dvostranih simuliranih intervala za koeficijent varijacije *skew-normalne* raspodele (nastavak)

β	5				5				5				
γ	0,7947				1,7928				4,8186				
α	2				2				2				
n	10	20	40	80	10	20	40	80	10	20	40	80	
C&P	0,872	0,931	0,941	0,937	0,909	0,933	0,931	0,956	0,896	0,914	0,966	0,956	
C&P trim	0,879	0,937	0,940	0,931	0,894	0,928	0,946	0,950	0,903	0,946	0,958	0,969	
C&P trim 10%	0,898	0,912	0,922	0,930	0,910	0,930	0,919	0,944	0,902	0,931	0,919	0,936	
RSS C&P	$k=2$	0,871	0,925	0,926	0,940	0,879	0,929	0,943	0,942	0,890	0,924	0,937	0,972
	$k=5$	0,945	0,967	0,965	0,974	0,905	0,961	0,967	0,964	0,950	0,971	0,976	0,989
RSS C&P trim	$k=2$	0,880	0,910	0,916	0,933	0,901	0,943	0,943	0,946	0,918	0,945	0,935	0,952
	$k=5$	0,938	0,934	0,971	0,977	0,953	0,958	0,961	0,972	0,958	0,970	0,974	0,987
RSS C&P trim 10%	$k=2$	0,895	0,913	0,908	0,949	0,907	0,941	0,934	0,927	0,884	0,928	0,966	0,953
	$k=5$	0,943	0,973	0,940	0,978	0,950	0,962	0,970	0,932	0,954	0,961	0,957	0,992
C&P bootstrap-t	0,859	0,886	0,915	0,934	0,893	0,886	0,940	0,954	0,949	0,961	0,963	0,967	
PRSS C&P	$r=1$	0,694	0,592	0,558	0,234	0,696	0,613	0,579	0,182	0,748	0,725	0,543	0,492
	$r=2$	0,792	0,801	0,818	0,791	0,830	0,840	0,823	0,772	0,822	0,855	0,876	0,831
McKay-ev	0,944	0,946	0,956	0,955	0,954	0,957	0,938	0,959	0,973	0,964	0,971	0,973	
McKay-ev trim	0,938	0,947	0,947	0,952	0,957	0,948	0,961	0,950	0,966	0,970	0,962	0,970	
McKay-ev trim 10%	0,925	0,951	0,952	0,940	0,956	0,947	0,965	0,960	0,970	0,961	0,974	0,931	
RSS McKay-ev	$k=2$	0,960	0,950	0,946	0,947	0,943	0,940	0,975	0,923	0,974	0,976	0,973	0,984
	$k=5$	0,966	0,959	0,978	0,974	0,982	0,969	0,956	0,983	0,992	0,995	0,989	0,985
RSS McKay-ev trim	$k=2$	0,942	0,957	0,951	0,943	0,944	0,961	0,926	0,939	0,971	0,980	0,975	0,972
	$k=5$	0,967	0,977	0,962	0,973	0,966	0,972	0,964	0,985	0,989	0,986	0,989	0,997
RSS McKay-ev trim 10%	$k=2$	0,935	0,926	0,927	0,925	0,955	0,970	0,965	0,950	0,974	0,973	0,965	0,973
	$k=5$	0,980	0,970	0,974	0,972	0,964	0,974	0,981	0,981	0,988	0,994	0,990	0,989
PRSS McKay-ev	$r=1$	0,909	0,785	0,505	0,257	0,920	0,867	0,693	0,319	0,950	0,912	0,697	0,477
	$r=2$	0,945	0,904	0,894	0,852	0,947	0,935	0,885	0,896	0,955	0,957	0,909	0,932

U nastavku će se uporediti širine postojećih intervala poverenja (Miller-ov, Curto-ov i Pinto-ov i McKay-ev interval poverenja zasnovan na prostom slučajnom uzorkovanju i rangiranom uzorkovanju) sa širinama postojećih intervala koji su zasnovani na odsečnoj sredini. U Tabeli 3.12 predstavljena je prosečna širina 95% dvostranog intervala za koeficijent varijacije gama raspodele.

Tabela 3.12: Prosečna širina 95% dvostranog intervala za koeficijent varijacije gama raspodele

<i>a</i>		100				25				6,25			
<i>b</i>		2				2				2			
<i>n</i>		10	20	40	80	10	20	40	80	10	20	40	80
<i>t</i>		0,097	0,066	0,045	0,031	0,198	0,133	0,093	0,064	0,400	0,266	0,177	0,124
Percentil		0,076	0,057	0,040	0,029	0,154	0,111	0,082	0,061	0,296	0,219	0,170	0,109
RSS Percentil	<i>k=2</i>	0,068	0,055	0,041	0,030	0,140	0,111	0,079	0,058	0,282	0,211	0,157	0,124
	<i>k=5</i>	0,049	0,043	0,034	0,025	0,098	0,089	0,069	0,051	0,189	0,170	0,142	0,112
Bootstrap-t		0,123	0,069	0,043	0,030	0,242	0,134	0,092	0,065	0,465	0,278	0,186	0,148
BC		0,071	0,054	0,041	0,031	0,141	0,110	0,082	0,060	0,278	0,215	0,173	0,133
RSS BC	<i>k=2</i>	0,070	0,054	0,039	0,029	0,134	0,108	0,083	0,061	0,268	0,217	0,166	0,116
	<i>k=5</i>	0,047	0,042	0,034	0,026	0,094	0,085	0,069	0,053	0,185	0,174	0,147	0,112
Bca		0,073	0,056	0,042	0,031	0,140	0,112	0,084	0,061	0,296	0,242	0,175	0,127
RSS Bca	<i>k=2</i>	0,070	0,055	0,042	0,031	0,140	0,111	0,081	0,063	0,277	0,220	0,173	0,129
	<i>k=5</i>	0,045	0,044	0,035	0,026	0,089	0,088	0,075	0,050	0,182	0,170	0,136	0,112
Jackknife		0,107	0,068	0,046	0,031	0,212	0,131	0,092	0,063	0,414	0,258	0,197	0,135
Miller-ov		0,091	0,064	0,045	0,032	0,184	0,129	0,091	0,063	0,402	0,271	0,204	0,148
Miller-ov trim		0,092	0,064	0,044	0,032	0,185	0,130	0,093	0,064	0,422	0,290	0,201	0,141
Miller-ov trim 10%		0,089	0,060	0,045	0,032	0,191	0,129	0,091	0,064	0,416	0,278	0,200	0,146
RSS Miller-ov	<i>k=2</i>	0,093	0,064	0,044	0,030	0,198	0,137	0,091	0,064	0,427	0,296	0,203	0,145
	<i>k=5</i>	0,093	0,065	0,046	0,031	0,191	0,135	0,092	0,064	0,426	0,300	0,195	0,142
RSS Miller-ov trim	<i>k=2</i>	0,093	0,064	0,044	0,031	0,190	0,132	0,092	0,066	0,436	0,305	0,205	0,140
	<i>k=5</i>	0,095	0,064	0,044	0,032	0,197	0,132	0,093	0,065	0,432	0,293	0,203	0,137
RSS Miller-ov trim 10%	<i>k=2</i>	0,091	0,062	0,044	0,031	0,190	0,132	0,093	0,065	0,422	0,297	0,210	0,141
	<i>k=5</i>	0,097	0,063	0,044	0,032	0,201	0,130	0,091	0,066	0,418	0,294	0,209	0,139
Miller-ov bootstrap-t		0,128	0,069	0,046	0,031	0,252	0,139	0,091	0,068	0,614	0,318	0,219	0,157
PRSS Miller-ov	<i>r=1</i>	0,075	0,053	0,038	0,026	0,155	0,107	0,076	0,053	0,339	0,230	0,163	0,112
	<i>r=2</i>	0,083	0,059	0,041	0,028	0,175	0,121	0,086	0,060	0,369	0,252	0,180	0,128

Izvor: Preuzeto iz Ivković i Rajić (2019)

Tabela 3.12: Prosečna širina 95% dvostranog intervala za koeficijent varijacije gama raspodele (nastavak)

a	100				25				6,25				
b	2				2				2				
n	10	20	40	80	10	20	40	80	10	20	40	80	
C&P	0,085	0,063	0,044	0,030	0,182	0,126	0,091	0,065	0,397	0,289	0,197	0,148	
C&P trim	0,086	0,061	0,045	0,032	0,178	0,125	0,088	0,063	0,391	0,293	0,206	0,140	
C&P trim 10%	0,085	0,061	0,044	0,031	0,176	0,128	0,087	0,064	0,382	0,281	0,200	0,141	
RSS C&P	$k=2$	0,088	0,063	0,043	0,031	0,178	0,127	0,089	0,063	0,406	0,280	0,199	0,147
	$k=5$	0,091	0,063	0,044	0,031	0,184	0,128	0,089	0,065	0,416	0,287	0,203	0,144
RSS C&P trim	$k=2$	0,090	0,064	0,045	0,031	0,187	0,126	0,092	0,065	0,422	0,285	0,206	0,139
	$k=5$	0,090	0,063	0,044	0,032	0,185	0,128	0,090	0,065	0,413	0,285	0,202	0,142
RSS C&P trim 10%	$k=2$	0,087	0,063	0,044	0,031	0,181	0,125	0,091	0,065	0,403	0,282	0,202	0,146
	$k=5$	0,093	0,064	0,045	0,032	0,185	0,131	0,093	0,063	0,416	0,286	0,202	0,142
C&P bootstrap-t	0,122	0,067	0,048	0,031	0,245	0,139	0,096	0,067	0,548	0,322	0,225	0,148	
PRSS C&P	$r=1$	0,072	0,049	0,036	0,026	0,147	0,106	0,074	0,052	0,318	0,223	0,160	0,116
	$r=2$	0,080	0,056	0,041	0,029	0,161	0,119	0,081	0,059	0,368	0,261	0,186	0,130
McKay-ev	0,115	0,071	0,046	0,032	0,239	0,150	0,097	0,067	0,702	0,347	0,224	0,150	
McKay-ev trim	0,112	0,069	0,047	0,032	0,249	0,145	0,099	0,067	0,785	0,362	0,227	0,149	
McKay-ev trim 10%	0,114	0,070	0,047	0,033	0,237	0,154	0,096	0,066	0,730	0,356	0,210	0,148	
RSS McKay-ev	$k=2$	0,115	0,071	0,046	0,033	0,245	0,146	0,096	0,068	0,741	0,348	0,218	0,153
	$k=5$	0,120	0,071	0,048	0,033	0,257	0,153	0,099	0,067	0,773	0,345	0,226	0,150
RSS McKay-ev trim	$k=2$	0,117	0,073	0,046	0,032	0,247	0,148	0,097	0,066	0,836	0,353	0,221	0,149
	$k=5$	0,117	0,071	0,047	0,033	0,258	0,149	0,099	0,066	0,741	0,358	0,235	0,148
RSS McKay-ev trim 10%	$k=2$	0,115	0,073	0,046	0,032	0,249	0,149	0,097	0,069	0,702	0,351	0,228	0,146
	$k=5$	0,119	0,071	0,047	0,032	0,259	0,147	0,102	0,065	0,800	0,355	0,227	0,149
PRSS McKay-ev	$r=1$	0,092	0,058	0,039	0,027	0,193	0,117	0,080	0,055	0,492	0,272	0,167	0,124
	$r=2$	0,102	0,065	0,043	0,030	0,225	0,138	0,091	0,059	0,640	0,328	0,205	0,139

Prosečna širina 95% dvostranog intervala za koeficijent varijacije Weibull-ove raspodele prikazana je u Tabeli 3.13.

Tabela 3.13: Prosečna širina 95% dvostranog intervala za koeficijent varijacije Weibull-ove raspodele

<i>a</i>	12,1543				5,80				2,70				
<i>b</i>	2				2				2				
<i>n</i>	10	20	40	80	10	20	40	80	10	20	40	80	
<i>t</i>	0,096	0,064	0,041	0,031	0,204	0,129	0,090	0,063	0,394	0,263	0,182	0,124	
Percentil	0,080	0,062	0,048	0,035	0,164	0,129	0,094	0,066	0,326	0,228	0,169	0,122	
RSS													
Percentil	<i>k=2</i>	0,079	0,062	0,049	0,037	0,154	0,118	0,084	0,063	0,295	0,231	0,161	0,117
	<i>k=5</i>	0,054	0,051	0,041	0,029	0,099	0,095	0,070	0,054	0,190	0,174	0,135	0,097
Bootstrap-t		0,138	0,080	0,053	0,038	0,256	0,140	0,100	0,067	0,473	0,280	0,184	0,126
BC		0,078	0,064	0,047	0,035	0,159	0,126	0,088	0,066	0,317	0,235	0,170	0,121
RSS BC	<i>k=2</i>	0,075	0,062	0,047	0,037	0,155	0,118	0,091	0,067	0,284	0,225	0,165	0,118
	<i>k=5</i>	0,052	0,049	0,039	0,031	0,097	0,094	0,072	0,055	0,184	0,169	0,132	0,098
Bca		0,076	0,066	0,050	0,039	0,157	0,127	0,098	0,066	0,325	0,241	0,170	0,123
RSS Bca	<i>k=2</i>	0,082	0,066	0,051	0,040	0,150	0,129	0,094	0,064	0,288	0,235	0,154	0,117
	<i>k=5</i>	0,047	0,052	0,045	0,035	0,097	0,100	0,081	0,054	0,190	0,172	0,133	0,100
Jackknife		0,119	0,078	0,054	0,039	0,231	0,147	0,102	0,069	0,441	0,271	0,185	0,122
Miller-ov		0,089	0,060	0,044	0,031	0,193	0,132	0,091	0,065	0,415	0,295	0,200	0,145
Miller-ov trim		0,096	0,064	0,046	0,031	0,196	0,132	0,094	0,064	0,428	0,289	0,193	0,141
Miller-ov trim 10%		0,089	0,063	0,044	0,031	0,186	0,125	0,091	0,063	0,445	0,286	0,202	0,136
RSS Miller-ov	<i>k=2</i>	0,094	0,062	0,045	0,033	0,193	0,140	0,094	0,064	0,427	0,291	0,209	0,145
	<i>k=5</i>	0,096	0,066	0,044	0,033	0,190	0,129	0,091	0,066	0,441	0,297	0,211	0,144
RSS Miller-ov trim	<i>k=2</i>	0,096	0,062	0,045	0,032	0,196	0,131	0,088	0,064	0,426	0,298	0,202	0,145
	<i>k=5</i>	0,097	0,067	0,044	0,032	0,194	0,126	0,091	0,064	0,433	0,304	0,203	0,142
RSS Miller-ov trim 10%	<i>k=2</i>	0,091	0,067	0,045	0,031	0,193	0,130	0,090	0,066	0,429	0,304	0,201	0,142
	<i>k=5</i>	0,095	0,067	0,044	0,032	0,199	0,133	0,094	0,064	0,462	0,301	0,200	0,147
Miller-ov bootstrap-t		0,137	0,085	0,058	0,038	0,297	0,171	0,103	0,073	0,570	0,335	0,212	0,146
PRSS Miller-ov	<i>r=1</i>	0,075	0,052	0,035	0,026	0,153	0,106	0,075	0,053	0,338	0,237	0,163	0,114
	<i>r=2</i>	0,082	0,058	0,041	0,029	0,175	0,120	0,085	0,060	0,388	0,264	0,188	0,135

Izvor: Preuzeto iz Ivković i Rajić (2019)

Tabela 3.13: Prosečna širina 95% dvostranog intervala za koeficijent varijacije Weibull-ove raspodele (nastavak)

<i>a</i>	12,1543				5,80				2,70				
<i>b</i>	2				2				2				
<i>n</i>	10	20	40	80	10	20	40	80	10	20	40	80	
C&P	0,084	0,062	0,043	0,031	0,185	0,127	0,093	0,066	0,411	0,282	0,201	0,143	
C&P trim	0,089	0,064	0,045	0,031	0,178	0,131	0,088	0,065	0,421	0,272	0,194	0,137	
C&P trim 10%	0,086	0,062	0,044	0,031	0,180	0,131	0,095	0,063	0,409	0,279	0,196	0,145	
RSS C&P	<i>k=2</i>	0,089	0,060	0,043	0,031	0,189	0,129	0,091	0,064	0,391	0,273	0,201	0,145
	<i>k=5</i>	0,090	0,063	0,045	0,031	0,185	0,124	0,092	0,063	0,412	0,295	0,209	0,145
RSS C&P trim	<i>k=2</i>	0,086	0,062	0,045	0,031	0,180	0,128	0,089	0,067	0,405	0,288	0,204	0,143
	<i>k=5</i>	0,089	0,065	0,044	0,031	0,183	0,131	0,093	0,065	0,418	0,285	0,207	0,143
RSS C&P trim 10%	<i>k=2</i>	0,083	0,062	0,047	0,031	0,187	0,129	0,092	0,063	0,405	0,280	0,204	0,134
	<i>k=5</i>	0,087	0,062	0,047	0,032	0,181	0,129	0,093	0,064	0,416	0,288	0,202	0,142
C&P bootstrap-t	0,134	0,081	0,051	0,038	0,280	0,157	0,108	0,071	0,562	0,333	0,204	0,147	
PRSS C&P	<i>r=1</i>	0,069	0,050	0,035	0,025	0,147	0,104	0,074	0,052	0,330	0,231	0,156	0,114
	<i>r=2</i>	0,078	0,055	0,042	0,029	0,165	0,120	0,084	0,060	0,377	0,253	0,185	0,129
McKay-ev	0,112	0,070	0,044	0,032	0,249	0,148	0,096	0,069	0,811	0,369	0,219	0,149	
McKay-ev trim	0,111	0,072	0,047	0,032	0,247	0,141	0,095	0,068	0,795	0,336	0,234	0,150	
McKay-ev trim 10%	0,105	0,070	0,046	0,033	0,255	0,149	0,100	0,066	0,781	0,358	0,234	0,146	
RSS McKay-ev	<i>k=2</i>	0,113	0,070	0,045	0,033	0,250	0,149	0,095	0,064	0,701	0,358	0,227	0,155
	<i>k=5</i>	0,121	0,072	0,046	0,032	0,258	0,146	0,097	0,066	0,733	0,349	0,221	0,149
RSS McKay-ev trim	<i>k=2</i>	0,123	0,069	0,047	0,032	0,255	0,144	0,099	0,066	0,758	0,370	0,231	0,144
	<i>k=5</i>	0,117	0,072	0,046	0,032	0,260	0,152	0,100	0,067	0,774	0,368	0,223	0,145
RSS McKay-ev trim 10%	<i>k=2</i>	0,116	0,072	0,047	0,032	0,248	0,152	0,096	0,067	0,745	0,351	0,226	0,153
	<i>k=5</i>	0,119	0,071	0,047	0,032	0,260	0,151	0,099	0,067	0,754	0,358	0,222	0,155
PRSS McKay-ev	<i>r=1</i>	0,092	0,057	0,038	0,026	0,202	0,120	0,079	0,053	0,501	0,268	0,179	0,119
	<i>r=2</i>	0,103	0,065	0,042	0,029	0,235	0,133	0,087	0,061	0,596	0,318	0,203	0,133

U Tabeli 3.14 predstavljena je prosečna širina 95% dvostranog intervala za koeficijent varijacije log-normalne raspodele.

Tabela 3.14: Prosečna širina 95% dvostranog intervala za koeficijent varijacije log-normalne raspodele

μ	2				2				2				
σ	0,0998				0,1980				0,3853				
n	10	20	40	80	10	20	40	80	10	20	40	80	
t	0,101	0,067	0,046	0,032	0,196	0,133	0,088	0,062	0,378	0,252	0,179	0,124	
Percentil	0,074	0,055	0,042	0,030	0,146	0,112	0,085	0,063	0,288	0,228	0,189	0,135	
RSS													
Percentil	$k=2$	0,071	0,058	0,042	0,031	0,134	0,111	0,089	0,062	0,270	0,221	0,176	0,137
	$k=5$	0,048	0,044	0,034	0,026	0,094	0,092	0,070	0,057	0,187	0,178	0,145	0,115
<i>Bootstrap-t</i>		0,120	0,069	0,046	0,032	0,254	0,134	0,096	0,064	0,516	0,298	0,197	0,161
BC		0,066	0,056	0,039	0,031	0,136	0,114	0,081	0,062	0,279	0,221	0,177	0,131
RSS BC	$k=2$	0,071	0,054	0,040	0,031	0,140	0,109	0,084	0,062	0,250	0,227	0,178	0,145
	$k=5$	0,049	0,049	0,036	0,026	0,096	0,089	0,073	0,057	0,182	0,182	0,146	0,132
Bca		0,072	0,057	0,042	0,030	0,138	0,110	0,085	0,064	0,282	0,234	0,188	0,147
RSS Bca	$k=2$	0,072	0,057	0,040	0,032	0,136	0,108	0,083	0,060	0,268	0,214	0,179	0,140
	$k=5$	0,048	0,046	0,037	0,027	0,090	0,091	0,076	0,058	0,179	0,178	0,175	0,126
<i>Jackknife</i>		0,105	0,070	0,045	0,031	0,207	0,143	0,093	0,062	0,436	0,301	0,190	0,153
Miller-ov		0,093	0,064	0,044	0,031	0,187	0,134	0,090	0,065	0,401	0,276	0,201	0,140
Miller-ov trim		0,089	0,061	0,045	0,032	0,187	0,133	0,094	0,064	0,430	0,290	0,206	0,135
Miller-ov trim 10%		0,092	0,062	0,045	0,031	0,184	0,131	0,088	0,064	0,409	0,293	0,202	0,144
RSS													
Miller-ov	$k=2$	0,096	0,064	0,045	0,033	0,186	0,134	0,094	0,064	0,396	0,265	0,204	0,143
	$k=5$	0,093	0,065	0,045	0,031	0,199	0,133	0,096	0,066	0,418	0,284	0,215	0,137
RSS													
Miller-ov trim	$k=2$	0,090	0,064	0,045	0,032	0,188	0,129	0,090	0,065	0,416	0,301	0,207	0,150
	$k=5$	0,095	0,065	0,045	0,032	0,197	0,131	0,094	0,064	0,440	0,300	0,198	0,141
RSS													
Miller-ov trim 10%	$k=2$	0,091	0,063	0,043	0,031	0,188	0,133	0,091	0,066	0,423	0,277	0,201	0,144
	$k=5$	0,096	0,066	0,044	0,031	0,196	0,134	0,090	0,068	0,451	0,302	0,210	0,1454
Miller-ov bootstrap-t		0,128	0,071	0,045	0,031	0,268	0,149	0,097	0,065	0,592	0,372	0,242	0,168
PRSS													
Miller-ov	$r=1$	0,078	0,052	0,038	0,025	0,154	0,109	0,074	0,053	0,307	0,230	0,164	0,116
	$r=2$	0,084	0,058	0,042	0,029	0,169	0,123	0,081	0,059	0,389	0,258	0,184	0,134

Izvor: Preuzeto iz Ivković i Rajić (2019)

Tabela 3.14: Prosečna širina 95% dvostranog intervala za koeficijent varijacije log-normalne raspodele (nastavak)

μ	2				2				2				
σ	0,0998				0,1980				0,3853				
n	10	20	4	80	10	20	40	80	10	20	40	80	
C&P	0,086	0,062	0,044	0,030	0,173	0,123	0,088	0,066	0,396	0,271	0,192	0,149	
C&P trim	0,087	0,061	0,044	0,031	0,183	0,129	0,090	0,065	0,393	0,273	0,197	0,143	
C&P trim 10%	0,088	0,062	0,044	0,032	0,184	0,124	0,091	0,063	0,374	0,283	0,206	0,139	
RSS C&P	$k=2$	0,088	0,061	0,043	0,032	0,175	0,127	0,096	0,065	0,396	0,271	0,192	0,145
	$k=5$	0,090	0,063	0,044	0,031	0,185	0,131	0,089	0,065	0,404	0,287	0,201	0,140
RSS C&P trim	$k=2$	0,088	0,063	0,044	0,029	0,183	0,128	0,092	0,065	0,413	0,276	0,195	0,144
	$k=5$	0,090	0,063	0,044	0,031	0,191	0,131	0,090	0,063	0,398	0,288	0,190	0,143
RSS C&P trim 10%	$k=2$	0,085	0,062	0,044	0,032	0,182	0,127	0,088	0,064	0,383	0,302	0,200	0,148
	$k=5$	0,086	0,063	0,044	0,032	0,186	0,133	0,096	0,065	0,405	0,289	0,199	0,137
C&P bootstrap-t	0,118	0,068	0,045	0,030	0,261	0,143	0,098	0,066	0,612	0,315	0,245	0,165	
PRSS C&P	$r=1$	0,071	0,051	0,035	0,025	0,144	0,101	0,072	0,052	0,301	0,224	0,154	0,113
	$r=2$	0,078	0,056	0,041	0,029	0,166	0,115	0,083	0,060	0,355	0,255	0,187	0,129
McKay-ev	0,109	0,071	0,046	0,032	0,238	0,142	0,095	0,067	0,676	0,363	0,222	0,155	
McKay-ev trim	0,113	0,071	0,047	0,033	0,236	0,144	0,097	0,068	0,686	0,347	0,232	0,153	
McKay-ev trim 10%	0,116	0,071	0,047	0,032	0,252	0,145	0,098	0,067	0,680	0,347	0,220	0,146	
RSS McKay-ev	$k=2$	0,118	0,067	0,046	0,032	0,250	0,147	0,097	0,067	0,707	0,356	0,222	0,151
	$k=5$	0,116	0,072	0,046	0,032	0,250	0,146	0,099	0,066	0,686	0,347	0,230	0,147
RSS McKay-ev trim	$k=2$	0,109	0,070	0,048	0,033	0,238	0,145	0,101	0,065	0,740	0,373	0,233	0,143
	$k=5$	0,115	0,071	0,046	0,032	0,247	0,146	0,098	0,066	0,777	0,368	0,218	0,146
RSS McKay-ev trim 10%	$k=2$	0,118	0,073	0,047	0,032	0,243	0,145	0,094	0,066	0,718	0,342	0,228	0,150
	$k=5$	0,121	0,075	0,049	0,032	0,249	0,148	0,097	0,068	0,776	0,355	0,226	0,142
PRSS McKay-ev	$r=1$	0,095	0,059	0,038	0,027	0,198	0,119	0,081	0,054	0,481	0,269	0,173	0,119
	$r=2$	0,107	0,064	0,043	0,030	0,220	0,137	0,090	0,059	0,567	0,338	0,204	0,130

Prosečna širina 95% dvostranog intervala za koeficijent varijacije *skew*-normalne raspodele prikazana je u Tabeli 3.15.

Tabela 3.15: Prosečna širina 95% dvostranog intervala za koeficijent varijacije *skew-normalne* raspodele

β	5				5				5				
γ	0,7947				1,7928				4,8186				
α	2				2				2				
n	10	20	40	80	10	20	40	80	10	20	40	80	
t	0,097	0,064	0,045	0,031	0,205	0,132	0,090	0,062	0,393	0,264	0,179	0,125	
Percentil	0,073	0,055	0,042	0,031	0,152	0,115	0,086	0,060	0,328	0,243	0,168	0,126	
RSS	$k=2$	0,071	0,056	0,041	0,031	0,147	0,111	0,080	0,065	0,294	0,233	0,172	0,123
Percentil	$k=5$	0,049	0,046	0,036	0,026	0,093	0,093	0,069	0,052	0,198	0,183	0,137	0,107
Bootstrap-t		0,123	0,067	0,047	0,032	0,247	0,137	0,091	0,064	0,499	0,281	0,184	0,137
BC		0,069	0,055	0,040	0,029	0,138	0,111	0,081	0,059	0,320	0,242	0,169	0,124
RSS BC	$k=2$	0,070	0,055	0,041	0,029	0,142	0,108	0,082	0,060	0,275	0,234	0,170	0,121
	$k=5$	0,048	0,048	0,035	0,027	0,097	0,083	0,071	0,051	0,192	0,186	0,151	0,104
Bca		0,068	0,058	0,041	0,032	0,143	0,110	0,086	0,062	0,339	0,257	0,178	0,129
RSS Bca	$k=2$	0,072	0,059	0,045	0,031	0,141	0,113	0,086	0,061	0,296	0,227	0,169	0,125
	$k=5$	0,046	0,047	0,034	0,028	0,096	0,084	0,073	0,053	0,188	0,189	0,142	0,110
Jackknife		0,103	0,067	0,047	0,033	0,206	0,139	0,095	0,063	0,435	0,280	0,203	0,134
Miller-ov		0,090	0,063	0,044	0,031	0,186	0,131	0,091	0,065	0,421	0,289	0,205	0,142
Miller-ov trim		0,090	0,063	0,044	0,031	0,196	0,129	0,093	0,064	0,419	0,294	0,213	0,145
Miller-ov trim 10%		0,092	0,063	0,044	0,031	0,185	0,132	0,094	0,065	0,412	0,292	0,213	0,145
RSS Miller-ov	$k=2$	0,089	0,065	0,045	0,031	0,188	0,129	0,094	0,065	0,414	0,291	0,203	0,140
	$k=5$	0,097	0,065	0,045	0,031	0,196	0,128	0,090	0,065	0,448	0,300	0,208	0,144
RSS Miller-ov trim	$k=2$	0,094	0,063	0,044	0,030	0,195	0,133	0,092	0,065	0,433	0,287	0,212	0,135
	$k=5$	0,093	0,065	0,044	0,032	0,195	0,135	0,093	0,063	0,441	0,295	0,198	0,148
RSS Miller-ov trim 10%	$k=2$	0,090	0,065	0,044	0,031	0,190	0,133	0,091	0,063	0,449	0,298	0,206	0,141
	$k=5$	0,092	0,066	0,044	0,031	0,190	0,131	0,094	0,062	0,446	0,296	0,204	0,141
Miller-ov bootstrap-t		0,133	0,076	0,045	0,032	0,255	0,148	0,100	0,065	0,625	0,348	0,221	0,147
PRSS Miller-ov	$r=1$	0,074	0,052	0,036	0,025	0,154	0,107	0,077	0,053	0,331	0,226	0,165	0,112
	$r=2$	0,083	0,059	0,041	0,030	0,174	0,123	0,082	0,060	0,378	0,267	0,189	0,126

Izvor: Preuzeto iz Ivković i Rajić (2019)

Tabela 3.15: Prosečna širina 95% dvostranog intervala za koeficijent varijacije *skew*-normalne raspodele (nastavak)

β	5				5				5				
γ	0,7947				1,7928				4,8186				
α	2				2				2				
n	10	20	40	8	10	20	40	80	10	20	40	80	
C&P	0,087	0,061	0,044	0,031	0,178	0,130	0,093	0,066	0,388	0,276	0,204	0,145	
C&P trim	0,084	0,062	0,044	0,030	0,185	0,129	0,090	0,063	0,405	0,283	0,196	0,133	
C&P trim 10%	0,086	0,061	0,044	0,032	0,178	0,127	0,089	0,066	0,394	0,270	0,202	0,146	
RSS C&P	$k=2$	0,086	0,061	0,044	0,031	0,174	0,128	0,089	0,065	0,402	0,288	0,202	0,138
	$k=5$	0,092	0,063	0,045	0,031	0,182	0,130	0,091	0,066	0,422	0,302	0,203	0,149
RSS C&P trim	$k=2$	0,084	0,062	0,044	0,031	0,183	0,129	0,088	0,064	0,407	0,283	0,199	0,147
	$k=5$	0,089	0,064	0,044	0,031	0,186	0,130	0,092	0,063	0,429	0,292	0,206	0,143
RSS C&P trim 10%	$k=2$	0,089	0,061	0,044	0,031	0,181	0,129	0,091	0,063	0,405	0,296	0,204	0,146
	$k=5$	0,089	0,063	0,045	0,032	0,186	0,131	0,093	0,066	0,426	0,292	0,201	0,144
C&P bootstrap-t	0,131	0,069	0,044	0,032	0,261	0,149	0,095	0,067	0,565	0,347	0,211	0,146	
PRSS C&P	$r=1$	0,073	0,051	0,036	0,026	0,146	0,101	0,075	0,053	0,313	0,233	0,159	0,112
	$r=2$	0,079	0,057	0,040	0,030	0,164	0,117	0,083	0,059	0,363	0,262	0,183	0,130
McKay-ev	0,113	0,069	0,047	0,032	0,237	0,146	0,096	0,067	0,742	0,346	0,220	0,152	
McKay-ev trim	0,113	0,070	0,047	0,033	0,253	0,152	0,098	0,068	0,775	0,371	0,222	0,152	
McKay-ev trim 10%	0,112	0,071	0,046	0,033	0,254	0,151	0,100	0,068	0,779	0,357	0,226	0,152	
RSS McKay-ev	$k=2$	0,115	0,071	0,046	0,032	0,254	0,146	0,099	0,069	0,732	0,357	0,238	0,150
	$k=5$	0,118	0,073	0,047	0,033	0,263	0,152	0,099	0,067	0,860	0,336	0,218	0,151
RSS McKay-ev trim	$k=2$	0,118	0,070	0,048	0,032	0,252	0,149	0,098	0,068	0,744	0,376	0,228	0,150
	$k=5$	0,121	0,072	0,048	0,032	0,251	0,151	0,100	0,067	0,815	0,360	0,217	0,154
RSS McKay-ev trim 10%	$k=2$	0,116	0,072	0,046	0,032	0,249	0,143	0,100	0,064	0,790	0,358	0,218	0,146
	$k=5$	0,116	0,071	0,049	0,032	0,266	0,151	0,098	0,069	0,902	0,354	0,225	0,151
PRSS McKay-ev	$r=1$	0,091	0,056	0,039	0,027	0,197	0,120	0,080	0,055	0,522	0,271	0,178	0,121
	$r=2$	0,105	0,065	0,043	0,030	0,222	0,136	0,090	0,061	0,650	0,310	0,202	0,135

Utvrđeno je da intervali poverenja zasnovani na odsečenoj sredini imaju približno istu širinu kao i oni koji nisu zasnovani na odsečenoj sredini.

Posmatrajući predstavljene rezultate, može se zaključiti da su za gama, Weibull-ovu, log-normalnu i *skew*-normalnu raspodelu, u većini slučajeva, proporcije simuliranih intervala za koeficijent varijacije zasnovanih na odsečenoj sredini najbliže nominalnom nivou pouzdanosti, za sve veličine uzoraka. Takođe, uočava se da intervali poverenja zasnovani na odsečenoj sredini imaju približno istu širinu kao i oni koji nisu zasnovani na odsečenoj sredini.

3.4 Analiza i primena u korporativnim finansijama

Koristeći različite promenljive u ovom delu će se vršiti analiza pouzdanosti intervala poverenja za mere disperzije. Reč je o meri sistematskog rizika i dividendama. U odeljcima 3.4.3 i 3.4.4 će biti prikazani dobijeni rezultati.

3.4.1 Mera sistematskog rizika

Mera sistematskog rizika (beta koeficijent) predmet je proučavanja različitih autora: Sharpe et al. (1999), Ross et al. (2003), Brealey et al. (2011), Bodie et al. (2009), Damodaran (2015), Besley i Brigham (2018) i drugih.

Kompanije različito reaguju na dejstvo ekonomskih (tržišnih) faktora, kao što su promena kamatne stope i potrošačkih cena. Dodavanje različitih hartija od vrednosti u dobro diversifikovan portfolio ne bi imalo isti efekat na rizik portfolija, jer različite hartije od vrednosti imaju različit stepen sistematskog rizika. Pitanje je kako izmeriti sistematski rizik pojedinačne hartije od vrednosti. Svi rizici osim onih koji se odnose na tržišna kretanja mogu se dobro diversifikovati. Rizik koji preostaje nakon diversifikacije naziva se tržišni rizik. Tržišni rizik preostaje u portfoliju koji se sastoji od sve finansijske i realne aktive koja je dostupna na tržištu. Neka se pretpostavi da postoji mogućnost za kupovinu sve finansijske aktive. Ako se raspolaže dobro diversifikovanim portfolijom, povraćaj na portfolio samo je pod uticajem tržišnih faktora koji utiču na sve kompanije (na primer, potrošačke cene). Taj portfolio nije pod uticajem onih faktora koje utiču na određene kompanije (na primer, loša marketinška kampanja). Znači, dobro diversifikovan portfolio je samo pod uticajem sistematskog, odnosno tržišnog rizika. Može se zaključiti da postoji mogućnost da se meri tržišni rizik pojedinačne hartije od vrednosti tako što se analizira njeno kretanje i kretanje na tržištu. Mera osetljivosti akcije na tržišne fluktuacije naziva se beta koeficijent (Besley & Brigham, 2018).

Sistematski rizik je pod uticajem tržišnih kretanja i utiče na sve hartije od vrednosti, a specifični faktori rizika kompanije mere specifičnosti vezane samo za konkretnu kompaniju. U svrhu procene sistematskog i specifičnog rizika određene hartije od vrednosti koriste se faktorski modeli. Za potrebe definisanja faktorskog modela, polazi se od dodatnog prinosa neke hartije od vrednosti, koji će biti označen sa R_i . Da bi se uočila razlika između tržišnih i specifičnih faktora, dodatni prinos se razlaže na sledeće komponente (Bodie et al., 2009, str. 183):

$$R_i = E(R_i) + \beta_i M + \varepsilon_i, i = 1, \dots, n, \quad (3.15)$$

gde je:

- $E(R_i)$ - očekivani dodatni prinos u periodu investiranja,
- M - pokazatelj tržišnih kretanja tokom perioda investiranja,
- β_i - pokazatelj osetljivosti hartije od vrednosti na tržišni faktor,
- ε_i - pokazatelj uticaja neočekivanih događaja na kompaniju.

Faktorski model prinosa na hartije od vrednosti se koristi samo u slučaju kada je moguće izmeriti uticaj faktora koji deluje na prinose hartija od vrednosti. Na primer, kao aproksimaciju tržišnog faktora moguće je koristiti stopu prinosa tržišnog indeksa S&P500. Model prinosa na akcije koji na osnovu tržišnog indeksa pokazuje uticaj sistematskih faktora rizika zove se indeksni model. Na osnovu ovog modela se uočava veza između prinosa na hartije od vrednosti i tržišnog indeksa. Koristeći indeksni model, dodatni prinos hartije od vrednosti se može razložiti na sistematsku i specifičnu komponentu na sledeći način (Bodie et al., 2009, str. 185):

$$R_i = \alpha_i + \beta_i R_M + \varepsilon_i, i = 1, \dots, n. \quad (3.16)$$

Jednačina (3.16) predstavlja regresionu jednačinu, gde je:

- R_i - dodatni prinos na hartiju od vrednosti (zavisna promenljiva),
- α_i - dodatni prinos na akciju kada je dodatni prinos tržišta jednak nuli (odsečak),
- β_i - mera osetljivosti prinosa hartija od vrednosti na promene u prinosu na tržištu (koeficijent nagiba),
- R_M - tržišni dodatni prinos koji zavisi od kretanja na celokupnom tržištu (objašnjavajuća promenljiva),
- ε_i - deo prinosa koji je rezultat neočekivanih događaja koji su od značaja samo za datu hartiju od vrednosti (rezidualni prinos hartije od vrednosti).

Na osnovu regresione analize, koristeći podatke o istorijskim prinosima, određuje se beta koeficijent β_i , koji pokazuje za koliko procentnih poena će se promeniti prinos hartija od vrednosti ako se tržišni prinos poveća za jedan procenat.

Iz jednačine (3.16) se uočava da postoje dva rizika hartije od vrednosti: sistematski rizik, $\beta_i R_M$, koji je posledica osetljivosti hartije od vrednosti na kretanja na ukupnom tržištu i specifični rizik, ε_i , koji predstavlja neizvesnost koja nije pod uticajem tržišnog faktora.

Pošto se nesistematski rizik može eliminisati diversifikacijom, nagrada za snošenje rizika zavisi samo od nivoa sistematskog rizika (Ross et al., 2003). Beta za prosečnu hartiju od vrednosti uzima vrednost 1. Kada je beta veća od 1, hartija od vrednosti je volatilnija od tržišnog indeksa. U slučaju kada je beta manja od 1, hartije od vrednosti predstavljaju tzv. defanzivne instrumente, jer se njihovim dodavanjem u portfolio smanjuje rizik portfolia (Sharpe et al., 1999, str. 184). Hartija od vrednosti može imati i negativnu betu, što se interpretira na sledeći način: kada stanje u privredi beleži poboljšanje očekuje se manji prinos hartije od vrednosti, i suprotno, u slučaju negativnog tržišnog dodatnog prinosa hartija od vrednosti će ostvariti veći prinos. Sledi da hartija od vrednosti sa negativnom betom obezbeđuje zaštitu od sistematskog rizika (Bodie et al., 2009).

Varijansa dodatnog prinosa akcije može se zapisati na sledeći način (Bodie et al., 2009):

$$\begin{aligned}
 \text{Var}(R_i) &= \text{Var}(\alpha_i + \beta_i R_M + \varepsilon_i) \\
 &= \text{Var}(\beta_i R_M) + \text{Var}(\varepsilon_i) \\
 &= \beta_i^2 \sigma_M^2 + \sigma^2(\varepsilon_i) \\
 &= \text{Sistematski rizik} + \text{Specifični rizik},
 \end{aligned}
 \tag{3.17}$$

za $i = 1, \dots, n$.

Iz (3.17) sledi da ukupan varijabilitet stope prinosa svake hartije od vrednosti zavisi od varijanse koja vodi poreklo od neizvesnosti koja se odnosi na celokupno tržište (sistematski rizik svake akcije zavisi od σ_M^2 i od osetljivosti akcije koja je određena sa β_i) i varijanse koja vodi poreklo od specifičnih faktora rizika (njen uticaj se meri pomoću ε_i).

Jedan od načina da se utvrdi relativna važnost sistematskog rizika je da se izračuna odnos sistematske varijanse i ukupne varijanse (Bodie et al., 2009, str. 187):

$$\rho^2 = \frac{\text{Sistematska varijansa}}{\text{Ukupna varijansa}} = \frac{\beta_i^2 \sigma_M^2}{(\beta_i^2 \sigma_M^2 + \sigma^2(\varepsilon_i))}, \quad i = 1, \dots, n.$$

Ukoliko koeficijent korelacije (ρ) između R_M i R_i iznosi 1 ili -1, prinos hartije od vrednosti se u potpunosti objašnjava tržišnim prinosom. Visoka apsolutna vrednost ovog koeficijenta znači da tržišni faktor ima značajan uticaj u objašnjavanju varijanse hartije od vrednost.

U nastavku će biti analizirani podaci o meri sistematskog rizika⁴⁷ u S&P500 kompanijama (razmatraće se 490 kompanija, jer podaci nisu dostupni za sve kompanije) na dan 5. jul 2017. godine.

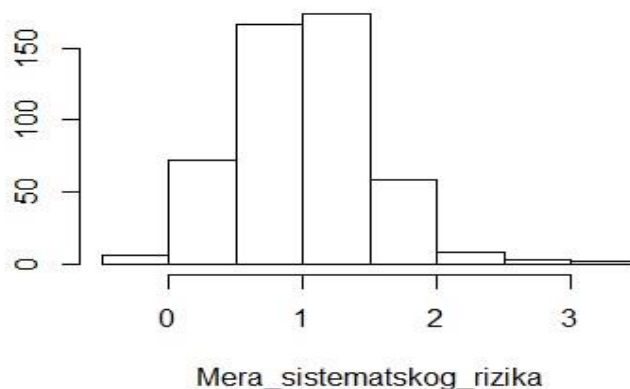
Deskriptivne mere za analizirane podatke navedene su u Tabeli 3.16.

Tabela 3.16: Deskriptivne mere za meru sistematskog rizika

Promenljiva	Broj kompanija	Sredina	Standardna devijacija	Koeficijent asimetrije
Mera sistematskog rizika	490	1,02	0,52	0,43

Izvor: Preuzeto iz Ivković i Rajić (2018)

Slikom 3.1 predstavljen je histogram raspodele mere sistematskog rizika.



Slika 3.1: Histogram raspodele mere sistematskog rizika

Izvor: Preuzeto iz Ivković i Rajić (2018)

Uočeno je da analizirana promenljiva nije normalno raspodeljena. Takođe, korišćen je Shapiro-Wilk test normalnosti da bi se ispitalo da li je mera sistematskog rizika normalno raspodeljena. Sprovedeni test pokazao je isti rezultat kao histogram (H_0 : Mera sistematskog rizika je normalno raspodeljena, a p -vrednost je približno 0).

3.4.2 Dividenda

Za kompaniju je od velike važnosti da donese odluku o tome kako raspodeliti neto dobitak na dividendu i akumulirani dobitak (Sharpe et al., 1999; Ross et al., 2003; Todorović & Ivanišević, 2018).

Svaka kompanija u određenom vremenskom periodu donosi odluku da li će vratiti gotovinu svojim akcionarima, i ako vrati koji deo će biti u obliku dividendi. Slično, vlasnik privatne

⁴⁷ Podaci su preuzeti sa sajta: <http://finance.yahoo.com/> i biće navedeni u PRILOGU 9. U daljem radu pretpostavljaće se da je reč o osnovnom skupu.

kompanije donosi odluku koliko novca povlači iz poslovanja i ponovo investira. Ove odluke nazivaju se odlukama o dividendama i pod direktnim su uticajem dividendne politike koja se primenjuje u kompaniji (Damodaran, 2015).

Izbor dividendne politike nam daje odgovor na različita pitanja: kako problem raspodele rešiti u praksi, da li dividendna politika utiče na tržišnu vrednost kompanije, da li postoji potreba za eksternim kapitalom, i druga. Postoje tri različita stanovišta o dividendnoj politici. Prema prvom stanovištu dividende nemaju važnost, jer ne utiču na vrednost kompanije. Ova tvrdnja počiva na sledeće dve pretpostavke: prva pretpostavka je da nema nepovoljnog poreskog tretmana za investitora koji prima dividendu nasuprot kapitalnom dobitku, a druga je da kompanija može da prikupi kapital na tržištu za nove investicije bez snošenja previsokih troškova emisije. Prema drugom stanovištu dividende nisu dobre za prosečnog akcionara zbog nepovoljnog poreskog tretmana, što za rezultat ima nižu vrednost kompanije. Pristalice trećeg stanovišta smatraju da su dividende dobre, jer su akcionari zainteresovani za njihovu isplatu i u skladu sa tim reaguju na povećanje dividendi (Damodaran, 2015).

Osnovne vrste novčanih dividendi su redovna, dodatna, posebna i likvidirajuća dividenda. Najčešće se isplaćuje redovna dividenda. Nekada kompanija isplaćuje i redovnu i dodatnu dividendu. Dodatna dividenda podrazumeva isplatu koja se može, a ne mora ponoviti. Posebna dividenda se uglavnom smatra jednokratnom, dok se likvidirajuća dividenda isplaćuje kada kompanija završava sa poslovanjem (Ross et al., 2003, str. 606).

Iznos isplate dividendi odvija se u skladu sa ranije usvojenom dividendnom politikom. Dividende se najčešće isplaćuju jednom, dva puta ili četiri puta godišnje. Razlikuju se četiri dividendne politike: politika rezidualne dividende, politika konstantnog racija plaćanja dividende, politika stabilnih dividendi i politika niskih regularnih plus ekstradividende (Todorović & Ivanišević, 2018, str. 216-218).

- Politika rezidualne dividende - cilj kompanije je maksimiziranje vrednosti, pa se raspoloživi kapital koristi da bi se finansirali isplativi investicioni projekti, dok se ostatak isplaćuje u obliku dividende. Kao rezultat dividende će iz godine u godinu konstantno oscilirati, što ne nailazi na odobravanje akcionara.
- Politika konstantnog racija plaćanja dividende - kompanija unapred utvrđuje procenat neto dobitka koji će se isplaćivati kao dividenda. Problem se javlja u slučaju kada je neto dobitak vrlo nizak ili kada kompanija posluje sa gubitkom. Tada će dividende biti znatno smanjenje, ili neće biti isplaćene, što dovodi do toga da investitori gube poverenje, pri čemu mogu pasti tržišne cene akcija kompanije. Stoga, ukoliko su u kompaniji izražene promene u visini ostvarenog neto dobitka, ne bi trebalo primenjivati ovu politiku.
- Politika stabilnih dividendi - menadžeri imaju za cilj da cene akcija održe stabilnim. Vođenje ove politike, koja podrazumeva visoke stabilne dividende, može dovesti do finansijskog opterećenja kompanije, posebno u situacijama kada kompanija nije u mogućnosti da generiše gotovinu u istom obimu kao ranije ili se investicione potrebe povećaju. U tom slučaju kompanija će odustati od investicionih projekata, isplatu dividendi će finansirati iz novih zaduženja ili će je čak prekinuti, što vodi do velikog pada cena akcija.
- Politika niskih regularnih plus ekstradividende - isplatom visokih regularnih dividendi povećava se finansijsko opterećenje kompanije, a investitori očekuju da će se dividende povećavati. Iz tog razloga kompanija može odabrati da vodi politiku isplate ujednačenih

dividendi, ali u znatno manjem iznosu u odnosu na politiku stabilnih dividendi. U periodu kada kompanija bude raspolagala velikim zalihama gotovine, pored regularne dividende isplatiće i dodatne, tzv. ekstradividende. Primena ove politika je veoma pogodna za kompanije kod kojih su izražene visoke fluktuacije neto novčanog toka, a istovremeno sprečava formiranje visokih očekivanja od strane akcionara.

Izbor dividendne politike zavisi od različitih faktora, od kojih se mogu izdvojiti sledeći (Todorović & Ivanišević, 2018, str. 218-219):

- Raspoloživost gotovine - iznos gotovine koji se koristi za isplatu dividendi zavisi od zaliha gotovine, gotovinskih ekvivalenata, itd.
- Raspoloživost alternativnih izvora kapitala i troškovi njegovog privlačenja - ukoliko su troškovi emisije akcija visoki, cena eksternog sopstvenog kapitala će biti veća od cene neraspoređenog dobitka, zbog čega će dividende biti niske.
- Nivo i stabilnost neto dobitaka, prošlih i očekivanih - kompanije čiji je godišnji neto dobitak visok i stabilan isplaćivaće veći procenat neto dobitka kao dividendu, dok će kompanije sa nestabilnim neto dobitkom zadržati veći deo neto dobitka da bi se održao stabilan nivo dividendi.
- Iznos neraspoređenog dobitka - u nekim zemljama je zakonom utvrđeno da iznos dividendi ne može biti veći od iznosa neraspoređenog dobitka u bilansu stanja.
- Preferencije akcionara - različiti akcionari mogu plaćati porez po različitim stopama. Akcionari koji plaćaju porez po višim poreskim stopama pokušavaju da veći deo neto dobitka akumuliraju, dok je suprotna situacija sa akcionarima koji su oslobođeni plaćanja poreza na prihode od dividende.

Pored navedenih faktora na izbor dividendne politike utiču i investicione potrebe, nivo zaduženosti, kontrola nad kompanijom, itd.

U nastavku će biti analizirani podaci o dividendama⁴⁸ u S&P500 kompanijama (razmatraće se 436 kompanija, jer podaci nisu dostupni za sve kompanije) na dan 5. jul 2017. godine.

Deskriptivne mere za analizirane podatke navedene su u Tabeli 3.17.

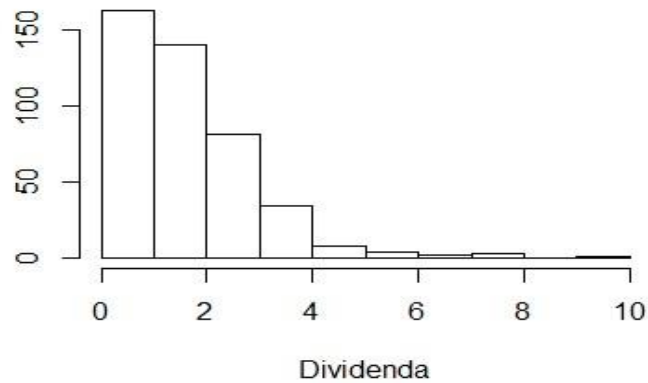
Tabela 3.17: Deskriptivne mere za dividendu

Promenljiva	Broj kompanija	Sredina	Standardna devijacija	Koeficijent asimetrije
Dividenda	436	1,63	1,34	1,92

Izvor: Izlaz iz programskog jezika R

Histogram raspodele analizirane promenljive prikazan je na Slici 3.2.

⁴⁸ Podaci su preuzeti sa sajta: <http://finance.yahoo.com/> i biće navedeni u PRILOGU 10. U daljem radu pretpostavljaće se da je reč o osnovnom skupu.



Slika 3.2: Histogram raspodele dividende

Izvor: Izlaz iz programskog jezika R

Primećuje se da dividenda nije normalno raspodeljena. U prilog tome govori i Shapiro-Wilk test normalnosti (prilikom testiranja nulte hipoteze koja glasi H_0 : Dividenda je normalno raspodeljena dobijena je p -vrednost koja je približno 0).

3.4.3 Analiza pouzdanosti intervala poverenja za varijansu

U cilju analize pouzdanosti intervala poverenja biće prikazane dobijene proporcije simuliranih intervala za varijansu zasnovane na podacima o meri sistematskog rizika i dividendama. Na kraju će biti analizirani i podaci o štetama nastalim usled požara u Danskoj (engl. Danish fire loss data).

Primer 1 Na osnovu podataka o meri sistematskog rizika⁴⁹ analizirane su proporcije simuliranih 95% dvostranih intervala za varijansu.

Iz Tabele 3.18 se uočava da su najbolji rezultati za uzorke veličine 15 dobijeni primenom RSS χ^2 intervala, dok je za uzorke veličine 20 najbolji izbor bio RSS ADF interval. Za uzorke veličine 50 primenom RSS Steve-ovog intervala dobijeni rezultat je najbliži 0,95, a za uzorke veličine 80 se kao najbolji pokazao ADF interval.

Tabela 3.18: Proporcije simuliranih 95% dvostranih intervala za varijansu mere sistematskog rizika

n	15	20	50	80
χ^2	0,934	0,918	0,921	0,911
RSS χ^2	0,945	0,924	0,943	0,921
Bonett-ov	0,962	0,972	0,976	0,971
RSS Bonett-ov	0,975	0,977	0,986	0,968
Steve-ov	0,811	0,878	0,919	0,938
RSS Steve-ov	0,833	0,893	0,955	0,931
LOG	0,839	0,873	0,904	0,913
RSS LOG	0,879	0,898	0,920	0,915
ADF	0,907	0,926	0,921	0,946
RSS ADF	0,919	0,935	0,943	0,938

Izvor: Preuzeto iz Ivković i Rajić (2018)

⁴⁹ Podaci su preuzeti sa sajta: <http://finance.yahoo.com/> i biće navedeni u PRILOGU 9.

Analiza na podacima o meri sistematskog rizika pokazala je da su, uglavnom, proporcije simuliranih intervala za varijansu zasnovanih na rangiranom uzorkovanju bile bliže nominalnom nivou pouzdanosti u odnosu na proporcije simuliranih intervala zasnovanih na prostom slučajnom uzorkovanju.

Primer 2 Na osnovu podataka o dividendama⁵⁰ analizirane su proporcije simuliranih 95% dvostranih intervala za varijansu.

Iz Tabele 3.19 se zaključuje da je za male uzorke (veličine 15 i 20), RSS Bonett-ov interval obezbedio najbolje rezultate. Kod uzoraka veličine 50, RSS Bonett-ov interval je bio najbolji izbor, dok je za uzorke veličine 80, proporcija simuliranih Bonett-ovih intervala bila najbliža nominalnom nivou pouzdanosti (i iznosi 0,915).

Tabela 3.19: Proporcije simuliranih 95% dvostranih intervala za varijansu dividende

<i>n</i>	15	20	50	80
χ^2	0,751	0,728	0,670	0,668
RSS χ^2	0,778	0,784	0,684	0,676
Bonett-ov	0,910	0,894	0,891	0,915
RSS Bonett-ov	0,926	0,922	0,901	0,900
Steve-ov	0,602	0,593	0,749	0,859
RSS Steve-ov	0,637	0,637	0,756	0,846
LOG	0,675	0,679	0,755	0,868
RSS LOG	0,720	0,751	0,780	0,816
ADF	0,761	0,754	0,824	0,867
RSS ADF	0,791	0,792	0,807	0,822

Izvor: Izlaz iz programskog jezika R

Sprovedena analiza na podacima o dividendama dovela je do zaključka da su, uglavnom, proporcije simuliranih intervala za varijansu zasnovanih na rangiranom uzorkovanju bliže nominalnom nivou pouzdanosti u odnosu na proporcije simuliranih intervala zasnovanih na prostom slučajnom uzorkovanju.

*Primer 3*⁵¹. Predmet analize su i proporcije simuliranih intervala za varijansu zasnovane na podacima koji se odnose na štete nastale usled požara u Danskoj. Ovi podaci su analizirani od strane mnogih autora i može se naći mnogo radova u kojima se proučavaju navedeni podaci. Baza podataka se sastoji od 2167 iznosa šteta u državi Danskoj i pokriva period od 03.01.1980. do 31.12.1990. godine (Pfaff et al., 2018)⁵².

Eling (2012) je u svom obimnom istraživanju pokazao da je *skew*-Student-ova raspodela pogodna za opisivanje logaritma iznosa šteta.

Deskriptivne mere analiziranih podataka date su u Tabeli 3.20.

⁵⁰ Podaci su preuzeti sa sajta: <http://finance.yahoo.com/> i biće navedeni u PRILOGU 10.

⁵¹ Primer 3 je hipotetički primer.

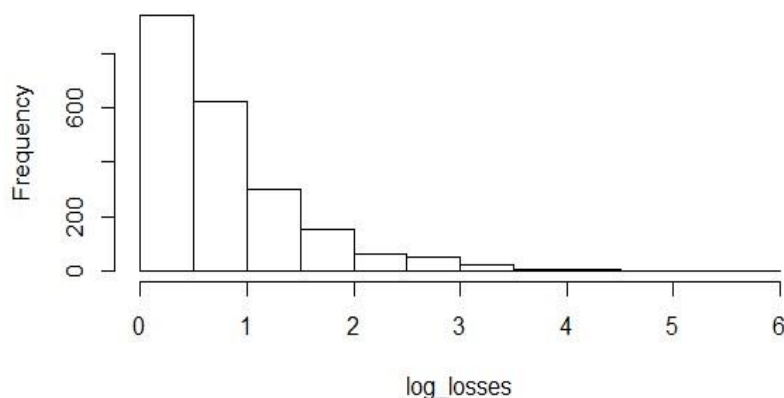
⁵² U daljem radu pretpostavljaće se da je reč o osnovnom skupu.

Tabela 3.20: Deskriptivne mere za logaritama iznosa šteta

Promenljiva	Broj šteta	Sredina	Standardna devijacija	Koeficijent asimetrije
Logaritama iznosa šteta	2167	0,79	0,72	1,76

Izvor: Preuzeto iz Rajić i Ivković (2018)

Slikom 3.3 predstavljen je histogram raspodele logaritama iznosa šteta.



Slika 3.3: Histogram raspodele logaritama iznosa šteta

Izvor: Preuzeto iz Rajić i Ivković (2018)

Proporcije simuliranih 95% dvostranih intervala za varijansu logaritama iznosa šteta su navedene u Tabeli 3.21.

Tabela 3.21: Proporcije simuliranih 95% dvostranih intervala za varijansu logaritama iznosa šteta

n	15	20	50	80
χ^2	0,774	0,781	0,790	0,755
RSS χ^2	0,808	0,805	0,804	0,788
Bonett-ov	0,893	0,902	0,932	0,944
RSS Bonett-ov	0,908	0,926	0,959	0,968
Steve-ov	0,591	0,680	0,842	0,918
RSS Steve-ov	0,624	0,717	0,881	0,926
LOG	0,722	0,761	0,864	0,905
RSS LOG	0,761	0,812	0,900	0,924
ADF	0,831	0,835	0,926	0,928
RSS ADF	0,821	0,864	0,938	0,937

Izvor: Preuzeto iz Rajić i Ivković (2018)

U slučaju malih uzoraka RSS Bonett-ov interval obezbedio je rezultate najbliže nominalnom nivou pouzdanosti. Za uzorke veličine 50 RSS Bonett-ov interval je bio najbolji izbor, dok je za uzorke veličine 80, proporcija simuliranih Bonett-ovih intervala bila najbliža nominalnom nivou pouzdanosti (iznosi 0,944).

Na osnovu dobijenih rezultata zaključuje se da su, u većini slučajeva, proporcije simuliranih intervala za varijansu zasnovanih na rangiranom uzorkovanju bile bliže nominalnom nivou pouzdanosti u odnosu na proporcije simuliranih intervala zasnovanih na prostom slučajnom uzorkovanju.

Na osnovu rezultata koji su predstavljeni u odeljcima 3.3.1, 3.3.2 i 3.4.3 potvrđuje se hipoteza

H₃: Proporcije simuliranih intervala poverenja za varijansu osnovnog skupa, konstruisanih na osnovu rangiranog uzorkovanja, bliže su nominalnom nivou pouzdanosti od proporcija intervala poverenja konstruisanih na osnovu prostog slučajnog uzorkovanja.

3.4.4 Analiza pouzdanosti intervala poverenja za koeficijent varijacije

Primer 1 Na osnovu podataka o meri sistematskog rizika⁵³ predstavljene su proporcije simuliranih intervala za koeficijent varijacije. Iz Tabele 3.22 se uočava da je za uzorke veličine 10, proporcija simuliranih RSS Miller-ovih intervala zasnovanih na odsečenoj sredini bila najbliža nominalnom nivou pouzdanosti (iznosi 0,949). Za uzorke veličine 20, RSS C&P interval zasnovan na odsečenoj sredini obezbedio je najbolji rezultat. Kada je reč o uzorcima veličine 40, najbolji rezultat ostvaren je primenom RSS Miller-ovog intervala zasnovanog na odsečenoj sredini, dok je za uzorke veličine 80 primena C&P i RSS Miller-ovog intervala koji su zasnovani na odsečenoj sredini dala najbolje rezultate.

⁵³ Podaci su preuzeti sa sajta: <http://finance.yahoo.com/> i biće navedeni u PRILOGU 9.

Tabela 3.22: Proporcije simuliranih 95% dvostranih intervala za koeficijent varijacije mere sistematskog rizika

n		10	20	40	80
t		0,892	0,910	0,902	0,924
Percentil		0,792	0,891	0,920	0,939
RSS Percentil	$k=2$	0,808	0,869	0,924	0,929
	$k=5$	0,704	0,843	0,894	0,904
Bootstrap- t		0,895	0,914	0,939	0,938
BC		0,830	0,885	0,928	0,953
RSS BC	$k=2$	0,817	0,890	0,909	0,930
	$k=5$	0,716	0,853	0,876	0,917
BCa		0,826	0,903	0,920	0,953
RSS Bca	$k=2$	0,835	0,890	0,915	0,931
	$k=5$	0,685	0,855	0,902	0,936
Jackknife		0,900	0,914	0,930	0,968
Miller-ov		0,897	0,907	0,942	0,945
Miller-ov trim		0,894	0,933	0,944	0,956
Miller-ov trim 10%		0,872	0,913	0,933	0,969
RSS Miller-ov	$k=2$	0,897	0,925	0,935	0,954
	$k=5$	0,943	0,961	0,964	0,979
RSS Miller-ov trim	$k=2$	0,907	0,918	0,950	0,954
	$k=5$	0,949	0,965	0,973	0,965
RSS Miller-ov trim 10%	$k=2$	0,917	0,932	0,942	0,949
	$k=5$	0,952	0,970	0,973	0,968
Miller-ov bootstrap- t		0,971	0,970	0,980	0,985
PRSS	$r=1$	0,743	0,655	0,536	0,373
Miller-ov	$r=2$	0,830	0,831	0,837	0,819
C&P		0,880	0,912	0,936	0,961

n		10	20	40	80
C&P trim		0,877	0,931	0,939	0,959
C&P trim 10%		0,875	0,910	0,934	0,949
RSS C&P	$k=2$	0,888	0,927	0,931	0,956
	$k=5$	0,945	0,967	0,964	0,973
RSS C&P trim	$k=2$	0,885	0,929	0,955	0,941
	$k=5$	0,946	0,954	0,968	0,969
RSS C&P trim 10%	$k=2$	0,921	0,935	0,946	0,957
	$k=5$	0,936	0,959	0,954	0,970
C&P bootstrap- t		0,965	0,977	0,983	0,986
PRSS C&P	$r=1$	0,708	0,637	0,510	0,367
	$r=2$	0,804	0,841	0,820	0,786
McKay-ev		0,880	0,965	0,968	0,971
McKay-ev trim		0,823	0,963	0,961	0,966
McKay-ev trim 10%		0,805	0,964	0,970	0,980
RSS McKay-ev	$k=2$	0,810	0,971	0,955	0,964
	$k=5$	0,854	0,985	0,991	0,984
RSS McKay-ev trim	$k=2$	0,806	0,971	0,974	0,952
	$k=5$	0,865	0,985	0,986	0,983
RSS McKay-ev trim 10%	$k=2$	0,830	0,965	0,967	0,970
	$k=5$	0,849	0,992	0,993	0,983
PRSS McKay-ev	$r=1$	0,887	0,850	0,733	0,485
	$r=2$	0,870	0,936	0,916	0,895

Izvor: Preuzeto iz Ivković i Rajić (2019)

Na podacima o meri sistematskog rizika ispitane su i prosečne širine intervala poverenja, prikazane u Tabeli 3.23. Rezultati su slični već navedenim u odeljku 3.3.3, odnosno zaključeno je da intervali poverenja zasnovani na odsečenoj sredini pokazuju približno istu širinu kao i oni koji nisu zasnovani na odsečenoj sredini.

Tabela 3.23: Prosečna širina 95% dvostranog intervala poverenja za koeficijent varijacije mere sistematskog rizika

<i>n</i>		10	20	40	80
<i>t</i>		0,510	0,335	0,231	0,161
Percentil		0,483	0,351	0,258	0,188
RSS Percentil	<i>k=2</i>	0,420	0,322	0,244	0,180
	<i>k=5</i>	0,274	0,251	0,201	0,152
<i>Bootstrap-t</i>		0,716	0,418	0,284	0,195
BC		0,497	0,360	0,264	0,191
RSS BC	<i>k=2</i>	0,418	0,327	0,243	0,179
	<i>k=5</i>	0,264	0,254	0,199	0,152
BCa		0,538	0,382	0,275	0,195
RSS Bca	<i>k=2</i>	0,451	0,342	0,252	0,184
	<i>k=5</i>	0,259	0,256	0,203	0,155
<i>Jackknife</i>		0,622	0,406	0,286	0,199
Miller-ov		0,602	0,403	0,281	0,198
Miller-ov trim		0,587	0,406	0,280	0,197
Miller-ov trim 10%		0,594	0,404	0,281	0,198
RSS Miller-ov	<i>k=2</i>	0,591	0,404	0,281	0,197
	<i>k=5</i>	0,594	0,409	0,282	0,198
RSS Miller-ov trim	<i>k=2</i>	0,598	0,410	0,281	0,196
	<i>k=5</i>	0,604	0,406	0,285	0,198
RSS Miller-ov trim 10%	<i>k=2</i>	0,598	0,404	0,282	0,197
	<i>k=5</i>	0,602	0,406	0,282	0,198
Miller-ov bootstrap-t		0,976	0,533	0,361	0,244
PRSS Miller-ov	<i>r=1</i>	0,451	0,307	0,215	0,151
	<i>r=2</i>	0,525	0,362	0,251	0,178
C&P		0,558	0,389	0,278	0,194

<i>n</i>		10	20	40	80
C&P trim		0,579	0,393	0,278	0,196
C&P trim 10%		0,563	0,395	0,276	0,196
RSS C&P	<i>k=2</i>	0,564	0,394	0,275	0,196
	<i>k=5</i>	0,580	0,393	0,278	0,195
RSS C&P trim	<i>k=2</i>	0,571	0,395	0,278	0,197
	<i>k=5</i>	0,572	0,402	0,278	0,196
RSS C&P trim 10%	<i>k=2</i>	0,565	0,392	0,280	0,195
	<i>k=5</i>	0,579	0,394	0,280	0,197
C&P bootstrap-t		0,893	0,528	0,353	0,242
PRSS C&P	<i>r=1</i>	0,434	0,302	0,214	0,151
	<i>r=2</i>	0,497	0,351	0,249	0,175
McKay-ev		1,140	0,562	0,322	0,214
McKay-ev trim		1,116	0,567	0,326	0,214
McKay-ev trim 10%		1,064	0,573	0,324	0,213
RSS McKay-ev	<i>k=2</i>	1,222	0,568	0,327	0,213
	<i>k=5</i>	1,864	0,561	0,324	0,212
RSS McKay-ev trim	<i>k=2</i>	1,297	0,564	0,325	0,213
	<i>k=5</i>	1,152	0,556	0,327	0,215
RSS McKay-ev trim 10%	<i>k=2</i>	1,192	0,551	0,326	0,213
	<i>k=5</i>	1,363	0,575	0,327	0,214
PRSS McKay-ev	<i>r=1</i>	0,796	0,393	0,238	0,161
	<i>r=2</i>	1,023	0,479	0,287	0,190

Izvor: Preuzeto iz Ivković i Rajić (2019)

Izvršena analiza na podacima o meri sistematskog rizika pokazala je da su za sve veličine uzoraka, u većini slučajeva, proporcije simuliranih intervala za koeficijent varijacije zasnovanih na odsečenoj sredini koristeći rangirano uzorkovanje bile najbliže nominalnom nivou pouzdanosti. Kada je reč o širinama intervala, zaključak je da intervali poverenja zasnovani na odsečenoj sredini pokazuju približno istu širinu kao i oni koji nisu zasnovani na odsečenoj sredini.

Primer 2 Na osnovu podataka o dividendama⁵⁴ prikazane su proporcije simuliranih intervala za koeficijent varijacije. Iz Tabele 3.24 se može primetiti da je za uzorke veličine 10, RSS Miller-ov interval zasnovan na odsečenoj sredini obezbedio najbolje rezultate, dok je za uzorke veličine 20 najbolji izbor bio RSS C&P interval zasnovan na odsečenoj sredini. Za uzorke veličine 40, najbolji rezultati dobijeni su primenom RSS Miller-ovog intervala zasnovanog na odsečenoj sredini, dok je Miller-ov interval zasnovan na odsečenoj sredini dao rezultate najbliže 0,95 za uzorke veličine 80.

⁵⁴ Podaci su preuzeti sa sajta: <http://finance.yahoo.com/> i biće navedeni u PRILOGU 10.

Tabela 3.24: Proporcije simuliranih 95% dvostranih intervala za koeficijent varijacije dividende

n		10	20	40	80
t		0,877	0,823	0,827	0,833
Percentil		0,813	0,827	0,837	0,888
RSS Percentil	$k=2$	0,764	0,771	0,800	0,841
	$k=5$	0,629	0,706	0,747	0,797
<i>Bootstrap-t</i>		0,898	0,833	0,845	0,882
BC		0,848	0,876	0,845	0,901
RSS BC	$k=2$	0,775	0,809	0,812	0,828
	$k=5$	0,634	0,696	0,746	0,828
BCa		0,852	0,852	0,876	0,896
RSS Bca	$k=2$	0,801	0,855	0,828	0,853
	$k=5$	0,639	0,741	0,757	0,826
<i>Jackknife</i>		0,468	0,430	0,471	0,461
Miller-ov		1,000	1,000	1,000	1,000
Miller-ov trim		0,904	0,909	0,918	0,950
Miller-ov trim 10%		0,895	0,918	0,929	0,955
RSS Miller-ov	$k=2$	1,000	1,000	1,000	1,000
	$k=5$	1,000	1,000	1,000	1,000
RSS Miller-ov trim	$k=2$	0,927	0,928	0,931	0,938
	$k=5$	0,950	0,959	0,951	0,946
RSS Miller-ov trim 10%	$k=2$	0,897	0,930	0,941	0,932
	$k=5$	0,955	0,952	0,956	0,953
Miller-ov <i>bootstrap-t</i>		0,982	0,979	0,966	0,968
PRSS Miller-ov	$r=1$	0,730	0,688	0,538	0,453
	$r=2$	0,810	0,829	0,809	0,781
C&P		0,889	0,892	0,920	0,938

n		10	20	40	80
C&P trim		0,888	0,906	0,938	0,941
C&P trim 10%		0,892	0,913	0,920	0,946
RSS C&P	$k=2$	0,904	0,908	0,906	0,926
	$k=5$	0,931	0,946	0,957	0,940
RSS C&P trim	$k=2$	0,914	0,925	0,935	0,932
	$k=5$	0,956	0,946	0,952	0,958
RSS C&P trim 10%	$k=2$	0,894	0,909	0,899	0,935
	$k=5$	0,951	0,950	0,958	0,957
C&P <i>bootstrap-t</i>		0,970	0,963	0,964	0,972
PRSS C&P	$r=1$	0,746	0,650	0,496	0,423
	$r=2$	0,819	0,829	0,811	0,784
McKay-ev		0,876	0,911	0,922	0,941
McKay-ev trim		0,314	0,758	0,982	0,976
McKay-ev trim 10%		0,293	0,767	0,976	0,979
RSS McKay-ev	$k=2$	0,892	0,909	0,914	0,941
	$k=5$	0,936	0,945	0,947	0,960
RSS McKay-ev trim	$k=2$	0,304	0,780	0,981	0,966
	$k=5$	0,239	0,789	0,993	0,987
RSS McKay-ev trim 10%	$k=2$	0,259	0,771	0,980	0,974
	$k=5$	0,241	0,777	0,997	0,991
PRSS McKay-ev	$r=1$	0,537	0,852	0,810	0,643
	$r=2$	0,426	0,799	0,938	0,921

Izvor: Izlaz iz programskog jezika R

Na osnovu sprovedene analize na podacima o dividendama zaključuje se da su za sve veličine uzoraka, u većini slučajeva, proporcije simuliranih intervala za koeficijent varijacije zasnovanih na odsečenoj sredini koristeći rangirano uzorkovanje najbliže nominalnom nivou pouzdanosti.

Proučavajući rezultate predstavljene u odeljcima 3.3.3 i 3.4.4 može se potvrditi hipoteza

H_4 : Proporcije simuliranih intervala poverenja za koeficijent varijacije osnovnog skupa, zasnovanih na transformaciji sa odsečenom sredinom, bliže su nominalnom nivou pouzdanosti u odnosu na proporcije ostalih ispitivanih intervala.

Na osnovu rezultata prikazanih u odeljcima 2.5.4, 2.5.5, 3.4.3 i 3.4.4 potvrđuje se hipoteza

H_5 : Analiza pouzdanosti na realnim podacima, u oblasti korporativnih finansija, pokazuje da su proporcije predloženih intervala poverenja bliže nominalnom nivou pouzdanosti u odnosu na proporcije intervala koji su do sada korišćeni u literaturi.

3.5 R kodovi za izabrane intervale

U odeljcima 3.5.1 i 3.5.2 predstavljeni su originalni kodovi iz programskog jezika *R* za izabrane intervale poverenja za varijansu i koeficijent varijacije.

3.5.1 R kodovi za intervale poverenja za varijansu

U nastavku slede originalni kodovi za χ^2 , Bonett-ov interval, interval poverenja sa korigovanim stepenima slobode, RSS Bonett-ov interval, Steve-ov interval poverenja zasnovan na normalnoj aproksimaciji u slučaju velikih uzoraka i interval poverenja zasnovan na logaritamskoj transformaciji za varijansu osnovnog skupa.

Tabela 3.25: Kodovi za χ^2 interval i Bonett-ov interval za varijansu

χ^2 interval poverenja	Bonett-ov interval poverenja
<pre>x<-rsn(2000,3,2,5) brsim<-1000 n<-80 s2<-rep(NA,brsim) l95<-rep(NA,brsim) u95<-rep(NA,brsim) for (i in 1:brsim){ uzorak<-sample(1:2000,n) x1<-x[uzorak] suma_x1=sum(x1) x1kv=x1^2 suma_x1kv=sum(x1kv) s2[i]=(suma_x1kv-((suma_x1)^2)/n)/(n-1) } l95<-((n-1)*s2)/qchisq(0.975,n-1) u95<-((n-1)*s2)/qchisq(0.025,n-1) mean((l95<=1.55)&(1.55<=u95))</pre>	<pre>x<-rsn(2000,3,2,5) brsim<-1000 n<-80 trimp=1/(2*((n-4)^0.5)) c=n/(n-qnorm(0.975)) s2<-rep(NA,brsim) se<-rep(NA,brsim) l95<-rep(NA,brsim) u95<-rep(NA,brsim) for (i in 1:brsim){ uzorak<-sample(1:2000,n) x1<-x[uzorak] suma_x1=sum(x1) x1kv=x1^2 suma_x1kv=sum(x1kv) s2[i]=(suma_x1kv-((suma_x1)^2)/n)/(n-1) sred=mean(x1) m=mean(x1,trim=trimp) g4_bar=(n*sum((x1-m)^4))/((sum((x1-sred)^2))^2) se[i]=c*((g4_bar*(n-3)/n)/(n-1))^0.5 } l95=exp(log(c*s2)-qnorm(0.975)*se) u95=exp(log(c*s2)+qnorm(0.975)*se) mean((l95<=1.55)&(1.55<=u95))</pre>

Tabela 3.26: Kodovi za interval poverenja sa korigovanim stepenima slobode i RSS Bonett-ov interval za varijansu

Interval poverenja sa korigovanim stepenima slobode	RSS Bonett-ov interval poverenja
<pre> x<-rsn(2000,3,2,5) brsim<-1000 n<-80 s2<-rep(NA,brsim) s<-rep(NA,brsim) ge_kapa<-rep(NA,brsim) r_kapa<-rep(NA,brsim) l95<-rep(NA,brsim) u95<-rep(NA,brsim) for (i in 1:brsim){ uzorak<-sample(1:2000,n) x1<-x[uzorak] suma_x1=sum(x1) x1kv=x1^2 suma_x1kv=sum(x1kv) s2[i]=(suma_x1kv-((suma_x1)^2)/n)/(n-1) s[i]=sqrt(s2[i]) sred=mean(x1) ge_kapa[i]=(((n*(n+1))/((n-1)*(n-2)*(n-3)))*((sum((x1- sred)^4))/(s[i]^4)))-((3*(n-1)^2)/((n-2)*(n-3))) r_kapa[i]=(2*n)/(ge_kapa[i]+((2*n)/(n-1))) } l95<-(r_kapa*s2)/qchisq(0.975,r_kapa) u95<-(r_kapa*s2)/qchisq(0.025,r_kapa) l95[is.na(l95)] <- 0 u95[is.na(u95)] <- 0 mean((l95<=1.55)&(1.55<=u95)) </pre>	<pre> x<-rsn(2000,3,2,5) brsim<-1000 k=4 m=20 n=m*k trimp=1/(2*((n-4)^0.5)) c=n/(n-qnorm(0.975)) rss<-rep(NA,k) rss_uzorak<-rep(NA,m) rss_all<-rep(NA,brsim) sred_rss<-rep(NA,brsim) s2_rss<-rep(NA,brsim) se<-rep(NA,brsim) l95<-rep(NA,brsim) u95<-rep(NA,brsim) for (b in 1:brsim){ List = list() for (j in 1:m){ uzorak=sample(x, k^2) for (i in 1:k){ pod_uzorak=uzorak[((i-1)*k)+1):(i*k)] pod_uzorak_sort=sort(pod_uzorak) rss[i]=pod_uzorak_sort[i] } } rss_uzorak=c(rss) List[[length(List)+1]] = rss_uzorak } rss_all=c(do.call("cbind",List)) sred_rss=(1/(k*m))*sum(rss_all) tm=mean(rss_all,trim=trimp) g4_bar=(n*sum((rss_all-tm)^4))/((sum((rss_all- sred_rss)^2))^2) s2_rss[b]=((1/((k*m)-1))*sum((rss_all-sred_rss)^2)) se[b]=c*(((g4_bar*(n-3)/n)/(n-1))^0.5) } l95=exp(log(c*s2_rss)-qnorm(0.975)*se) u95=exp(log(c*s2_rss)+qnorm(0.975)*se) mean((l95<=1.55)&(1.55<=u95)) </pre>

Tabela 3.27: Kodovi za Steve-ov interval zasnovan na normalnoj aproksimaciji u slučaju velikih uzoraka i interval zasnovan na logaritamskoj transformaciji za varijansu

Steve-ov interval poverenja zasnovan na normalnoj aproksimaciji u slučaju velikih uzoraka	Interval poverenja zasnovan na logaritamskoj transformaciji
<pre> x<-rsn(2000,3,2,5) brsim<-1000 n<-80 s2<-rep(NA,brsim) g_kapa<-rep(NA,brsim) l95<-rep(NA,brsim) u95<-rep(NA,brsim) for (i in 1:brsim){ uzorak<-sample(1:2000,n) x1<-x[uzorak] suma_x1=sum(x1) x1kv=x1^2 suma_x1kv=sum(x1kv) s2[i]=(suma_x1kv-((suma_x1)^2)/n)/(n-1) sred=mean(x1) g_kapa[i]=(n*sum((x1-sred)^4))/((sum((x1-sred)^2))^2) } l95=(s2/(1-(qnorm(0.025)*sqrt((g_kapa-1)/n)))) u95=(s2/(1+(qnorm(0.025)*sqrt((g_kapa-1)/n)))) l95[is.na(l95)] <- 0 u95[is.na(u95)] <- 0 mean((l95<=1.55)&(1.55<=u95)) </pre>	<pre> x<-rsn(2000,3,2,5) brsim<-1000 n<-80 s2<-rep(NA,brsim) g_kapa<-rep(NA,brsim) l95<-rep(NA,brsim) u95<-rep(NA,brsim) for (i in 1:brsim){ uzorak<-sample(1:2000,n) x1<-x[uzorak] suma_x1=sum(x1) x1kv=x1^2 suma_x1kv=sum(x1kv) s2[i]=(suma_x1kv-((suma_x1)^2)/n)/(n-1) sred=mean(x1) g_kapa[i]=(n*sum((x1-sred)^4))/((sum((x1-sred)^2))^2) } l95=(s2*exp(qnorm(0.025)*sqrt((g_kapa-1)/n))) u95=(s2*exp(-qnorm(0.025)*sqrt((g_kapa-1)/n))) #l95[is.na(l95)] <- 0 #u95[is.na(u95)] <- 0 mean((l95<=1.55)&(1.55<=u95)) </pre>

3.5.2 R kodovi za intervale poverenja za koeficijent varijacije

U nastavku slede originalni kodovi⁵⁵ za sledeće intervale za koeficijent varijacije osnovnog skupa: *t*, percentil, *jackknife*, *bootstrap-t*, BC, BCa, RSS BCa, RSS percentil, Miller-ov, Miller-ov zasnovan na Bonett-ovoj odsečenoj sredini, Miller-ov zasnovan na odsečenoj sredini koja eliminiše 10% vrednosti, Miller-ov *bootstrap-t*, RSS Miller-ov, RSS Miller-ov zasnovan na Bonett-ovoj odsečenoj sredini, RSS Miller-ov zasnovan na odsečenoj sredini koja eliminiše 10% vrednosti, PRSS Miller-ov, Curto-ov & Pinto-ov i McKay-ev.

Tabela 3.28: Kodovi za *t* interval i percentil interval za koeficijent varijacije

Interval poverenja zasnovan na <i>t</i> statistici	Percentil interval poverenja
<pre>x<-rnorm(20,5,2) brsim<-4 n<-6 cv<-rep(NA,brsim) scv<-rep(NA,brsim) l95<-rep(NA,brsim) u95<-rep(NA,brsim) for (i in 1:brsim){ uzorak<-sample(1:20,n) x1<-x[uzorak] suma_x1=sum(x1) x1kv=x1^2 suma_x1kv=sum(x1kv) s=sqrt((suma_x1kv-((suma_x1)^2)/n)/(n-1)) m=mean(x1) cv[i]<-s/m scv[i]<-cv[i]/sqrt(2*n) } l95<-cv-qt(0.975,n-1)*scv u95<-cv+qt(0.975,n-1)*scv mean((l95<=0.4)&(0.4<=u95)) sirina=u95-l95 mean(sirina)</pre>	<pre>x<-rnorm(20,5,2) brsim<-5 B<-5 n<-5 cv<-rep(NA,brsim) boot_cv<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) for(j in 1:brsim){ uzorak1<-sample(1:20,n) x1<-x[uzorak1] suma_x1=sum(x1) x1kv=x1^2 suma_x1kv=sum(x1kv) s=sqrt((suma_x1kv-((suma_x1)^2)/n)/(n-1)) m=mean(x1) cv[j]<-s/m for (i in 1:B){ uzorak2<-sample(uzorak1,replace=T) x1b<-x[uzorak2] suma_x1b=sum(x1b) x1bkv=x1b^2 suma_x1bkv=sum(x1bkv) boot_s=sqrt((suma_x1bkv-((suma_x1b)^2)/n)/(n-1)) boot_m=mean(x1b) boot_cv[j,i]<-boot_s/boot_m } } t.lowcv=apply(boot_cv, 1, quantile, probs = c(0.025), na.rm = TRUE) t.highcv=apply(boot_cv, 1, quantile, probs = c(0.975), na.rm = TRUE) l95<-t.lowcv u95<-t.highcv mean((l95<=0.4)&(0.4<=u95)) sirina=u95-l95 mean(sirina)</pre>

⁵⁵ Kodovi predstavljeni u Tabelama 3.28-3.36 preuzeti su iz rada Ivković i Rajić (2019).

Tabela 3.29: Kodovi za *jackknife* interval i *bootstrap-t* interval za koeficijent varijacije

<i>Jackknife</i> interval poverenja	<i>Bootstrap-t</i> interval poverenja
<pre> x<-rnorm(20,5,2) brsim<-5 n<-5 cv<-rep(NA,brsim) partial_cv<-matrix(rep(NA,brsim*n),nrow=brsim,ncol=n) pseudo_cv<-matrix(rep(NA,brsim*n),nrow=brsim,ncol=n) for(j in 1:brsim){ uzorak1 <- sample(1:20,n) x1<-x[uzorak1] suma_x1=sum(x1) x1kv=x1^2 suma_x1kv=sum(x1kv) s=sqrt((suma_x1kv-((suma_x1)^2)/n)/(n-1)) m=mean(x1) cv[j]<-s/m for(i in 1:n){ uzorak_jack<-sample(uzorak1[-i],n-1) x_jack<-x[uzorak_jack] suma_x_jack=sum(x_jack) x_jackkv=x_jack^2 suma_x_jackkv=sum(x_jackkv) s_jack=sqrt((suma_x_jackkv-((suma_x_jack)^2)/(n-1)))/(n-2)) m_jack=mean(x_jack) partial_cv[j,i]<- s_jack/m_jack pseudo_cv[j,i]<-n*cv[j]-(n-1)*partial_cv[j,i] } } jack_cv<-apply(pseudo_cv,1,mean) jack_se_cv<-sqrt((apply(((pseudo_cv-jack_cv)^2),1,sum)))/(n*(n-1))) l95_cv<-jack_cv-qt(0.975,n-1)*jack_se_cv u95_cv<-jack_cv+qt(0.975,n-1)*jack_se_cv mean((l95_cv<=0.4)&(0.4<=u95_cv)) sirina=u95_cv-l95_cv mean(sirina) </pre>	<pre> x<-rnorm(20,5,2) brsim<-5 B<-5 n<-5 cv<-rep(NA,brsim) scv<-rep(NA,brsim) boot_cv<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) boot_scv<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) for(j in 1:brsim){ uzorak1 <- sample(1:20,n) x1<-x[uzorak1] suma_x1=sum(x1) x1kv=x1^2 suma_x1kv=sum(x1kv) s=sqrt((suma_x1kv-((suma_x1)^2)/n)/(n-1)) m=mean(x1) cv[j]<-s/m scv[j]<-cv[j]/sqrt(2*n) for (i in 1:B){ uzorak2 <-sample(uzorak1,replace=T) x1b <- x[uzorak2] suma_x1b=sum(x1b) x1bkv=x1b^2 suma_x1bkv=sum(x1bkv) boot_s=sqrt((suma_x1bkv-((suma_x1b)^2)/n)/(n-1)) boot_m=mean(x1b) boot_cv[j,i] <-boot_s/boot_m boot_scv[j,i] <-boot_cv[j,i]/sqrt(2*n) } } boot_tcv=(boot_cv-cv)/boot_scv t.lowcv=apply(boot_tcv, 1, quantile, probs = c(0.025), na.rm = TRUE) t.highcv=apply(boot_tcv, 1, quantile, probs = c(0.975), na.rm = TRUE) l95<-cv-t.highcv*scv u95<-cv-t.lowcv*scv mean((l95<=0.4)&(0.4<=u95)) sirina=u95-l95 mean(sirina) </pre>

Tabela 3.30: Kodovi za BC interval i BCa interval za koeficijent varijacije

BC interval poverenja	BCa interval poverenja
<pre> x<-rnorm(20,5,2) brsim<-5 B<-8 n<-5 alfa=0.05 cv<-rep(NA,brsim) boot_cv<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) for(j in 1:brsim){ uzorak1 <- sample(1:20,n) x1<-x[uzorak1] suma_x1=sum(x1) x1kv=x1^2 suma_x1kv=sum(x1kv) s=sqrt((suma_x1kv-((suma_x1)^2)/n)/(n-1)) m=mean(x1) cv[j]<-s/m for (i in 1:B){ uzorak2 <-sample(uzorak1,replace=T) x1b <- x[uzorak2] suma_x1b=sum(x1b) x1bkv=x1b^2 suma_x1bkv=sum(x1bkv) boot_s=sqrt((suma_x1bkv-((suma_x1b)^2)/n)/(n-1)) boot_m=mean(x1b) boot_cv[j,i] <-boot_s/boot_m } } br_manji_cv=rowSums(boot_cv<cv, na.rm=TRUE) cumbr_cv=br_manji_cv/B z0_kapa_cv=qnorm(cumbr_cv) alfa1_cv=pnorm(2*z0_kapa_cv+qnorm(alfa/2)) alfa2_cv=pnorm(2*z0_kapa_cv+qnorm(1-alfa/2)) low_cv=apply(boot_cv, 1, quantile, probs = c(alfa1_cv), na.rm = TRUE) lcv=diag(low_cv) high_cv=apply(boot_cv, 1, quantile, probs = c(alfa2_cv), na.rm = TRUE) hcv=diag(high_cv) mean((lcv<=0.4)&(0.4<=hcv)) sirina=hcv-lcv mean(sirina) </pre>	<pre> x<-rnorm(20,5,2) brsim<-5 B<-10 n<-5 alfa=0.05 cv<-rep(NA,brsim) cv_jack<-matrix(rep(NA,brsim*n),nrow=brsim,ncol=n) cv_tacka<-rep(NA,brsim) boot_cv<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) for(j in 1:brsim){ uzorak1 <- sample(1:20,n) x1<-x[uzorak1] suma_x1=sum(x1) x1kv=x1^2 suma_x1kv=sum(x1kv) s=sqrt((suma_x1kv-((suma_x1)^2)/n)/(n-1)) m=mean(x1) cv[j]<-s/m for(k in 1:n){ uzorak_jack<-sample(uzorak1[-k],n-1) x_jack<-x[uzorak_jack] suma_x_jack=sum(x_jack) x_jackkv=x_jack^2 suma_x_jackkv=sum(x_jackkv) s_jack=sqrt((suma_x_jackkv-((suma_x_jack)^2)/(n-1)))/(n-2)) m_jack=mean(x_jack) cv_jack[j,k]<-s_jack/m_jack } } cv_tacka=apply(cv_jack,1,mean) for (i in 1:B){ uzorak2 <-sample(uzorak1,replace=T) x1b <- x[uzorak2] suma_x1b=sum(x1b) x1bkv=x1b^2 suma_x1bkv=sum(x1bkv) boot_s=sqrt((suma_x1bkv-((suma_x1b)^2)/n)/(n-1)) boot_m=mean(x1b) boot_cv[j,i] <-boot_s/boot_m } } a_kapa_cv=(apply(((cv_tacka- cv_jack)^3),1,sum))/(6*(apply(((cv_tacka- cv_jack)^2),1,sum)^(3/2))) br_manji_cv=rowSums(boot_cv<cv, na.rm=TRUE) cumbr_cv=br_manji_cv/B z0_kapa_cv=qnorm(cumbr_cv) alfa1_cv=pnorm(z0_kapa_cv+((z0_kapa_cv+qnorm(alfa/2))/(1- a_kapa_cv*(z0_kapa_cv+qnorm(alfa/2)))))) alfa2_cv=pnorm(z0_kapa_cv+((z0_kapa_cv+qnorm(1- alfa/2))/(1-a_kapa_cv*(z0_kapa_cv+qnorm(1-alfa/2)))))) low_cv=apply(boot_cv, 1, quantile, probs = c(alfa1_cv), na.rm = TRUE) lcv=diag(low_cv) high_cv=apply(boot_cv, 1, quantile, probs = c(alfa2_cv), na.rm = TRUE) hcv=diag(high_cv) mean((lcv<=0.4)&(0.4<=hcv)) sirina=hcv-lcv mean(sirina) </pre>

Tabela 3.31: Kodovi za RSS BCa interval i RSS percentil interval za koeficijent varijacije

RSS BCa interval poverenja	RSS percentil interval poverenja
<pre> x<-rnorm(20,5,2) brsim<-3 B<-8 k=3 m=3 n=m*k alfa=0.05 rss<-rep(NA,k) rss_uzorak<-rep(NA,m) rss_all<-rep(NA,brsim) boot_uzorak_matrix<-matrix(rep(NA,k*m),nrow=k,ncol=m) boot_rss_cv<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) sred_rss<-rep(NA,brsim) s_rss<-rep(NA,brsim) cv_rss<-rep(NA,brsim) cv_jack<-matrix(rep(NA,brsim*n),nrow=brsim,ncol=n) cv_tacka<-rep(NA,brsim) for (b in 1:brsim){ List = list() for (j in 1:m){ uzorak=sample(x, k^2) for (i in 1:k){ pod_uzorak=uzorak[((i-1)*k)+1):(i*k)] pod_uzorak_sort=sort(pod_uzorak) rss[i]=pod_uzorak_sort[i] } rss_uzorak=c(rss) List[[length(List)+1]] = rss_uzorak } rss_all=c(do.call("cbind",List)) sred_rss=(1/(k*m))*sum(rss_all) s_rss=sqrt(((1/((k*m)-1))*sum((rss_all-sred_rss)^2))) cv_rss[b]<-s_rss/sred_rss for(r in 1:length(rss_all)){ x_jack<-sample(rss_all[-r],length(rss_all)-1) suma_x_jack=sum(x_jack) x_jackkv=x_jack^2 suma_x_jackkv=sum(x_jackkv) s_jack=sqrt((suma_x_jackkv-((suma_x_jack^2)/(length(rss_all)-1)))/(length(rss_all)-2)) m_jack=mean(x_jack) cv_jack[b,r]<-s_jack/m_jack } cv_tacka=apply(cv_jack,1,mean) datalist=list() for (i in 1:k){ rss_all_k=rss_all[seq(i,length(rss_all),k)] datalist[[i]]<-rss_all_k } rss_all_k_matrix=do.call(rbind,datalist) for (s in 1:B){ boot_uzorak_matrix<- t(apply(rss_all_k_matrix,1,sample,replace=T)) boot_rss=as.vector(t(boot_uzorak_matrix)) suma_boot_rss=sum(boot_rss) boot_rsskv=boot_rss^2 suma_boot_rsskv=sum(boot_rsskv) boot_rss_s=sqrt((suma_boot_rsskv-((suma_boot_rss)^2)/n)/(n-1)) boot_rss_m=mean(boot_rss) boot_rss_cv[b,s] <-boot_rss_s/boot_rss_m } } lcv=apply(boot_rss_cv, 1, quantile, probs = c(0.025), na.rm = TRUE) hcv=apply(boot_rss_cv, 1, quantile, probs = c(0.975), na.rm = TRUE) mean((lcv<=0.4)&(0.4<=hcv)) sirina=hcv-lcv mean(sirina) </pre>	<pre> x<-rnorm(20,5,2) brsim<-3 B<-8 k=3 m=3 n=m*k rss<-rep(NA,k) rss_uzorak<-rep(NA,m) rss_all<-rep(NA,brsim) boot_uzorak_matrix<- matrix(rep(NA,k*m),nrow=k,ncol=m) boot_rss_cv<- matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) sred_rss<-rep(NA,brsim) s_rss<-rep(NA,brsim) cv_rss<-rep(NA,brsim) for (b in 1:brsim){ List = list() for (j in 1:m){ uzorak=sample(x, k^2) for (i in 1:k){ pod_uzorak=uzorak[((i-1)*k)+1):(i*k)] pod_uzorak_sort=sort(pod_uzorak) rss[i]=pod_uzorak_sort[i] } rss_uzorak=c(rss) List[[length(List)+1]] = rss_uzorak } rss_all=c(do.call("cbind",List)) sred_rss=(1/(k*m))*sum(rss_all) s_rss=sqrt(((1/((k*m)-1))*sum((rss_all-sred_rss)^2))) cv_rss[b]<-s_rss/sred_rss datalist=list() for (i in 1:k){ rss_all_k=rss_all[seq(i,length(rss_all),k)] datalist[[i]]<-rss_all_k } rss_all_k_matrix=do.call(rbind,datalist) for (s in 1:B){ boot_uzorak_matrix<- t(apply(rss_all_k_matrix,1,sample,replace=T)) boot_rss=as.vector(t(boot_uzorak_matrix)) suma_boot_rss=sum(boot_rss) boot_rsskv=boot_rss^2 suma_boot_rsskv=sum(boot_rsskv) boot_rss_s=sqrt((suma_boot_rsskv-((suma_boot_rss)^2)/n)/(n-1)) boot_rss_m=mean(boot_rss) boot_rss_cv[b,s] <-boot_rss_s/boot_rss_m } } lcv=apply(boot_rss_cv, 1, quantile, probs = c(0.025), na.rm = TRUE) hcv=apply(boot_rss_cv, 1, quantile, probs = c(0.975), na.rm = TRUE) mean((lcv<=0.4)&(0.4<=hcv)) sirina=hcv-lcv mean(sirina) </pre>

Tabela 3.32: Kodovi za Miller-ov interval i Miller-ov interval zasnovan na Bonett-ovoj odsečenoj sredini za koeficijent varijacije

Miller-ov interval poverenja	Miller-ov interval poverenja zasnovan na Bonett-ovoj odsečenoj sredini
<pre>x<-rnorm(20,5,2) brsim<-10 n<-5 cv<-rep(NA,brsim) l95<-rep(NA,brsim) u95<-rep(NA,brsim) for (i in 1:brsim){ uzorak<-sample(1:20,n) x1<-x[uzorak] suma_x1=sum(x1) x1kv=x1^2 suma_x1kv=sum(x1kv) s=sqrt((suma_x1kv-((suma_x1)^2)/n)/(n-1)) m=mean(x1) cv[i]<-s/m } l95<-cv-qnorm(0.975)*sqrt(((cv^2)/(n-1))*(0.5+(cv^2))) u95<-cv+qnorm(0.975)*sqrt(((cv^2)/(n-1))*(0.5+(cv^2))) mean((l95<=0.4)&(0.4<=u95)) sirina=u95-l95 mean(sirina)</pre>	<pre>x<-rnorm(20,5,2) brsim<-5 n<-10 trimp=1/(2*((n-4)^0.5)) cv<-rep(NA,brsim) l95<-rep(NA,brsim) u95<-rep(NA,brsim) for (i in 1:brsim){ uzorak<-sample(1:20,n) x1<-x[uzorak] trm=mean(x1,trim=trimp) s=sqrt((sum((x1-trm)^2)/(n-1))) m=mean(x1) cv[i]<-s/m } l95<-cv-qnorm(0.975)*sqrt(((cv^2)/(n-1))*(0.5+(cv^2))) u95<-cv+qnorm(0.975)*sqrt(((cv^2)/(n-1))*(0.5+(cv^2))) mean((l95<=0.4)&(0.4<=u95)) sirina=u95-l95 mean(sirina)</pre>

Tabela 3.33: Kodovi za Miller-ov interval zasnovan na odsečenoj sredini koja eliminiše 10% vrednosti i Miller-ov *bootstrap-t* interval za koeficijent varijacije

Miller-ov interval poverenja zasnovan na odsečenoj sredini koja eliminiše 10% vrednosti	Miller-ov <i>bootstrap-t</i> interval poverenja
<pre>x<-rnorm(20,5,2) brsim<-5 n<-10 cv<-rep(NA,brsim) l95<-rep(NA,brsim) u95<-rep(NA,brsim) for (i in 1:brsim){ uzorak<-sample(1:20,n) x1<-x[uzorak] trm=mean(x1,trim=0.1) s=sqrt((sum((x1-trm)^2)/(n-1))) m=mean(x1) cv[i]<-s/m } l95<-cv-qnorm(0.975)*sqrt(((cv^2)/(n-1))*(0.5+(cv^2))) u95<-cv+qnorm(0.975)*sqrt(((cv^2)/(n-1))*(0.5+(cv^2))) mean((l95<=0.4)&(0.4<=u95)) sirina=u95-l95 mean(sirina)</pre>	<pre>x<-rnorm(20,5,2) brsim<-10 B<-5 n<-5 cv<-rep(NA,brsim) scv<-rep(NA,brsim) boot_cv<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) boot_scv<-matrix(rep(NA,brsim*B),nrow=brsim,ncol=B) for(j in 1:brsim){ uzorak1 <- sample(1:20,n) x1<-x[uzorak1] suma_x1=sum(x1) x1kv=x1^2 suma_x1kv=sum(x1kv) s=sqrt((suma_x1kv-((suma_x1)^2)/n)/(n-1)) m=mean(x1) cv[j]<-s/m scv[j]<-cv[j]/sqrt(2*n) for (i in 1:B){ uzorak2 <- sample(uzorak1,replace=T) x1b <- x[uzorak2] suma_x1b=sum(x1b) x1bkv=x1b^2 suma_x1bkv=sum(x1bkv) boot_s=sqrt((suma_x1bkv-((suma_x1b)^2)/n)/(n-1)) boot_m=mean(x1b) boot_cv[j,i] <-boot_s/boot_m boot_scv[j,i] <-boot_cv[j,i]/sqrt(2*n) } boot_tcv=(boot_cv-cv)/boot_scv t.lowcv=apply(boot_tcv, 1, quantile, probs = c(0.025), na.rm = TRUE) t.highcv=apply(boot_tcv, 1, quantile, probs = c(0.975), na.rm = TRUE) l95<-cv-t.highcv*sqrt(((cv^2)/(n-1))*(0.5+(cv^2))) u95<-cv-t.lowcv*sqrt(((cv^2)/(n-1))*(0.5+(cv^2))) mean((l95<=0.4)&(0.4<=u95)) sirina=u95-l95 mean(sirina)</pre>

Tabela 3.34: Kodovi za RSS Miller-ov interval i RSS Miller-ov interval zasnovan na Bonett-ovoj odsečenoj sredini za koeficijent varijacije

RSS Miller-ov interval poverenja	RSS Miller-ov interval poverenja zasnovan na Bonett-ovoj odsečenoj sredini
<pre> x<-rnorm(20,5,2) brsim<-5 k=3 m=4 n=m*k rss<-rep(NA,k) rss_uzorak<-rep(NA,m) rss_all<-rep(NA,brsim) sred_rss<-rep(NA,brsim) s_rss<-rep(NA,brsim) cv<-rep(NA,brsim) I95<-rep(NA,brsim) u95<-rep(NA,brsim) for (b in 1:brsim){ List = list() for (j in 1:m){ uzorak=sample(x, k^2) for (i in 1:k){ pod_uzorak=uzorak[((i-1)*k)+1:(i*k)] pod_uzorak_sort=sort(pod_uzorak) rss[i]=pod_uzorak_sort[i] } rss_uzorak=c(rss) List[[length(List)+1]] = rss_uzorak } rss_all=c(do.call("cbind",List)) sred_rss=(1/(k*m))*sum(rss_all) s_rss=sqrt(((1/(k*m)-1))*sum((rss_all-sred_rss)^2))) cv[b]<-s_rss/sred_rss } I95<-cv-qnorm(0.975)*sqrt(((cv^2)/(n-1))*(0.5+(cv^2))) u95<-cv+qnorm(0.975)*sqrt(((cv^2)/(n-1))*(0.5+(cv^2))) mean((I95<=0.4)&(0.4<=u95)) sirina=u95-I95 mean(sirina) </pre>	<pre> x<-rnorm(20,5,2) brsim<-5 k=2 m=5 n=m*k trimp=1/(2*((n-4)^0.5)) rss<-rep(NA,k) rss_uzorak<-rep(NA,m) rss_all<-rep(NA,brsim) sred_rss<-rep(NA,brsim) s_rss<-rep(NA,brsim) cv<-rep(NA,brsim) I95<-rep(NA,brsim) u95<-rep(NA,brsim) for (b in 1:brsim){ List = list() for (j in 1:m){ uzorak=sample(x, k^2) for (i in 1:k){ pod_uzorak=uzorak[((i-1)*k)+1:(i*k)] pod_uzorak_sort=sort(pod_uzorak) rss[i]=pod_uzorak_sort[i] } rss_uzorak=c(rss) List[[length(List)+1]] = rss_uzorak } rss_all=c(do.call("cbind",List)) trm=mean(rss_all,trim=trimp) s_rss=sqrt((sum((rss_all-trm)^2))/(n-1)) sred_rss=(1/(k*m))*sum(rss_all) cv[b]<-s_rss/sred_rss } I95<-cv-qnorm(0.975)*sqrt(((cv^2)/(n-1))*(0.5+(cv^2))) u95<-cv+qnorm(0.975)*sqrt(((cv^2)/(n-1))*(0.5+(cv^2))) mean((I95<=0.4)&(0.4<=u95)) sirina=u95-I95 mean(sirina) </pre>

Tabela 3.35: Kodovi za RSS Miller-ov interval zasnovan na odsečenoj sredini koja eliminiše 10% vrednosti i PRSS Miller-ov interval za koeficijent varijacije

RSS Miller-ov interval poverenja zasnovan na odsečenoj sredini koja eliminiše 10% vrednosti	PRSS Miller-ov interval poverenja
<pre> x<-rnorm(20,5,2) brsim<-5 k=2 m=5 n=m*k rss<-rep(NA,k) rss_uzorak<-rep(NA,m) rss_all<-rep(NA,brsim) sred_rss<-rep(NA,brsim) s_rss<-rep(NA,brsim) cv<-rep(NA,brsim) l95<-rep(NA,brsim) u95<-rep(NA,brsim) for (b in 1:brsim){ List = list() for (j in 1:m){ uzorak=sample(x, k^2) for (i in 1:k){ pod_uzorak=uzorak[(((i-1)*k)+1):(i*k)] pod_uzorak_sort=sort(pod_uzorak) rss[i]=pod_uzorak_sort[i] } rss_uzorak=c(rss) List[[length(List)+1]] = rss_uzorak } rss_all=c(do.call("cbind",List)) trm=mean(rss_all,trim=0.1) s_rss=sqrt((sum((rss_all-trm)^2))/(n-1)) sred_rss=(1/(k*m))*sum(rss_all) cv[b]<-s_rss/sred_rss } l95<-cv-qnorm(0.975)*sqrt(((cv^2)/(n-1))*(0.5+(cv^2))) u95<-cv+qnorm(0.975)*sqrt(((cv^2)/(n-1))*(0.5+(cv^2))) mean((l95<=0.4)&(0.4<=u95)) sirina=u95-l95 mean(sirina) </pre>	<pre> x<-rnorm(20,5,2) brsim<-3 k=4 m=3 r=1 n=k*m srs<-rep(NA,(2*r)) uzorakSRS<-rep(NA,m) uzorakSRS_all<-rep(NA,brsim) rss<-rep(NA,(k-2*r)) uzorakRSS<-rep(NA,m) uzorakRSS_all<-rep(NA,brsim) sred_prss<-rep(NA,brsim) s_prss<-rep(NA,brsim) cv<-rep(NA,brsim) l95<-rep(NA,brsim) u95<-rep(NA,brsim) for (b in 1:brsim){ List = list() List1 = list() for (j in 1:m){ for (t in 1:(2*r)){ uzorak=sample(x,1) srs[t]=uzorak } List2 = list() for (z in 1:(k-2*r)){ uzorak1=sample(x,k) List2[[z]]<-uzorak1 } uzorakMATRIX=do.call(rbind,List2) uzorakMATRIXsort=t(apply(uzorakMATRIX,1,sort)) uzorakVEC=as.vector(t(uzorakMATRIXsort)) for (z in 1:(k-2*r)){ pod_uzorak=uzorakVEC[(((z-1)*k)+1):(z*k)] rss[z]=pod_uzorak[r+z] } uzorakSRS=c(srs) List1[[length(List1)+1]] = uzorakSRS uzorakRSS=c(rss) List[[length(List)+1]] = uzorakRSS } uzorakSRS_all=c(do.call("cbind",List1)) uzorakRSS_all=c(do.call("cbind",List)) uzorakPRSS=c(uzorakSRS_all,uzorakRSS_all) sred_prss=(1/n)*sum(uzorakPRSS) s_prss=sqrt(((1/(n-1))*sum((uzorakPRSS-sred_prss)^2))) cv[b]<-s_prss/sred_prss } l95<-cv-qnorm(0.975)*sqrt(((cv^2)/(n-1))*(0.5+(cv^2))) u95<-cv+qnorm(0.975)*sqrt(((cv^2)/(n-1))*(0.5+(cv^2))) mean((l95<=0.4)&(0.4<=u95)) sirina=u95-l95 mean(sirina) </pre>

Tabela 3.36: Kodovi za Curto-ov & Pinto-ov interval i McKay-ev interval za koeficijent varijacije

Curto-ov & Pinto-ov interval poverenja	Mckay-ev interval poverenja
<pre> x<-rnorm(20,5,2) brsim<-10 n<-5 cv<-rep(NA,brsim) l95<-rep(NA,brsim) u95<-rep(NA,brsim) for (i in 1:brsim){ uzorak<-sample(1:20,n) x1<-x[uzorak] suma_x1=sum(x1) x1kv=x1^2 suma_x1kv=sum(x1kv) s=sqrt((suma_x1kv-((suma_x1)^2)/n)/(n-1)) m=mean(x1) cv[i]<-s/m } l95<-cv-qnorm(0.975)*sqrt((1/n)*(cv^4+0.5*cv^2)) u95<-cv+qnorm(0.975)*sqrt((1/n)*(cv^4+0.5*cv^2)) mean((l95<=0.4)&(0.4<=u95)) sirina=u95-l95 mean(sirina) </pre>	<pre> x<-rnorm(20,5,2) brsim<-10 n<-10 cv<-rep(NA,brsim) l95<-rep(NA,brsim) u95<-rep(NA,brsim) for (i in 1:brsim){ uzorak<-sample(1:20,n) x1<-x[uzorak] suma_x1=sum(x1) x1kv=x1^2 suma_x1kv=sum(x1kv) s=sqrt((suma_x1kv-((suma_x1)^2)/n)/(n-1)) m=mean(x1) cv[i]<-s/m } l95<-cv*(((qchisq(0.975,n-1)/n)-1)*cv^2+(qchisq(0.975,n-1)/(n-1)))^(-1/2) u95<-cv*(((qchisq(0.025,n-1)/n)-1)*cv^2+(qchisq(0.025,n-1)/(n-1)))^(-1/2) #u95[is.na(u95)] <- 0 mean((l95<=0.4)&(0.4<=u95)) sirina=u95-l95 mean(sirina) </pre>

IV MARKOV CHAIN MONTE CARLO (MCMC) METODE KOD BAYES-OVSKOG OCENJIVANJA

Osnovna ideja Bayes-ovske statistike je da se parametrima koji su predmet proučavanja pridruže raspodele verovatnoća koje pokazuje apriorna uverenja o tim parametrima. Na osnovu apriornih uverenja i prikupljenih podataka izražavaju se nova znanja o parametrima koja su predstavljena aposteriornim raspodelama. Određivanje aposteriorne raspodele, a zatim i izbor uzorka iz te raspodele često može biti jako komplikovano. Ovaj problem se prevazilazi upotrebom Markov Chain Monte Carlo (skraćeno MCMC) metoda.

Na početku ove glave će biti razmatrane osnovne razlike između klasičnog i Bayes-ovskog pristupa u statistici, a zatim će biti prikazana osnovna ideja Bayes-ovskog zaključivanja zasnovana na apriornoj i aposteriornoj raspodeli. U nastavku će biti predstavljena svojstva lanaca Markova, kao i ideja na kojoj je utemeljena primena Markov Chain Monte Carlo metoda. Detaljno će biti objašnjeni osnovni MCMC algoritmi: Metropolis-Hastings-ov algoritam i njegove varijante, kao i Gibbs-ov algoritam. Zatim će biti analizirani osnovni pojmovi na osnovu kojih se mogu doneti zaključci o konstruisanim lancima Markova. Način konstruisanja lanaca Markova na realnom primeru biće prikazan na kraju glave IV.

4.1 Aposteriorna raspodela i Bayes-ovsko zaključivanje

Predmet razmatranja biće klasični i Bayes-ovski pristup u statistici i osnovne razlike između njih. Zatim će pažnja biti posvećena apriornoj i aposteriornoj raspodeli i njihovoj ulozi u Bayes-ovskom zaključivanju. Na kraju odeljka biće prikazan ilustrativni primer radi jednostavnijeg razumevanja prethodno izloženog sadržaja. O Bayes-ovskoj analizi se može više pročitati u Box i Tiao (1973), Berger (1985), Antelman (1996), Bayarri i Berger (2004), Lee (2004), Albert (2007), Bolstad (2007), Lynch (2007), Jackman (2009), Held i Bové (2014), Liu i Wasserman (2014), Marin i Robert (2014).

4.1.1 Poređenje klasičnog i Bayes-ovskog pristupa

Moguće je razlikovati dva pristupa u statistici: klasični (frekventistički) i Bayes-ovski pristup. Tvorac Bayes-ovske teorije bio je Thomas Bayes. Neki od pristalica Bayes-ovskog načina razmišljanja bili su Simon Laplace, Harold Jeffreys, Jimmy Savage. Sve do početka 90-tih godina XX veka primena Bayes-ovskih metoda u analizi realnih problema je bila jako ograničena, jer su se neophodna izračunavanja mogla izvršiti analitički samo prilikom rešavanja jednostavnih problema.

Osnovna razlika između navedenih pristupa se može predstaviti sledećim primerom.

Primer 1 Neka je predmet istraživanja određivanje verovatnoće javljanja pisma prilikom bacanja nehomogenog novčića.

- Prema klasičnom pristupu, verovatnoća se određuje na sledeći način: Novčić se baca veliki broj puta (N) i beleži se broj javljanja pisma (N_1). Odnos N_1/N predstavlja aproksimaciju verovatnoće da će se prilikom bacanja novčića javiti pismo.
- Pristalice Bayes-ovskog pristupa određuju verovatnoću prateći sledeće korake:

Koristeći subjektivnu verovatnoću, na osnovu sopstvenih uverenja, znanja ili prošlog iskustva pretpostavlja se da, na primer, verovatnoća javljanja pisma iznosi 0,5. Ova verovatnoća se zove apriorna verovatnoća. Zatim se putem eksperimenta sakupljaju podaci i vrši se korekcija apriorne verovatnoće. Kao rezultat se dobija aposteriorna verovatnoća javljanja pisma.

U nastavku se navode osnovne karakteristike klasičnog pristupa (Cowles, 2013):

- relativna frekvencija dobijena ponavljanjem eksperimenta veliki broj puta se koristi kao aproksimacija verovatnoće;
- parametri skupa su fiksni (konstante);
- podaci se smatraju slučajnim; tačkaste i intervalne ocene nepoznatih parametara posmatraju se kao funkcija podataka;
- ne koriste se apriorna (prethodna) znanja i iskustva;
- izračunavanja uglavnom nisu komplikovana.

Osnovne karakteristike Bayes-ovskog pristupa su (Cowles, 2013):

- verovatnoća predstavlja stepen uverenja da će se neki događaj desiti;
- parametri skupa su slučajne promenljive i određeni su raspodelom verovatnoća (postoje uverenja o mogućim raspodelama i njihovim parametrima);
- podaci se smatraju fiksnim (nakon što se jednom prikupe koriste se da bi se korigovala apriorna uverenja);
- apriorna znanja o parametrima mogu biti subjektivna, pri čemu pristrasnost može dovesti do loših rezultata;
- aposteriorna znanja o parametrima se mogu korigovati novim podacima;
- izračunavanja mogu biti vrlo komplikovana, ukoliko se vrši integracija preko velikog broja nepoznatih parametara.

4.1.2 Osnovna ideja Bayes-ovskog zaključivanja

Neka je X_1, \dots, X_n uzorak izabran iz neke raspodele sa parametrom θ . Verovatnoća dobijanja datih podataka ako je dat parametar θ biće označena sa $p(X_1, \dots, X_n | \theta)$. Kod Bayes-ovskog zaključivanja θ je nepoznati parametar koji predstavlja slučajnu promenljivu. Raspodela kojom se izražava uverenje o nepoznatom parametru θ pre nego što se prikupe podaci, odnosno pre sprovođenja eksperimenta, naziva se apriorna raspodela (engl. prior distribution) parametra θ . Apriorna raspodela od θ će biti označena sa $p(\theta)$.

Nakon sprovođenja eksperimenta i prikupljenih podataka X_1, \dots, X_n koriguju se uverenja o nepoznatom parametru θ i definiše (odnosno izračunava) aposteriorna raspodela (engl. posterior distribution) parametra θ . Aposteriorna raspodela parametra θ , odnosno aposteriorna verovatnoća od θ , za dati uzorak X_1, \dots, X_n biće označena sa $p(\theta | X_1, \dots, X_n)$.

Aposteriorna raspodela kombinuje apriorna uverenja o nepoznatom parametru θ i informacije o θ koje su sadržane u uzorku X_1, \dots, X_n . Na ovoj raspodeli se zasniva Bayes-ovsko zaključivanje, odnosno moguće je odrediti Bayes-ovske tačkaste i intervalne ocene nepoznatih parametara.

Koristeći Bayesovu teoremu, aposteriorna raspodela se može zapisati u sledećoj formi (Liu & Wasserman, 2014, str. 302):

$$p(\theta|X_1, \dots, X_n) = \frac{p(X_1, \dots, X_n|\theta)p(\theta)}{p(X_1, \dots, X_n)} = \frac{L_n(\theta)p(\theta)}{c_n} \propto L_n(\theta)p(\theta), \quad (4.1)$$

gde je

$$c_n = p(X_1, \dots, X_n) = \int p(X_1, \dots, X_n|\theta)p(\theta)d\theta = \int \prod_{i=1}^n p(X_i|\theta)p(\theta)d\theta$$

normalizaciona konstanta. Simbol \propto predstavlja proporcionalnost. U izrazu (4.1) $L_n(\theta) = \prod_{i=1}^n p(X_i|\theta)$ je funkcija verodostojnosti. Može se reći da je aposteriorna raspodela zapravo proporcionalna proizvodu funkcije verodostojnosti i apriorne raspodele.

Ako apriorna i aposteriorna raspodela pripadaju istoj familiji raspodela, tada se za ove raspodele kaže da su konjugovane raspodele. Ukoliko apriorna i aposteriorna raspodela ne pripadaju istoj familiji raspodela, tada izračunavanje konstante c_n može predstavljati problem. Jedan od načina za rešavanje ovog problema jeste primena MCMC metoda. Primenom MCMC metoda moguće je birati uzorke iz aposteriorne raspodele $p(\theta|X_1, \dots, X_n)$, pri čemu nije potrebno poznavanje konstante c_n .

4.2 Osnovna ideja MCMC metoda

Markov Chain Monte Carlo metode predstavljaju metode generisanja slučajnih uzoraka iz određene raspodele. U cilju što jednostavnijeg razumevanja ideje MCMC metoda, na samom početku biće definisani pojmovi slučajan proces, proces Markova i lanac Markova. Biće razmatrana svojstva lanaca Markova i na kraju odeljka će biti predstavljena osnovna ideja MCMC metoda. Ovom tematikom su se bavili Geyer (1992), Raftery i Lewis (1996), Roberts (1996), Green (2001), Gamerman i Lopes (2006), Cowles (2013), Gelman et al. (2014), McElreath (2020).

Definicija 1. (Nešić, 2010, str. 15) Neka $\{X_t\}_{t \in T}$ predstavlja familiju slučajnih promenljivih koja je definisana na nekom prostoru verovatnoća i neka je T beskonačan skup parametara $t \in T$. Ako je skup T podskup skupa \mathbf{R} , tada se $\{X_t\}_{t \in T}$ naziva slučajni proces.

Definicija 2. (Nešić, 2010, str. 17) Slučajni proces $\{X_t\}_{t \in T}$ predstavlja proces Markova ako za svaki prirodan broj n i za svaki niz tačaka $s_1 < s_2 < \dots < s_n < t$ iz skupa T važi tvrđenje:

$$P(X_t \in B | X_{s_1}, X_{s_2}, \dots, X_{s_n}) = P(X_t \in B | X_{s_n}),$$

za svaki Borel-ov skup B .

Navedenom definicijom se tvrdi da su procesi Markova slučajni procesi za koje važi da budućnost zavisi isključivo od sadašnjeg stanja. Dakle, ukoliko je poznato sadašnje stanje, poznavanje prošlosti nije neophodno.

Lanci Markova su vrsta procesa Markova sa diskretnim vremenom koji mogu uzimati konačno mnogo različitih vrednosti. Neka je predmet razmatranja sistem koji se u vremenskim trenucima t_0, t_1, \dots može naći u nekom od n mogućih stanja. Zapis $X_k = i$ znači da se u nekom trenutku t_k sistem nalazi u stanju i . Slučajni proces $\{X_t\}_{t \in T}$ se definiše kao lanac Markova ako važi sledeće tvrđenje (Nešić, 2010, str. 20):

$$P(X_{t+1} = j | X_0, X_1, \dots, X_t) = P(X_{t+1} = j | X_t).$$

Lanac Markova se može posmatrati kao niz slučajnih prelazaka iz jednog stanja u drugo pri čemu je potrebno znati verovatnoću prelaska u naredno stanje.

Vektor početnih verovatnoća $p^{(0)} = (p_1^{(0)}, \dots, p_n^{(0)})$, gde je $p_i^{(0)} = P(X_0 = i)$ i matrica prelaza Π (engl. transition kernel) određuju lanac Markova. Matrica prelaza obuhvata verovatnoće prelaza iz stanja i u trenutku t_k u stanje j u trenutku t_l . Ako verovatnoće prelaza ne zavise od vremena t_k i t_l , lanac Markova je homogen. U tom slučaju se matrica prelaza sastoji od verovatnoća (Nešić, 2010, str. 20):

$$p_{ij} = P(X_{t+1} = j | X_t = i), \quad t = 0, 1, \dots, i, \quad j = 1, 2, \dots, n,$$

a vektor verovatnoća u k -tom koraku određuje se kao proizvod vektora početnih vrednosti i matrice prelaza:

$$p^{(k)} = p^{(0)} \Pi^k, \quad k = 1, 2, \dots$$

gde je $p^{(k)} = (p_1^{(k)}, \dots, p_n^{(k)})$ i $p_i^{(k)} = P(X_k = i)$, $i = 1, 2, \dots, n$.

U nastavku će biti navedena osnovna svojstva lanaca Markova (Sahu, 2000):

- Stacionarnost: kada $n \rightarrow \infty$ lanac Markova konvergira u raspodeli ka svojoj ravnotežnoj (engl. stationary) raspodeli.
- Ireducibilnost (engl. irreducibility): Pod pretpostavkom da postoji ravnotežna raspodela, ona će biti jedinstvena ako je lanac ireducibilan. Ireducibilnost podrazumeva da je potreban konačan broj koraka da se iz bilo kog stanja dostigne bilo koji skup stanja.
- Aperiodičnost (engl. aperiodicity): Lanac Markova koji uzima konačan broj vrednosti je aperiodičan ukoliko najveći zajednički delilac vremena povratka na neko stanje j iznosi 1 (Na primer, neka se beleži neophodan broj koraka potreban za povratak u stanje j . Najveći zajednički delilac ovih brojeva bi trebalo da iznosi 1 ili u retkim slučajevima 2. Ako to nije slučaj, lanac nije aperiodičan.).

- Ergodičnost (engl. ergodicity): Pod uslovom da lanac Markova ima ravnotežnu raspodelu i da je aperiodičan i ireducibilan, smatra se da poseduje svojstvo ergodičnosti.

Svojstvo ergodičnosti lanca Markova ima značajnu ulogu u primeni MCMC metoda. Suština MCMC metoda se može predstaviti sledećom definicijom:

Definicija 3. (Robert & Casella, 2004, str. 268) Svaka metoda za simuliranje uzorka iz raspodele f kojom se generiše ergodički lanac Markova $\{X_t\}_{t \in T}$ čija je ravnotežna raspodela f predstavlja MCMC metodu.

MCMC metode se koriste za izračunavanja komplikovanih integrala, tako da se prethodna definicija može objasniti na sledeći način. Neka je predmet razmatranja integral $I = \int h(x) f(x) dx$ koji nije moguće analitički izračunati, već ga je potrebno aproksimirati. Dalje, neka $f(x)$ predstavlja gustinu raspodele iz koje se simulira uzorak. Dakle, funkcijom f je opisana ciljna raspodela (engl. target distribution). Da bi se aproksimirao integral I , uzorak X_1, X_2, \dots se ne mora direktno birati iz ciljne raspodele, već se može simulirati primenom Markov Chain Monte Carlo metoda. Ideja na kojoj su zasnovane MCMC metode podrazumeva da se za izabranu ciljnu raspodelu f , skup stanja i proizvoljno početno stanje X_0 konstruiše lanac Markova čija ravnotežna raspodela dobro aproksimira ciljnu raspodelu. Za skup stanja se posmatraju verovatnoće prelaza iz jednog u drugo stanje i po određenom kriterijumu se bira stanje koje se dodaje u lanac. Dobijeni lanac poseduje svojstvo ergodičnosti, što znači da će za $n \rightarrow \infty$ ravnotežna raspodela lanca težiti ciljnoj raspodeli. Ovo se može zapisati u formi ergodičke teoreme (Liu & Wasserman, 2014, str. 334):

$$\frac{1}{n} \sum_{i=1}^n h(X_i) \xrightarrow{P} E_f(h(X)) = I, n \rightarrow \infty.$$

Dakle, moguće je proceniti bilo koje očekivanje aposteriorne raspodele koristeći uzorak iz te raspodele (Veljović, 2019). Ergodička teorema tvrdi da se realizovane vrednosti lanca mogu koristiti kao uzorak iz ravnotežne raspodele. S obzirom da lanac poseduje svojstvo ergodičnosti, sledi da je ireducibilan, što znači da sva stanja međusobno komuniciraju. Dakle, lanac ne zavisi od početnog stanja, odnosno bez obzira na početno stanje lanac se u nekom trenutku sa određenom verovatnoćom može naći u stanju j .

4.3 Vrste MCMC algoritama

Metropolis et al. (1953) su razvili jedan od prvih MCMC algoritama pod nazivom Metropolis-ov algoritam. Hastings (1970) je izvršio generalizaciju Metropolis-ovog algoritma i algoritam je dobio naziv Metropolis-Hastings-ov algoritam. Na početku odeljka će biti objašnjen Metropolis-Hastings-ov algoritam, kao i njegove varijante nezavisni Metropolis-Hastings-ov algoritam (engl. Independence Metropolis-Hastings algorithm) i Metropolis-Hastings-ov algoritam sa slučajnim hodom (engl. random walk Metropolis-Hastings). Zatim će pažnja biti posvećena Gibbs-ovom algoritmu. U nastavku će biti reči o dijagnostici lanca Markova dobijenog primenom nekog od MCMC algoritama, kao i osnovnim pojmovima koji se u tu svrhu koriste. O MCMC algoritmima se diskutovalo u radovima Tierney (1994), Chib i Greenberg (1995), Carlin i Louis (2000), Robert i Casella (2004, 2010), Givens i Hoeting (2005), Lynch (2007), Kruschke (2010), Liu i Wasserman (2014).

4.3.1 Metropolis-Hastings-ov algoritam

Metropolis-Hastings-ov algoritam se primenjuje u situacijama kada je teško generisati uzorke iz neke raspodele f . Ali ako je moguće računati vrednosti neke druge funkcije \tilde{f} koja je proporcionalna funkciji gustine f , biće moguće i uzorkovanje iz ciljne raspodele f .⁵⁶

Metropolis-Hastings-ovim algoritmom se generiše lanac Markova koristeći tzv. raspodelu predloga (engl. proposal distribution) za predlaganje novih stanja koja će biti prihvaćena ili odbačena sa određenom verovatnoćom. U cilju konstruisanja lanca Markova, za svako stanje u kome se lanac može naći definiše se uslovna funkcija gustine $q(y|x)$ koja predlaže kandidata y za sledeću vrednost uzorka (za datu prethodnu vrednost uzorka x) (Veljović, 2019). Raspodela čija je funkcija gustine data sa $q(y|x)$ je raspodela predloga. Iz raspodele predloga je moguće direktno birati uzorke.

Postupak generisanja lanca Markova X_0, X_1, \dots, X_n koristeći Metropolis-Hastings-ov algoritam se predstavlja sledećim koracima (Liu & Wasserman, 2014, str. 334):

1. Proizvoljno izabrati početno stanje X_0 .
2. Za dato $X_i, i=1, \dots, n-1$ generisati X_{i+1} na sledeći način:
 - generisati predloženu vrednost $Y \sim q(y|X_i)$;
 - u cilju donošenja odluke o prihvatanju ili odbacivanju predložene vrednosti Y , određuje se verovatnoća $r \equiv r(X_i, Y)$ gde je

$$r(x, y) = \min\left(\frac{f(y) q(x|y)}{f(x) q(y|x)}, 1\right),^{57} \quad (4.2)$$

- iz uniformne raspodele⁵⁸, $U(0,1)$, generisati U . Ako je $U < r$ tada se prihvata predložena vrednost Y i novo stanje, odnosno naredna vrednost u lancu će biti $X_{i+1} = Y$. U suprotnom, odbacuje se predložena vrednost Y i naredna vrednost u lancu će biti određena trenutnim stanjem $X_{i+1} = X_i$.

⁵⁶„Činjenica da je potrebno da $\tilde{f}(x)$, bude samo proporcionalna željenoj gustini, ali ne i jednaka, u potpunosti nam omogućava ovo, jer je računanje normalizacione konstante obično najveći problem (posebno u Bajesovskom pristupu, kada nam je ciljna raspodela aposteriorna raspodela parametara od interesa).“ (Veljović, 2019).

⁵⁷ $r(x, y) = \min\left(\frac{\tilde{f}(y) q(x|y)}{\tilde{f}(x) q(y|x)}, 1\right) = \min\left(\frac{f(y) q(x|y)}{f(x) q(y|x)}, 1\right)$

⁵⁸Neprekidna slučajna promenljiva X ima uniformnu raspodelu sa parametrima lokacije a i b , tako da je $a < b$ (u oznaci $X \sim U(a, b)$), ako je njena funkcija gustine oblika (Đorić et al., 2007, str. 332):

$$f(x) = \frac{1}{b-a}, x \in [a, b].$$

Normalna raspodela se često koristi kao raspodela predloga. Specijalan slučaj Metropolis-Hastings-ovog algoritma je Metropolis-ov algoritam, kod koga je raspodela predloga simetrična. U slučaju kada je raspodela predloga simetrična, važi da je $q(y|x) = q(x|y)$. Tada je

$$r = \min\left(\frac{f(y)}{f(x)}, 1\right).$$

Algoritam se zaustavlja kada je sproveden dovoljan broj iteracija n , odnosno generisan lanac dovoljne veličine za potrebe istraživanja.

U nastavku će biti predstavljene dve varijante Metropolis-Hastings-ovog algoritma. Reč je o nezavisnom Metropolis-Hastings-ovom algoritmu i Metropolis-Hastings-ovom algoritmu sa slučajnim hodom.

Nezavisni Metropolis-Hastings-ov algoritam

Osnovna razlika nezavisnog Metropolis-Hastings-ovog algoritma u odnosu na Metropolis-Hastings-ov algoritam je da je raspodela predloga uvek ista. Dakle, raspodela predloga ne zavisi od trenutnog stanja u kom se lanac nalazi i biće označena sa $g(y)$. Sledi da je $q(y|x) = g(y)$. Iako su generisane vrednosti nezavisne, vrednosti u lancu su zavisne, jer verovatnoća prihvatanja stanja Y zavisi od trenutnog stanja.

Postupak generisanja lanca Markova X_0, X_1, \dots, X_n koristeći nezavisni Metropolis-Hastings-ov algoritam se zasniva na sledećim koracima (Liu & Wasserman, 2014, str. 338; Robert & Casella, 2004, str. 276):

1. Proizvoljno izabrati početno stanje X_0 .
2. Za dato $X_i, i = 1, \dots, n-1$ generisati X_{i+1} na sledeći način:
 - generisati predloženu vrednost $Y \sim g(y)$;
 - odrediti verovatnoću $r \equiv r(X_i, Y)$ gde je

$$r(x, y) = \min\left(\frac{f(y) g(x)}{f(x) g(y)}, 1\right);$$

- iz uniformne raspodele, $U(0,1)$, generisati U . Ako je $U < r$ prelazi se u novo stanje, inače se ostaje u trenutnom stanju.

Metropolis-Hastings-ov algoritam sa slučajnim hodom

Neka se predloženo stanje Y određuje na sledeći način:

$$Y = X_i + \varepsilon_i,$$

gde je ε_i slučajna promenljiva sa raspodelom g koja ne zavisi od trenutnog stanja X_i . U tom slučaju je $q(y|x) = g(y-x)$. Za raspodelu slučajne promenljive ε_i najčešće se biraju simetrične raspodele poput normalne, Student-ove t i uniformne raspodele.

Metropolis-Hastings-ov algoritam modifikovan na prethodno opisan način zove se Metropolis-Hastings-ov algoritam sa slučajnim hodom.

Na osnovu narednih koraka objašnjava se Metropolis-Hastings-ov algoritam sa slučajnim hodom (Liu & Wasserman, 2014, str. 338; Robert & Casella, 2004, str. 288):

1. Proizvoljno izabrati početno stanje X_0 .
2. Za dato $X_i, i = 1, \dots, n-1$ generisati X_{i+1} na sledeći način:
 - generisati predloženu vrednost $Y \sim g(|y-x_i|)$;
 - odrediti verovatnoću $r \equiv r(X_i, Y)$ gde je

$$r(x, y) = \min\left(\frac{f(y)}{f(x)}, 1\right);$$

- generisati U iz uniformne raspodele, $U(0,1)$. Ako je $U < r$ prelazi se u novo stanje, inače se ostaje u trenutnom stanju.

4.3.2 Gibbs-ov algoritam

Gibbs-ov algoritam je dobio naziv prema fizičaru Josiah Gibbsu. Koristi se ukoliko su sve uslovne raspodele poznate u zatvorenoj formi i ako se iz tih raspodela mogu direktno birati uzorci. Primenom Gibbs-ovog algoritma se omogućava da se višedimenzionalni problemi reše raščlanjivanjem na nekoliko jednodimenzionalnih. Konstruisanje lanca Markova pomoću Gibbs-ovog algoritma biće objašnjeno na primeru dvodimenzionalnog problema. Pretpostavlja se da (X, Y) ima funkciju gustine $f_{X,Y}(x, y)$. Neka su uslovne raspodele $f_{X|Y}(x|y)$ i $f_{Y|X}(y|x)$ poznate i neka je moguće izabrati uzorke iz ovih raspodela. Početne vrednosti će biti označene sa (X_0, Y_0) i neka u n -tom koraku lanac uzima vrednosti (X_n, Y_n) . Postupak generisanja (X_{n+1}, Y_{n+1}) koristeći Gibbs-ov algoritam se predstavlja sledećim koracima (Liu & Wasserman, 2014, str. 338):

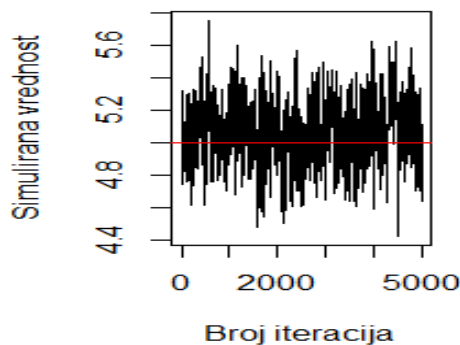
$$\begin{aligned} X_{n+1} &\sim f_{X|Y}(x|Y_n) \\ Y_{n+1} &\sim f_{Y|X}(y|X_{n+1}). \end{aligned}$$

Navedeni postupak je potrebno ponavljati sve dok raspodela lanca ne konvergira ka ciljnoj raspodeli.

4.3.3 Dijagnostika lanca Markova

Nakon što se konstruiše lanac Markova potrebno je doneti o njemu određene zaključke. Mogu se analizirati konvergencija lanca, prisustvo autokorelacije u lancu, efektivna veličina uzorka (engl. effective sample size, skraćeno ESS), Gelman-Rubinov pokazatelj i stopa prihvatanja (engl. acceptance rate). Za više informacija o ovoj temi pogledati Cowles i Carlin (1996), Brooks i Roberts (1998), Gamerman i Lopes (2006), Gelman et al. (2014).

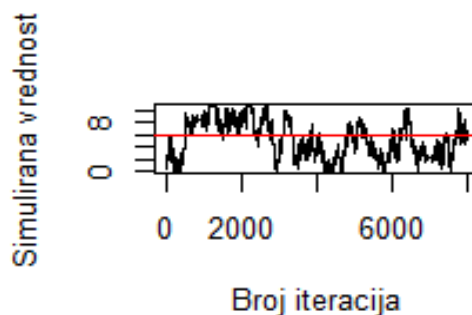
Jedan od načina da se utvrdi da li je lanac konvergirao je da se pogleda dijagram (engl. trace plot) koji prikazuje ponašanje lanca za određeni broj iteracija n^{59} . U nastavku slede tri primera. Na Slici 4.1 predstavljen je lanac koji brzo konvergira ka ciljnoj raspodeli. Ovaj lanac se dobro ponaša u smislu da u svakoj iteraciji prelazi iz jednog u drugo stanje sa velikim skokovima, zbog čega nije potrebno mnogo vremena da stigne sa jednog na drugi kraj raspodele (engl. good mixing).



Slika 4.1: Lanac Markova koji brzo konvergira ka ciljnoj raspodeli

Izvor: Izlaz iz programskog jezika R

Na Slici 4.2 prikazan je lanac kome je potreban veliki broj iteracija da konvergira (engl. poor mixing). Iako se u većini iteracija pomera iz jednog u drugo stanje, skokovi su mali.

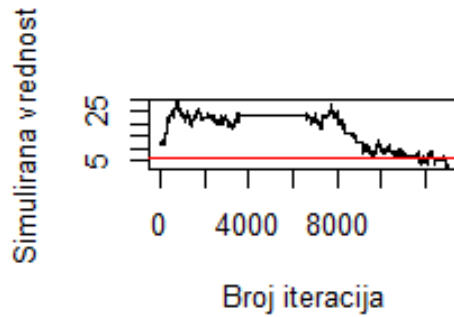


Slika 4.2: Lanac Markova koji sporo konvergira ka ciljnoj raspodeli

Izvor: Izlaz iz programskog jezika R

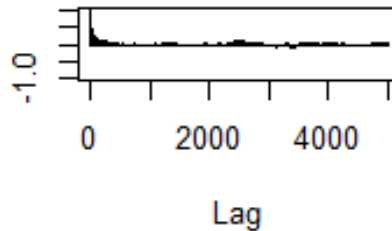
Primer lanca koji se ponaša veoma loše je lanac koji se zadržava u jednom stanju veliki broj iteracija. Takav lanac je prikazan na Slici 4.3. Uočava se da otprilike od 4000-te do 7000-te iteracije lanac nije prešao u naredno stanje.

⁵⁹ Sa povećanjem broja iteracija raspodela lanca konvergira ka ciljnoj raspodeli. Međutim, početne vrednosti lanca uglavnom ne potiču iz ciljne raspodele i zbog toga se mogu isključiti iz analize. Ove vrednosti predstavljaju “burn in” period.



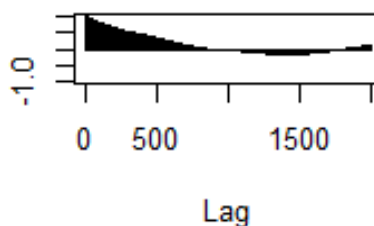
Slika 4.3: Lanac Markova koji se zadržava u jednom stanju veliki broj iteracija
Izvor: Izlaz iz programskog jezika R

Kao što je ranije objašnjeno, prelazak lanca u naredno stanje zavisi od trenutnog stanja, što znači da će u dobijenom uzorku postojati autokorelacija. Na Slici 4.4 se vidi da korelacija iščezava nakon manjeg broja pomaka (engl. lag). U tom slučaju raspodela lanca brže konvergira ka ciljnoj raspodeli.



Slika 4.4: Autokorelacija brzo iščezava
Izvor: Izlaz iz programskog jezika R

Na Slici 4.5 korelacija iščezava veoma sporo (nakon 1000 pomaka se i dalje održava se na niskom nivou), što ukazuje da će lanac sporo konvergirati ka ciljnoj raspodeli.



Slika 4.5: Autokorelacija sporo iščezava
Izvor: Izlaz iz programskog jezika R

S obzirom da su vrednosti u lancu korelisane, znači da će lanac veličine 10000 izabran iz ciljne raspodele dati manje informacija od slučajnog uzorka iste veličine iz iste raspodele u kome su vrednosti nekorelisane. Moguće je utvrditi veličinu lanca u kome su vrednosti nekorelisane, a koji će dati istu informaciju kao i lanac u kome su vrednosti korelisane. Kass et al. (1998, str. 99) uveli su termin efektivna veličina uzorka. Efektivna veličina uzorka određuje se formulom:

$$ESS = \frac{n}{1 + 2 \sum_{k=1}^{\infty} \rho(k)},$$

gde je n veličina lanca, a $\rho(k)$ je autokorelacija sa pomakom za k koraka. Uočava se da ukoliko su vrednosti nekorelisane efektivna veličina uzorka biće n . Preporučuje se povećanje broja iteracija sve dok efektivna veličina uzorka ne bude bar 100. Prilikom poređenja više lanaca boljim se smatra onaj čija je efektivna veličina uzorka veća.

Pored razmatranih dijagrama, da bi se utvrdilo da li lanac konvergira, Gelman i Rubin (1992) predložili su izračunavanje pokazatelja \hat{R} (engl. Potential Scale Reduction Factor). Potrebno je generisati bar dva lanca Markova čija će početna stanja biti različita (poželjne su što veće razlike između početnih stanja). S obzirom da se lanci zasnivaju na istoj aposteriornoj raspodeli, nakon konvergencije bi trebalo da varijanse između i unutar lanaca budu jednake. U tom slučaju je \hat{R} blizu 1. Ako varijanse između i unutar lanaca nisu jednake, \hat{R} je veće od 1. Gelman et al. (2014) su predložili da se za $\hat{R} < 1,1$ smatra da su lanci konvergirali.

Postavlja se pitanje šta preduzeti ukoliko lanac ne konvergira. Jedno rešenje je da se poveća broj iteracija. Drugi način da se ubrza konvergencija podrazumeva praktikovanje "good mixing" ponašanja lanca, što direktno zavisi od izbora raspodele predloga. Ako je varijansa raspodele predloga veoma mala, stopa prihvatanja⁶⁰ će biti visoka i lanac će sporo napredovati. Dakle, lanac će sa veoma malim skokovima prelaziti u naredno stanje, što će usporiti algoritam i dovesti do spore konvergencije lanca ka ciljnoj raspodeli (videti Sliku 4.2). Suprotno, ukoliko je varijansa veoma velika, stopa prihvatanja će biti niska. U tom slučaju će lanac biti „zaglavljen“, zbog čega će biti potrebno po nekoliko stotina iteracija da pređe u naredno stanje (Slika 4.3). I u ovom slučaju će lanac konvergirati veoma sporo (Lynch, 2007, str. 141). Optimalna stopa prihvatanja zavisi od predmeta istraživanja. Smatra se da je stopa prihvatanja u rasponu od 20% do 30% optimalna. Međutim, u praksi, čak i ako je stopa prihvatanja na ovom nivou, lanac može sporo konvergirati ako je varijansa raspodele predloga za jedan parametar veoma mala, a za drugi veoma velika.

4.4 Primena MCMC algoritama kod Bayesovskog ocenjivanja

Na jednom hipotetičkom primeru biće objašnjena konstrukcija lanaca Markova. MCMC metoda u Bayes-ovskoj statistici se može koristiti za dobijanje uzorka iz aposteriorne raspodele (Veljović, 2019).

Neka je generisan uzorak (X_1, \dots, X_n) iz log-normalne raspodele sa određenim vrednostima parametara $\omega = \omega_1$ i $CV = CV_1$.⁶¹

⁶⁰Stopa prihvatanja predstavlja odnos predloženih vrednosti koje su prihvaćene i n vrednosti koje čine lanac.

⁶¹Moguće je definisati neprekidnu slučajnu promenljivu X koja ima log-normalnu raspodelu sa parametrima ω i CV (u oznaci $X \sim LN(\omega, CV)$), gde ω predstavlja sredinu log-normalne raspodele, dok je CV koeficijent varijacije log-normalne raspodele. Ovo je zapravo alternativna parametrizacija za log-normalnu raspodelu. Veza između parametara μ i σ definisanih u odeljku 2.1.4 i parametara ω i CV se predstavlja sledećim formulama

$$\omega = \exp \left[\mu + \left(\frac{\sigma^2}{2} \right) \right] \text{ i } CV = \sqrt{\exp(\sigma^2) - 1}.$$

Dalje, kao što je poznato u Bayes-ovskoj statistici parametri se tretiraju kao slučajne promenljive i potrebno je odrediti njihove aposteriorne raspodele. Radi jednostavnosti, pretpostaviće se da je ω poznato ($\omega = \omega_1$) odnosno da je potrebno odrediti aposteriornu raspodelu za CV . Za parametar CV mora se izabrati apriorna raspodela. Neka je to Beta raspodela⁶² sa poznatim parametrima a i b . Ova raspodela je proizvoljno izabrana, odnosno predmet analize mogu biti bilo koje druge raspodele.

S obzirom da je veoma komplikovano odrediti oblik aposteriorne raspodele $p(CV|X_1, \dots, X_n)$ za parametar CV analitičkim putem biće primenjen jedan od MCMC algoritama, Metropolis-Hastings-ov algoritam.

Prateći korake za sprovođenje Metropolis-Hastings-ovog algoritma koji je objašnjen u odeljku 4.3.1 biće konstruisan lanac Markova lanac za koeficijent varijacije log-normalne raspodele. Početno stanje algoritma se bira proizvoljno i pretpostaviće se da je ono CV_0 . Neka raspodela predloga bude npr. normalna raspodela sa poznatim parametrima. Iz ove raspodele se generiše sledeća vrednost CV_1 (i tako u svakoj iteraciji). Izračunava se verovatnoća r data formulom (4.2)⁶³. Dalje se iz uniformne raspodele, $U(0,1)$, generiše slučajni broj U . Ako je $U < r$ prihvata se vrednost CV_1 . Moguće je konstruisati lance Markova veličine 10000 ili 20000 sa odgovarajućim “burn in“ periodima. Dobijene vrednosti daju uzorke iz aposteriorne raspodele.

Za svaki generisani lanac može se izračunati efektivna veličina uzorka. Takođe, na osnovu dobijenih uzoraka mogu biti konstruisane intervalne ocene za koeficijent varijacije log-normalne raspodele i izračunate njihove širine. Detaljnije o ovome biće objašnjeno u glavi V.

⁶² Nепрекидна slučajna promenljiva X ima dvoparametarsku beta raspodelu sa parametrima oblika $a > 0$ i $b > 0$ (u oznaci $X \sim B(a, b)$), ako je njena funkcija gustine oblika (Đorić et al., 2007, str. 30):

$$f(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}, 0 \leq x \leq 1.$$

⁶³Verovatnoća r se u prvoj iteraciji može zapisati na sledeći način $\frac{p(X_1, \dots, X_n | CV_1)p(CV_1)}{p(X_1, \dots, X_n | CV_0)p(CV_0)}$ (prilagođeno prema Veljović, 2019).

V POREĐENJE KLASIČNIH I BAYES-OVSKIH INTERVALA

Na početku ovog dela biće razmotreni Bayes-ovski intervali za nepoznati parametar: interval sa jednakim repovima (engl. equal-tailed interval) i interval sa najvećom aposteriornom gustinom (engl. highest posterior density interval, skraćeno HPD). Zatim će biti prikazani i analizirani rezultati za koeficijent varijacije log-normalne raspodele dobijeni prilikom konstruisanja lanaca Markova i to prosek aposteriorne raspodele, efektivna veličina uzorka, širina intervala sa jednakim repovima i širina intervala sa najvećom aposteriornom gustinom. U nastavku će biti dato poređenje širina nekih klasičnih intervala poverenja za koeficijent varijacije log-normalne raspodele sa širinama Bayes-ovskih intervala. Analiza će biti izvršena na simuliranim i realnim podacima.

5.1 Bayes-ovski intervali

Bayes-ovsko zaključivanje je zasnovano na aposteriornoj raspodeli. Bayes-ovski “intervali poverenja” se nazivaju intervali prekrivanja ili intervali verodostojnosti (engl. credible intervals) i određuju se na osnovu percentila aposteriorne raspodele. Neka je θ nepoznati parametar, a $1-\alpha$ verovatnoća da će interval sadržati nepoznati parametar.

Bayes-ovski $(1-\alpha)100\%$ interval sa jednakim repovima određuje se na osnovu $\alpha/2$ i $1-\alpha/2$ percentila aposteriorne raspodele $p(\theta|X_1, \dots, X_n)$ nepoznatog parametra θ i predstavlja se na sledeći način (Casella & Berger, 2002, str. 436):

$$I_\theta = [p(\theta|X_1, \dots, X_n)_{\alpha/2}, p(\theta|X_1, \dots, X_n)_{1-\alpha/2}].$$

Ako je aposteriorna raspodela $p(\theta|X_1, \dots, X_n)$ nepoznatog parametra θ unimodalna, tada se interval sa najvećom aposteriornom gustinom zapisuje kao (Casella & Berger, 2002, str. 448):

$$I_\theta = [\theta : p(\theta|X_1, \dots, X_n) \geq k],$$

gde je k izabrano tako da $\int_{\{\theta: p(\theta|X_1, \dots, X_n) \geq k\}} p(\theta|X_1, \dots, X_n) d\theta = 1-\alpha$.

Postoji razlika u interpretaciji klasičnih intervala poverenja i Bayes-ovskih intervala. U klasičnoj statistici se sa pouzdanošću od $(1-\alpha)100\%$ tvrdi da se parametar nalazi u dobijenom intervalu poverenja, odnosno ako se eksperiment ponovi veliki broj puta $(1-\alpha)100\%$ realizovanih intervala poverenja će sadržati nepoznati parametar. Verovatnoća da će se parametar naći u realizovanom intervalu iznosi 0 ili 1. U Bayes-ovskoj statistici $(1-\alpha)100\%$ odražava subjektivno uverenje odnosno osoba koja sprovodi eksperiment nakon kombinovanja apriornih informacija sa podacima je $(1-\alpha)100\%$ sigurna da će interval sadržati parametar. Ovde će se parametar naći unutar intervala sa nekom verovatnoćom, koja nije 0 niti 1 (Casella & Berger, 2002, str. 435).

5.2 Analiza klasičnih i Bayes-ovskih intervala za koeficijent varijacije log-normalne raspodele

U ovom delu rada će na hipotetičkom primeru biti objašnjena konstrukcija dvodimenzionalnih lanaca Markova za sredinu i koeficijent varijacije log-normalne raspodele primenom Metropolis-Hastings-ovog algoritma. Analiziraće se prosek aposteriorne raspodele za koeficijent varijacije log-normalne raspodele, efektivna veličina uzorka, kao i širine 95% intervala sa jednakim repovima i 95% intervala sa najvećom aposteriornom gustinom. Dalje će se izvršiti poređenje širina nekih klasičnih intervala poverenja za koeficijent varijacije log-normalne raspodele koji su predstavljeni u odeljku 3.2 sa širinama Bayes-ovskih intervala. Prvo će biti predstavljeni rezultati dobijeni na osnovu simulacione studije, a zatim rezultati dobijenim primenom realnih podataka.

5.2.1 Analiza lanaca Markova za koeficijent varijacije log-normalne raspodele

Neka se iz log-normalne raspodele sa vrednostima parametara $\omega = 6$ i $CV=2$ generišu uzorci veličine 10, 20, 50 i 80.

Dalje, neka su oba parametra nepoznata, odnosno neka je vektor parametara $\theta = (\omega, CV)$ nepoznat. Parametri ω i CV će biti razmatrani nezavisno, što znači da je zajednička apriorna raspodela $p(\theta)$ jednaka proizvodu apriornih raspodela $p(\omega)$ i $p(CV)$.

Neka apriorna raspodela za ω bude uniformna raspodela, $U(0,10)$. Što se tiče parametra CV razmatraće se tri apriorne raspodele:

- Beta raspodela, $B(3;0,4)$,
- Uniformna raspodela, $U(0,4)$ i
- Beta raspodela (četvoroparametarska)⁶⁴, $B_4(5;0,7;0;3)$.

Apriorne raspodele su proizvoljno izabrane, odnosno predmet analize može biti bilo koja druga raspodela.

Metropolis-Hastings-ov algoritam se primenjuje na isti način bez obzira da li je u pitanju jedan parametar ili vektor parametara ("Multi parameter Metropolis-Hastings", n.d.). Proizvoljno se bira početno stanje, na primer $\theta_0 = (8,3)$. Neka je θ_i vektor parametara (u koraku i) a θ' generisani vektor parametara iz raspodele predloga (u prvom slučaju raspodela predloga za sredinu i

⁶⁴Neprekidna slučajna promenljiva X ima četvoroparametarsku beta raspodelu sa parametrima oblika $a > 0$ i $b > 0$, i parametrima c i d koji predstavljaju najmanju i najveću vrednost slučajne promenljive (u oznaci $X \sim B_4(a,b,c,d)$), ako je njena funkcija gustine oblika (Đorić et al., 2007, str. 38):

$$f(x) = \frac{1}{B(a,b)} \frac{(x-c)^{a-1} (d-x)^{b-1}}{(d-c)^{a+b-1}}, c \leq x \leq d.$$

koeficijent varijacije je normalna⁶⁵ sa standardnim devijacijama 0,3 i 0,6, respektivno; u drugom slučaju raspodela predloga je normalna sa standardnim devijacijama 0,3 i 0,9).

Izračunava se verovatnoća $r \equiv r(\theta_i, \theta')$ po formuli (4.2)⁶⁶ i iz uniformne raspodele, $U(0,1)$, generiše slučajni broj U . Ukoliko je $U < r$ prihvata se vektor θ' .

Postupak se nastavlja sve dok se ne konstruišu dvodimenzionalni lanci Markova veličine 10000, odnosno 20000 sa “burn in” periodima koji iznose 1000, odnosno 5000, respektivno. Svaka dimenzija lanca Markova smatra se marginalnom aposteriornom raspodelom sredine odnosno koeficijenta varijacije a generisane vrednosti se mogu smatrati uzorcima iz marginalnih aposteriornih raspodela.

Na osnovu dobijenih uzoraka biće konstruisani intervali za koeficijent varijacije log-normalne raspodele i izračunate njihove širine. Takođe, biće izračunata efektivna veličina uzorka.

U Tabelama 5.1-5.3 predstavljeni su rezultati analize lanaca Markova za koeficijent varijacije log-normalne raspodele za različite apriorne raspodele. Prvo je analizirana beta⁶⁷ raspodela. U Tabeli 5.1 prikazani su rezultati analize lanaca Markova za koeficijent varijacije log-normalne raspodele ako je apriorna raspodela $B(3;0,4)$. Za uzorke veličine 10, prosek marginalne aposteriorne raspodele je najbliži vrednosti koeficijenta varijacije log-normalne raspodele ukoliko je raspodela predloga sa standardnim devijacijama 0,3 i 0,9 i ako je sprovedeno 10000 iteracija. Efektivna veličina uzorka je najveća za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija. Interval sa jednakim repovima je najuži za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija, dok je interval sa najvećom aposteriornom gustinom najuži za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija. Kada je reč o uzorcima veličine 20, prosek marginalne aposteriorne raspodele je najbliži vrednosti koeficijenta varijacije log-normalne raspodele za raspodelu predloga sa standardnim devijacijama 0,3 i 0,9 i 10000 iteracija, dok je efektivna veličina uzorka najveća za raspodelu predloga sa standardnim devijacijama 0,3 i 0,9 i 20000 iteracija. Oba razmatrana intervala su najuža za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija. Za uzorke veličine 50, prosek marginalne aposteriorne raspodele je najbliži vrednosti koeficijenta varijacije za raspodelu predloga sa standardnim devijacijama 0,3 i 0,9 i 10000 iteracija. Kada je reč o efektivnoj veličini uzorka i širinama intervala, najbolji rezultati su dobijeni za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija. Za uzorke veličine 80, prosek marginalne aposteriorne raspodele je najbliži vrednosti parametra za raspodelu predloga sa standardnim devijacijama 0,3 i 0,9 i 20000 iteracija. Što se tiče efektivne veličine uzorka i širina intervala, najbolji rezultati su ostvareni za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija.

⁶⁵ Sredina raspodele predloga predstavljena je trenutnom vrednošću parametara, dok se standardna devijacija za svaki od parametara bira proizvoljno (Ford, 2018).

⁶⁶ Verovatnoća r se može zapisati i na sledeći način $r(\theta_i, \theta') = \frac{p(X_1, \dots, X_n | \theta') p(\theta')}{p(X_1, \dots, X_n | \theta_i) p(\theta_i)}$ (“Multi parameter Metropolis-Hastings”, n.d.).

⁶⁷ Funkcija gustine beta raspodele nalazi se u odeljku 4.4.

Tabela 5.1: Rezultati analize lanaca Markova za koeficijent varijacije log-normalne raspodele ako je apriorna raspodela $B(3;0,4)$

Apriorna raspodela $B(3;0,4)$						
Broj iteracija/ Burn in	Standardne devijacije raspodele predloga	Veličina uzorka	Prosek marginalne aposteriorne raspodele	ESS	Širina intervala sa jednakim repovima	Širina intervala sa najvećom aposteriornom gustinom
10000/ 1000	0,3 i 0,6	10	1,38	140,24	2,521	2,390
		20	1,75	53,42	2,932	2,832
		50	1,42	84,83	2,785	2,714
		80	1,23	184,99	2,403	2,381
20000/ 5000	0,3 i 0,6	10	1,28	229,21	2,519	2,485
		20	1,32	196,51	2,613	2,519
		50	1,26	293,33	2,491	2,444
		80	1,53	118,02	2,994	2,641
10000/ 1000	0,3 i 0,9	10	1,74	57,69	3,342	3,230
		20	1,83	59,44	3,481	3,384
		50	1,95	85,93	3,480	3,375
		80	1,83	159,65	3,167	3,053
20000/ 5000	0,3 i 0,9	10	1,73	219,35	3,156	3,129
		20	1,51	200,82	3,033	2,967
		50	1,93	135,73	3,661	3,467
		80	1,91	117,75	3,542	3,446

Izvor: Izlaz iz programskog jezika R

U drugom slučaju je razmatrana uniformna⁶⁸ raspodela. U Tabeli 5.2 predstavljeni su rezultati analize lanaca Markova za koeficijent varijacije log-normalne raspodele ako je apriorna raspodela $U(0,4)$. Za uzorke veličine 10, prosek marginalne aposteriorne raspodele je najbliži vrednosti koeficijenta varijacije log-normalne raspodele i intervali su najuži ukoliko se koristi raspodela predloga sa standardnim devijacijama 0,3 i 0,6 i ako je sprovedeno 10000 iteracija. Efektivna veličina uzorka je najveća za raspodelu predloga sa standardnim devijacijama 0,3 i 0,9 i 20000 iteracija. Kada je reč o uzorcima veličine 20, prosek marginalne aposteriorne raspodele je najbliži vrednosti koeficijenta varijacije za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija. Efektivna veličina uzorka je najveća za raspodelu predloga sa standardnim devijacijama 0,3 i 0,9 i 20000 iteracija, dok su oba razmatrana intervala najuža za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija. Za uzorke veličine 50, prosek marginalne aposteriorne raspodele je najbliži vrednosti koeficijenta varijacije za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija. Kada je reč o efektivnoj veličini uzorka i širinama intervala, najbolji rezultati su dobijeni za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija. Za uzorke veličine 80, prosek marginalne aposteriorne raspodele je najbliži vrednosti parametra za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija. Takođe, što se tiče efektivne veličine uzorka i širine intervala sa jednakim repovima, najbolji rezultati su ostvareni za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija. Interval sa najvećom aposteriornom gustinom je najuži za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija.

⁶⁸Funkcija gustine uniformne raspodele nalazi se u odeljku 4.3.1.

Tabela 5.2: Rezultati analize lanaca Markova za koeficijent varijacije log-normalne raspodele ako je apriorna raspodela $U(0,4)$

Apriorna raspodela $U(0,4)$						
Broj iteracija/ Burn in	Standardne devijacije raspodele predloga	Veličina uzorka	Prosek marginalne aposteriorne raspodele	ESS	Širina intervala sa jednakim repovima	Širina intervala sa najvećom aposteriornom gustinom
10000/ 1000	0,3 i 0,6	10	2,79	173,16	5,165	5,110
		20	2,94	124,13	5,432	5,310
		50	2,55	161,28	5,296	5,247
		80	2,84	192,89	5,421	5,173
20000/ 5000	0,3 i 0,6	10	3,00	148,00	5,632	5,462
		20	2,77	112,75	5,696	5,664
		50	2,70	337,87	5,112	5,107
		80	2,82	280,87	5,370	5,233
10000/ 1000	0,3 i 0,9	10	2,89	315,14	5,661	5,603
		20	3,15	203,77	5,831	5,796
		50	3,27	267,62	6,002	5,850
		80	3,39	110,91	6,114	5,925
20000/ 5000	0,3 i 0,9	10	3,07	419,37	5,830	5,781
		20	3,28	310,71	6,137	6,087
		50	3,33	270,20	6,306	6,205
		80	3,54	173,19	6,558	6,409

Izvor: Izlaz iz programskog jezika R

Predmet analize u trećem slučaju je beta⁶⁹ raspodela. U Tabeli 5.3 predstavljeni su rezultati analize lanaca Markova za koeficijent varijacije log-normalne raspodele ako je apriorna raspodela $B_4(5;0,7;0;3)$. Kada je reč o uzorcima veličine 10, prosek marginalne aposteriorne raspodele je najbliži vrednosti koeficijenta varijacije za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija. Što se tiče efektivne veličina uzorka i širina intervala, najbolji rezultati su ostvareni za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija. Za uzorke veličine 20, prosek marginalne aposteriorne raspodele je najbliži vrednosti koeficijenta varijacije za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija, a efektivna veličina uzorka je najveća za raspodelu predloga sa standardnim devijacijama 0,3 i 0,9 i 20000 iteracija. Oba intervala su najuža za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija. Za uzorke veličine 50, najbolji rezultati po svim kriterijumima ostvareni su za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija. Kod uzoraka veličine 80, najbolji rezultati po svim kriterijumima dobijeni su za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija.

⁶⁹Funkcija gustine beta raspodele nalazi se u odeljku 5.2.1.

Tabela 5.3: Rezultati analize lanaca Markova za koeficijent varijacije log-normalne raspodele ako je apriorna raspodela $B_4(5;0,7;0;3)$

Apriorna raspodela $Beta_4(5;0,7;0;3)$						
Broj iteracija/ Burn in	Standardne devijacije raspodele predloga	Veličina uzorka	Prosek marginalne aposteriorne raspodele	ESS	Širina intervala sa jednakim repovima	Širina intervala sa najvećom aposteriornom gustinom
10000/ 1000	0,3 i 0,6	10	1,98	135,27	4,420	4,215
		20	2,05	318,30	4,116	3,975
		50	2,24	261,92	4,203	4,102
		80	2,07	282,21	4,226	4,196
20000/ 5000	0,3 i 0,6	10	2,21	309,11	4,226	4,169
		20	2,03	387,04	4,289	4,170
		50	2,13	463,79	4,165	4,099
		80	2,40	117,58	5,049	4,977
10000/ 1000	0,3 i 0,9	10	2,89	102,44	5,642	5,463
		20	2,60	288,56	4,919	4,908
		50	2,73	188,85	5,176	5,130
		80	2,78	181,19	5,116	5,076
20000/ 5000	0,3 i 0,9	10	2,67	279,15	5,535	5,102
		20	2,65	399,28	5,542	5,124
		50	2,69	291,91	5,381	5,250
		80	2,69	279,86	6,044	5,319

Izvor: Izlaz iz programskog jezika R

Za sve tri analizirane apriorne raspodele, zaključuje se da je u većini slučajeva interval sa jednakim repovima najuži za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija, dok je interval sa najvećom aposteriornom gustinom najuži za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija.

Dalje, izabrani su neki od klasičnih intervala poverenja za koeficijent varijacije, koji su predstavljeni u odeljku 3.2 i dobijene širine su prikazane u Tabeli 5.4. Uočava se da je za uzorke veličine 10, 20 i 50 najuži RSS BCa interval, dok je za uzorke veličine 80 najuži *bootstrap-t* interval.

Tabela 5.4: Širine 95% dvostranog intervala poverenja za koeficijent varijacije log-normalne raspodele

Veličina uzorka		10	20	50	80
<i>t</i>		1,326	1,271	0,543	1,049
Percentil		0,603	1,318	1,818	0,462
RSS Percentil	<i>k=2</i>	1,368	0,975	0,737	0,595
	<i>k=5</i>	0,872	0,957	0,477	0,484
<i>Bootstrap-t</i>		1,095	1,125	1,590	0,379
Bca		1,317	0,675	2,412	0,497
RSS BCa	<i>k=2</i>	1,136	0,997	0,298	0,481
	<i>k=5</i>	0,331	0,513	0,230	1,041
<i>Jackknife</i>		0,482	0,815	4,338	0,934
C&P		1,224	2,909	1,404	0,876
C&P trim 10%		1,062	2,745	1,828	1,103
RSS C&P	<i>k=2</i>	1,639	1,442	1,507	2,402
	<i>k=5</i>	1,708	2,256	0,992	2,055
RSS C&P trim 10%	<i>k=2</i>	0,811	0,734	0,807	1,155
	<i>k=5</i>	2,520	4,988	1,202	0,731

Izvor: Izlaz iz programskog jezika R

Za sve tri razmatrane apriorne raspodele, moguće je izvesti zaključak da su za simulirane podatke klasični intervali poverenja uži od Bayes-ovskih intervala.

5.2.2 Analiza lanaca Markova za koeficijent varijacije neto dobitka

Analiza na realnim podacima će biti sprovedena koristeći podatke o neto dobitku (neto dobitak je analiziran u odeljku 2.5.4). Upotrebom programa EasyFit utvrđeno je da se promenljiva neto dobitak može najbolje opisati log-normalnom raspodelom sa parametrima $\mu = 11,881$ i $\sigma = 1,7486$ (što potvrđuju rezultati prikazani u Tabeli 5.5).

Tabela 5.5: Kolmogorov-Smirnov i hi-kvadrat statistike testa

Kolmogorov-Smirnov		Hi-kvadrat	
Statistika testa	Rang	Statistika testa	Rang
0,03884	1	3,3511	1

Izvor: Izlaz iz programskog jezika EasyFit

Koristeći alternativnu parametrizaciju koja je objašnjena u odeljku 4.4 izračunati su parametri $\omega = 666503,7$ i $CV=4,5$.

Postupak koji se primenjuje identičan je postupku koji je objašnjen u odeljku 5.2.1. Razlika je u tome što će se pretpostaviti da je apriorna raspodela za sredinu uniformna raspodela, $U(0,1.000.000)$, dok će se za koeficijent varijacije razmatrati dve apriorne raspodele:

- Uniformna raspodela, $U(3,6)$ i
- Beta raspodela, $B_4(5;0,7;3;6)$.

U nastavku su u Tabelama 5.6-5.8 predstavljeni rezultati analize lanaca Markova za koeficijent varijacije neto dobitka za različite apriorne raspodele. Prvo je analizirana uniformna⁷⁰ raspodela. U Tabeli 5.6 prikazani su rezultati analize lanaca Markova za koeficijent varijacije neto dobitka ako je apriorna raspodela $U(3,6)$. Za uzorke veličine 10, prosek marginalne aposteriorne raspodele je najbliži vrednosti koeficijenta varijacije neto dobitka ukoliko se koristi raspodela predloga sa standardnim devijacijama 0,3 i 0,6 i ako je sprovedeno 10000 iteracija. Efektivna veličina uzorka je najveća za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija. Interval sa jednakim repovima je najuži za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija, dok je interval sa najvećom aposteriornom gustinom najuži za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija. Kada je reč o uzorcima veličine 20, prosek marginalne aposteriorne raspodele je najbliži vrednosti koeficijenta varijacije za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija. Efektivna veličina uzorka je najveća za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija, a oba razmatrana intervala su najuža za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija. Za uzorke veličine 50 i 80, prosek marginalne aposteriorne raspodele je najbliži vrednosti koeficijenta varijacije za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija, dok je efektivna veličina uzorka najveća za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija. Najbolji rezultati za intervale poverenja su dobijeni za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija.

Tabela 5.6: Rezultati analize lanaca Markova za koeficijent varijacije neto dobitka ako je apriorna raspodela $U(3,6)$

Apriorna raspodela $U(3,6)$						
Broj iteracija/ Burn in	Standardne devijacije raspodele predloga	Veličina uzorka	Prosek marginalne aposteriorne raspodele	ESS	Širina intervala sa jednakim repovima	Širina intervala sa najvećom aposteriornom gustinom
10000/ 1000	0,3 i 0,6	10	4,46	103,80	6,266	5,999
		20	4,77	108,74	6,230	6,030
		50	4,42	105,26	6,892	6,850
		80	5,55	51,35	8,197	8,149
20000/ 5000	0,3 i 0,6	10	4,84	95,94	6,261	5,647
		20	4,32	113,95	6,500	6,500
		50	4,22	110,79	7,990	7,520
		80	3,97	89,81	8,033	7,950
10000/ 1000	0,3 i 0,9	10	4,50	152,26	5,408	5,406
		20	4,38	264,68	5,584	5,525
		50	3,95	108,41	7,593	7,564
		80	4,11	113,09	8,349	7,895
20000/ 5000	0,3 i 0,9	10	4,33	91,18	5,885	5,688
		20	4,21	156,05	6,242	5,948
		50	4,05	86,24	8,055	7,628
		80	5,08	58,78	8,321	8,127

Izvor: Izlaz iz programskog jezika R

Predmet razmatranja u drugom slučaju je beta⁷¹ raspodela. U Tabeli 5.7 prikazani su rezultati analize lanaca Markova za koeficijent varijacije neto dobitka ako je apriorna raspodela $B_4(5;0,7;3;6)$. Za uzorke veličine 10, prosek marginalne aposteriorne raspodele je najbliži

⁷⁰Funkcija gustine uniformne raspodele nalazi se u odeljku 4.3.1.

⁷¹Funkcija gustine beta raspodele nalazi se u odeljku 5.2.1.

vrednosti koeficijenta varijacije neto dobitka ukoliko se koristi raspodela predloga sa standardnim devijacijama 0,3 i 0,9 i ako je sprovedeno 20000 iteracija. Efektivna veličina uzorka je najveća za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija. Interval sa jednakim repovima je najuži za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija, dok je interval sa najvećom aposteriornom gustinom najuži za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija. Kada je reč o uzorcima veličine 20, prosek marginalne aposteriorne raspodele je najbliži vrednosti koeficijenta varijacije za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija. Najbolji rezultati za efektivnu veličinu uzorka i širine intervala su ostvareni za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija. Za uzorke veličine 50, prosek marginalne aposteriorne raspodele je najbliži vrednosti koeficijenta varijacije u dva slučaja: za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija i raspodelu predloga sa standardnim devijacijama 0,3 i 0,9 i 20000 iteracija, dok je efektivna veličina uzorka najveća za raspodelu predloga sa standardnim devijacijama 0,3 i 0,9 i 20000 iteracija. Najbolji rezultati za intervale poverenja su dobijeni za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija. Kada je reč o uzorcima veličine 80, prosek marginalne aposteriorne raspodele je najbliži vrednosti koeficijenta varijacije za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija, dok je efektivna veličina uzorka najveća za raspodelu predloga sa standardnim devijacijama 0,3 i 0,9 i 20000 iteracija. Najuži intervali su dobijeni za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija.

Tabela 5.7: Rezultati analize lanaca Markova za koeficijent varijacije neto dobitka ako je apriorna raspodela $B_4(5;0,7;3;6)$

Apriorna raspodela $B_4(5;0,7;3;6)$						
Broj iteracija/ Burn in	Standardne devijacije raspodele predloga	Veličina uzorka	Prosek marginalne aposteriorne raspodele	ESS	Širina intervala sa jednakim repovima	Širina intervala sa najvećom aposteriornom gustinom
10000/ 1000	0,3 i 0,6	10	4,62	117,02	6,518	5,957
		20	4,26	130,63	6,157	6,072
		50	3,97	66,66	7,985	7,985
		80	4,45	110,21	8,428	7,976
20000/ 5000	0,3 i 0,6	10	5,11	120,24	5,839	5,626
		20	4,33	106,06	6,302	6,117
		50	4,20	101,81	7,913	7,413
		80	4,03	104,09	8,228	7,727
10000/ 1000	0,3 i 0,9	10	4,94	58,70	6,772	6,618
		20	4,64	84,61	6,731	6,452
		50	4,74	87,09	8,251	7,617
		80	4,36	188,41	7,686	7,657
20000/ 5000	0,3 i 0,9	10	3,92	63,79	6,089	5,825
		20	4,61	110,01	6,635	6,591
		50	4,23	66,12	7,797	7,605
		80	4,62	126,55	8,387	8,387

Izvor: Izlaz iz programskog jezika R

Za obe razmatrane apriorne raspodele, zaključuje se da su u većini slučajeva interval sa jednakim repovima i interval sa najvećom aposteriornom gustinom najuži za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija.

Kada je reč o klasičnim intervalima poverenja za koeficijent varijacije, dobijeni rezultati su prikazani u Tabeli 5.8. Primećuje se da je za uzorke veličine 10 najuži RSS BCa interval, dok je za

uzorke veličine 20 i 50 najuži RSS percentil interval. Za uzorke veličine 80 kao najbolji izbor se pokazao t interval.

Tabela 5.8: Širine 95% dvostranog intervala poverenja za koeficijent varijacije neto dobitka

Veličina uzorka		10	20	50	80
t		1,208	0,872	1,346	0,834
Percentil		1,051	1,048	2,926	2,805
RSS Percentil	$k=2$	0,804	1,974	2,183	1,188
	$k=5$	0,706	0,419	0,506	1,114
<i>Bootstrap-t</i>		1,190	1,709	1,538	1,410
Bca		0,751	1,904	2,543	1,756
RSS Bca	$k=2$	1,189	1,457	2,677	0,941
	$k=5$	0,650	1,127	2,311	0,841
<i>Jackknife</i>		2,422	1,254	1,593	1,896
C&P		4,603	4,353	4,341	2,369
C&P trim 10%		3,594	8,693	1,448	2,083
RSS C&P	$k=2$	2,204	5,982	2,047	3,633
	$k=5$	4,941	2,308	3,907	2,594
RSS C&P trim 10%	$k=2$	4,688	8,617	1,549	4,162
	$k=5$	2,849	2,086	1,587	2,757

Izvor: Izlaz iz programskog jezika R

Za realne podatke se zaključuje da su klasični intervali poverenja uži od Bayes-ovskih intervala. Ovaj zaključak je validan ako se kao apriorne raspodele posmatraju uniformna i beta raspodela.

ZAKLJUČAK

U radu su analizirane različite metode na osnovu kojih su konstruisani intervali poverenja za regresione koeficijente u prostom linearnom regresionom modelu i kvadratnom regresionom modelu, kao i intervali poverenja za apsolutnu i relativnu meru disperzije. Pažnja je bila posvećena konstrukciji intervala poverenja u slučaju kada neke od polaznih pretpostavki nisu ispunjene. Konkretno, razmatrane su situacije u kojima slučajna greška modela nije normalno raspodeljena odnosno kada je reč o merama disperzije predmet analize su bili slučajevi u kojima osnovni skup nije normalno raspodeljen.

Utvrđeno je koje metode obezbeđuju da proporcije simuliranih intervala koji sadrže nepoznati parametar budu najbliže nominalnom nivou pouzdanosti u slučaju neispunjenosti određenih pretpostavki.

Predmet istraživanja u prostom linearnom regresionom modelu bile su proporcije simuliranih intervala (za regresioni koeficijent) koji sadrže nepoznati parametar kada slučajna greška nije normalno raspodeljena. U prvom slučaju se pretpostavilo da slučajna greška sledi eksponencijalnu raspodelu, dok je u drugom slučaju predmet razmatranja bila Weibull-ova raspodela. Izveden je generalni zaključak za obe raspodele. Kada je reč o malim uzorcima, utvrđeno je da su proporcije simuliranih *bootstrap-t* intervala i *bootstrap-t* intervala sa transformacijama T_2 i T_3 bile najbliže nominalnom nivou pouzdanosti. Prilikom razmatranja velikih uzoraka, donet je zaključak da su proporcije simuliranih *bootstrap-t* intervala sa transformacijama T_2 i T_3 dale najbolje rezultate.

U kvadratnom regresionom modelu razmatrane su proporcije simuliranih intervala (za regresione koeficijente) koji sadrže nepoznati parametar kada slučajna greška sledi gama, Weibull-ovu odnosno eksponencijalnu raspodelu, respektivno. Predmet analize su bili nominalni nivoi pouzdanosti od 95% i 90%. Donet je generalni zaključak da je prilikom intervalnog ocenjivanja parametra β_1 najbolje koristiti percentil, *bootstrap-t* i *jackknife* intervale, a za parametar β_2 preporučena je primena *bootstrap-t* intervala sa transformacijom T_2^* i percentil intervala. Sprovedenim analizama je potvrđena hipoteza H_2 .

Pored simulacione studije, u prostom linearnom regresionom modelu i kvadratnom regresionom modelu sprovedena je analiza na podacima u oblasti korporativnih finansija. Predmet razmatranja u prostom linearnom regresionom modelu bile su proporcije simuliranih intervala za regresioni koeficijent zasnovane na podacima o indikatoru verovatnoće bankrotstva i količniku ukupnog duga u 157 kompanija u Srbiji u 2014. godini. Istraživanje je sprovedeno na preduzećima koja ostvaruju najveći nivo poslovne aktivnosti, odnosno na preduzećima čiji su poslovni prihodi u 2013. godini bili iznad 5 milijardi dinara. Utvrđeno je da su za male uzorke proporcije simuliranih *bootstrap-t* intervala bile najbolje među svim proučavanim intervalima, ali ni one nisu blizu nominalnog nivoa pouzdanosti. Kada je reč o velikim uzorcima, T_3 i *bootstrap-t* intervali su obezbedili najbolje rezultate. Dalje, u hipotetičkom primeru su konstruisani intervali poverenja za regresione koeficijente i izračunate njihove širine, koristeći podatke o zbirnim finansijskim indikatorima. Reč je o neto dobitku, broju zaposlenih i troškovima sirovina. Polazeći od pretpostavke da je neto dobitak zavisna promenljiva, a da su troškovi sirovina odnosno broj zaposlenih objašnjavajuće promenljive, respektivno, utvrđeno je da je u oba slučaja najuži interval

poverenja dobijen primenom *bootstrap* metoda u kombinaciji sa transformacijom T_1 . Zatim su u hipotetičkom primeru ispitivane proporcije simuliranih intervala za regresioni koeficijent na osnovu podataka o prosečnoj ceni leta i udaljenosti između gradova. Ustanovljeno je da su za male uzorke proporcije simuliranih *bootstrap-t* intervala bile najbliže nominalnom nivou pouzdanosti, dok je za velike uzorke T_3 interval obezbedio najbolje rezultate. Ovim su potvrđene hipoteze H_1 i H_5 .

Koristeći podatke o o neto dobitku i ukupnoj imovini u 105 opština u Srbiji u 2017. godini ispitivane su proporcije simuliranih intervala za regresione koeficijente u kvadratnom regresionom modelu. Analizom se došlo do sledećih zaključaka. Kada je reč o parametru β_1 , za uzorke veličine 10, proporcija simuliranih percentil intervala bila je najbliža nominalnom nivou pouzdanosti. Za uzorke veličine 20, *jackknife* interval je obezbedio najbolje rezultate, dok su za velike uzorke najbolji rezultati dobijeni primenom *bootstrap-t* intervala. Utvrđeno je da je za parametar β_2 , kada je reč o malim uzorcima, *jackknife* interval obezbedio najbolje rezultate, a da su za velike uzorke proporcije simuliranih T_2^* intervala bile najbliže nominalnom nivou pouzdanosti. Dalje, u hipotetičkom primeru su analizirane proporcije simuliranih intervala za regresione koeficijente bazirane na podacima o indikatoru verovatnoće bankrotstva i količniku ukupnog duga. Analizom rezultata su doneti sledeći zaključci koji se odnose na parametar β_1 . Za male uzorke proporcija simuliranih BCa intervala bila je najbliža nominalnom nivou pouzdanosti, dok je za uzorke veličine 20 utvrđeno da je *jackknife* interval obezbedio najbolje rezultate. Kada je reč o velikim uzorcima, ustanovljeno je da su proporcije simuliranih *jackknife* i percentil intervala bile najbolje. Za parametar β_2 je utvrđeno da je za uzorke veličine 10 najbolje rezultate obezbedio BCa interval, dok je za uzorke veličine 20 *jackknife* interval bio najbolji izbor. U slučaju uzoraka veličine 50 najbolji rezultati su ostvareni primenom intervala T_2^* , a za uzorke veličine 80 korišćenjem percentil intervala. Ovim su potvrđene hipoteze H_2 i H_5 .

U radu su analizirane proporcije simuliranih intervala za varijansu. Prvo, pod pretpostavkom da osnovni skup sledi gama, log-normalnu odnosno eksponencijalnu raspodelu, respektivno, razmatrane su proporcije simuliranih intervala za varijansu osnovnog skupa koristeći prosto slučajno uzorkovanje i rangirano uzorkovanje. Na osnovu dobijenih rezultata utvrđeno je da su za sve ispitivane raspodele, u većini slučajeva, najbolje proporcije simuliranih intervala za varijansu dobijene primenom rangiranog uzorkovanja. U drugom slučaju, polazeći od pretpostavke da osnovni skup sledi *skew*-normalnu, *skew*-Student-ovu, *skew*-Laplace-ovu odnosno *skew* generalizovanu t raspodelu izvršeno je poređenje proporcija simuliranih intervala za varijansu osnovnog skupa koristeći prosto slučajno uzorkovanje i rangirano uzorkovanje. Analizom rezultata donet je generalni zaključak za sve razmatrane raspodele. Utvrđeno je da su, u većini slučajeva, proporcije simuliranih intervala za varijansu zasnovanih na rangiranom uzorkovanju najbliže nominalnom nivou pouzdanosti. Sprovedenim analizama je potvrđena hipoteza H_3 .

Dalje, ispitivane su proporcije simuliranih intervala za koeficijent varijacije osnovnog skupa, koji predstavlja relativnu meru disperzije. U svrhe istraživanja pretpostavilo se da osnovni skup sledi gama, Weibull-ovu, log-normalnu odnosno *skew*-normalnu raspodelu, respektivno. Analizom rezultata je utvrđeno da su za sve razmatrane raspodele, u većini slučajeva, proporcije simuliranih intervala za koeficijent varijacije zasnovanih na odsečnoj sredini najbliže nominalnom nivou pouzdanosti. Takođe, ispitivanjem širina intervala poverenja za koeficijent varijacije ustanovljeno je da intervali poverenja zasnovani na odsečnoj sredini imaju približno istu širinu kao i oni koji nisu zasnovani na odsečnoj sredini. Ovim je potvrđena hipoteza H_4 .

Pored simulacione studije, analiza je sprovedena i na podacima u oblasti korporativnih finansija. Koristeći podatke o meri sistematskog rizika ispitivane su proporcije simuliranih intervala za varijansu. Utvrđeno je da su najbolji rezultati za uzorke veličine 15 dobijeni primenom RSS χ^2 intervala, dok je za uzorke veličine 20 najbolji izbor bio RSS ADF interval. Kada je reč o uzorcima veličine 50 primenom RSS Steve-ovog intervala dobijeni rezultat je najbliži nominalnom nivou pouzdanosti, a za uzorke veličine 80 se kao najbolji pokazao ADF interval. Ova analiza je pokazala da su, u većini slučajeva, proporcije simuliranih intervala za varijansu zasnovanih na rangiranom uzorkovanju bile bliže nominalnom nivou pouzdanosti u odnosu na proporcije simuliranih intervala zasnovanih na prostom slučajnom uzorkovanju. Dalje, na osnovu podataka o dividendama razmatrane su proporcije simuliranih intervala za varijansu. Za male uzorke i za uzorke veličine 50 RSS Bonett-ov interval je obezbedio najbolje rezultate, dok je za uzorke veličine 80, proporcija simuliranih Bonett-ovih intervala bila najbliža nominalnom nivou pouzdanosti. Sprovedena analiza dovela je do zaključka da su, u većini slučajeva, proporcije simuliranih intervala za varijansu zasnovanih na rangiranom uzorkovanju bliže nominalnom nivou pouzdanosti u odnosu na proporcije simuliranih intervala zasnovanih na prostom slučajnom uzorkovanju. Predmet analize u jednom hipotetičkom primeru bile su i proporcije simuliranih intervala za varijansu zasnovane na podacima koji se odnose na štete nastale usled požara u Danskoj. U slučaju malih uzoraka i uzoraka veličine 50 RSS Bonett-ov interval obezbedio je rezultate najbliže nominalnom nivou pouzdanosti, dok je za uzorke veličine 80 proporcija simuliranih Bonett-ovih intervala bila najbliža nominalnom nivou pouzdanosti. Generalni zaključak je da su, u većini slučajeva, proporcije simuliranih intervala za varijansu zasnovanih na rangiranom uzorkovanju bile bliže nominalnom nivou pouzdanosti u odnosu na proporcije simuliranih intervala zasnovanih na prostom slučajnom uzorkovanju. Ovim su potvrđene hipoteze H_3 i H_5 .

Koristeći podatke o meri sistematskog rizika predmet ispitivanja bile su proporcije simuliranih intervala za koeficijent varijacije. Za uzorke veličine 10 proporcija simuliranih RSS Miller-ovih intervala zasnovanih na odsečnoj sredini bila je najbliža nominalnom nivou pouzdanosti, dok je za uzorke veličine 20 RSS C&P interval zasnovan na odsečnoj sredini obezbedio najbolji rezultat. U slučaju uzoraka veličine 40, najbolji rezultat ostvaren je primenom RSS Miller-ovog intervala zasnovanog na odsečnoj sredini, dok je za uzorke veličine 80 upotreba C&P i RSS Miller-ovog intervala koji su zasnovani na odsečnoj sredini dala najbolje rezultate. Sprovedena analiza je pokazala da su, u većini slučajeva, proporcije simuliranih intervala za koeficijent varijacije zasnovanih na odsečnoj sredini koristeći rangirano uzorkovanje bile najbliže nominalnom nivou pouzdanosti. Na razmatranim podacima ispitane su i prosečne širine intervala poverenja, pri čemu je zaključeno da intervali poverenja zasnovani na odsečnoj sredini pokazuju približno istu širinu kao i oni koji nisu zasnovani na odsečnoj sredini. Dalje, na osnovu podataka o dividendama određivane su proporcije simuliranih intervala za koeficijent varijacije. Utvrđeno je da je za uzorke veličine 10 RSS Miller-ov interval zasnovan na odsečnoj sredini obezbedio najbolje rezultate, dok je za uzorke veličine 20 najbolji izbor bio RSS C&P interval zasnovan na odsečnoj sredini. Za uzorke veličine 40, najbolji rezultati ostvareni su primenom RSS Miller-ovog intervala zasnovanog na odsečnoj sredini, dok je Miller-ov interval zasnovan na odsečnoj sredini dao rezultate najbliže nominalnom nivou pouzdanosti za uzorke veličine 80. Izveden je zaključak da su, u većini slučajeva, proporcije simuliranih intervala za koeficijent varijacije zasnovanih na odsečnoj sredini koristeći rangirano uzorkovanje najbliže nominalnom nivou pouzdanosti. Sprovedenim analizama su potvrđene hipoteze H_4 i H_5 .

Na kraju, u jednom hipotetičkom primeru objašnjena je konstrukcija dvodimenzionalnih lanaca Markova za sredinu i koeficijent varijacije log-normalne raspodele primenom Metropolis-Hastings-ovog algoritma i na osnovu dobijenih rezultata su ispitivane širine intervala sa jednakim repovima i intervala sa najvećom aposteriornom gustinom za koeficijent varijacije. Dalje, izvršeno

je poređenje širina nekih klasičnih intervala poverenja za koeficijent varijacije log-normalne raspodele sa širinama intervala sa jednakim repovima i intervala sa najvećom aposteriornom gustinom.

Kao apriorne raspodele za koeficijent varijacije log-normalne raspodele razmatrane su dvoparameterska beta, uniformna odnosno četvoroparameterska beta raspodela, respektivno i izveden je zaključak da je, u većini slučajeva, interval sa jednakim repovima najuži za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 20000 iteracija, dok je interval sa najvećom aposteriornom gustinom najuži za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija. Zatim, izračunate su širine nekih klasičnih intervala poverenja za koeficijent varijacije, pri čemu je utvrđeno da je za male uzorke i uzorke veličine 50 najuži RSS BCa interval, dok je za uzorke veličine 80 najuži *bootstrap-t* interval. Donet je zaključak da su klasični intervali poverenja uži od Bayes-ovskih intervala za razmatrane apriorne raspodele.

Pored analize na simuliranim podacima, sprovedena je i analiza na podacima o neto dobitku. Koristeći uniformnu odnosno četvoroparametersku beta raspodelu kao apriorne raspodele za koeficijent varijacije log-normalne raspodele, izveden je zaključak da su, u većini slučajeva, interval sa jednakim repovima i interval sa najvećom aposteriornom gustinom najuži za raspodelu predloga sa standardnim devijacijama 0,3 i 0,6 i 10000 iteracija. Kada je reč o klasičnim intervalima poverenja za koeficijent varijacije, utvrđeno je da je za uzorke veličine 10 najuži RSS BCa interval, dok je za uzorke veličine 20 i 50 najuži RSS percentil interval. Kao najbolji izbor za uzorke veličine 80 pokazao se *t* interval. Na osnovu svih rezultata, donet je zaključak da su klasični intervali poverenja uži od Bayes-ovskih intervala za razmatrane apriorne raspodele.

Moguće je izdvojiti sledeće doprinose doktorske disertacije:

1. U radu je predložen novi interval poverenja za regresioni koeficijent u prostom linearnom regresionom modelu koji je adekvatan kada slučajna greška ne sledi normalnu raspodelu. Najpre, izveden je Edgeworth-ov razvoj raspodele *t*-statistike do reda $n^{-1/2}$, na osnovu kog je predložena transformacija *t*-statistike, T_3 . Ova transformacija je upotrebljena za konstruisanje *bootstrap-t* intervala poverenja za regresioni koeficijent.
2. Pod pretpostavkom da slučajna greška nema normalnu raspodelu, koristeći podatke koji se odnose na indikator verovatnoće bankrotstva i količnik ukupnog duga utvrđeno je da su proporcije simuliranih intervala poverenja za regresioni koeficijent u prostom linearnom regresionom modelu koji je zasnovan na novoj transformaciji *t* statistike, T_3 , za velike uzorke, bliže nominalnom nivou pouzdanosti u odnosu na proporcije ostalih proučavanih intervala.
3. U disertaciji je predložen novi interval poverenja za regresioni koeficijent β_2 u kvadratnom regresionom modelu koji je pogodan za primenu kada slučajna greška ne sledi normalnu raspodelu. Prvo, izveden je Edgeworth-ov razvoj raspodele *t*-statistike do reda n^{-1} , koji predstavlja osnovu za transformaciju *t*-statistike, T_2^* . Zatim je dobijena transformacija upotrebljena za konstruisanje *bootstrap-t* intervala poverenja za regresioni koeficijent β_2 .

4. Pod pretpostavkom da slučajna greška sledi gama, Weibull-ovu, odnosno eksponencijalnu raspodelu, utvrđeno je da su proporcije simuliranih intervala poverenja za regresioni koeficijent β_2 u kvadratnom regresionom modelu koji je zasnovan na novoj transformaciji t statistike, T_2^* , za velike uzorke, u većini razmatranih slučajeva, bliže nominalnom nivou pouzdanosti u odnosu na proporcije ostalih proučavanih intervala.
5. Pod pretpostavkom da slučajna greška nema normalnu raspodelu, koristeći podatke koji se odnose na neto dobitak i ukupnu imovinu, kao i podatke o indikatoru verovatnoće bankrotstva i količniku ukupnog duga, utvrđeno je da su proporcije simuliranih intervala poverenja za regresioni koeficijent β_2 u kvadratnom regresionom modelu koji je zasnovan na novoj transformaciji t statistike, T_2^* , za velike uzorke, bliže nominalnom nivou pouzdanosti u odnosu na proporcije ostalih proučavanih intervala.
6. Predložene su transformacije intervala poverenja za koeficijent varijacije osnovnog skupa koje su pogodne za primenu ukoliko osnovni skup ne sledi normalnu raspodelu. Reč je o transformaciji postojećih intervala poverenja zasnovanoj na odsečnoj sredini i *bootstrap* transformaciji. U prvom slučaju, upotrebljena je odsečena sredina umesto aritmetičke sredine da bi se izračunala varijansa uzorka, dok su se u drugom slučaju koristili percentili *bootstrap* raspodele umesto tradicionalnih kritičnih vrednosti.
7. Pod pretpostavkom da slučajna greška sledi gama, Weibull-ovu, log-normalnu, odnosno *skew-normalnu* raspodelu, utvrđeno je da su proporcije simuliranih intervala poverenja za koeficijent varijacije osnovnog skupa, zasnovanih na transformaciji sa odsečenom sredinom, u većini razmatranih slučajeva, bliže nominalnom nivou pouzdanosti u odnosu na proporcije ostalih ispitivanih intervala.
8. Pod pretpostavkom da mera sistematskog rizika, odnosno dividenda ne slede normalnu raspodelu, utvrđeno je da su proporcije simuliranih intervala poverenja za koeficijent varijacije osnovnog skupa, zasnovanih na transformaciji sa odsečenom sredinom, u većini razmatranih slučajeva, bliže nominalnom nivou pouzdanosti u odnosu na proporcije ostalih ispitivanih intervala.
9. Originalni kodovi koji su napisani u programskom jeziku R , pored primene u oblasti korporativnih finansija, mogu se upotrebiti u različitim eksperimentalnim situacijama.
10. Na osnovu rezultata simulacionih studija, u zavisnosti od stepena narušenosti polaznih pretpostavki, predložene su metode ocenjivanja koje najviše odgovaraju datom problemu. Takođe, analizom dobijenih rezultata u oblasti korporativnih finansija otkriveni su postojeći problemi prilikom intervalnog ocenjivanja različitih parametara, a zatim predložene mogućnosti njihovog rešavanja.

Pored navedenih doprinosa, potrebno je osvrnuti se i na ograničenja istraživanja. Ona se u najvećoj meri odnose na upotrebu razmatranih raspodela i podataka. Bitno je istaći da je izvršen proizvodjan izbor asimetričnih raspodela, kao i podataka koji zadovoljavaju pretpostavke samog istraživanja, odnosno bilo je moguće koristiti bilo koju asimetričnu raspodelu, odnosno podatke koji ispunjavaju zahtevane pretpostavke. Takođe, regresioni modeli su razmatrani pod pretpostavkom da se bira slučajan uzorak.

U daljem istraživanju bi bilo interesantno primeniti razmatrane metode na drugim raspodelama i podacima iz oblasti korporativnih finansija, ali i u drugim oblastima ekonomije. Jedna od mogućnosti odnosi se na analizu drugih mera disperzije, kao i regresionih modela višeg reda. Zatim, pažnja bi se mogla posvetiti Bayes-ovskom *bootstrap* metodu i njegovoj primeni prilikom konstrukcije intervala poverenja za različite parametre. Jedna od ideja bi bila i proučavanje Bayes-ovskog ocenjivanja parametara primenom rangiranog uzorkovanja. Dalje, bilo bi interesantno analizirati nebalansirano rangirano uzorkovanje (engl. unbalanced ranked set sampling), ekstremno rangirano uzorkovanje (engl. extreme ranked set sampling), rangirano uzorkovanje zasnovano na medijani (engl. median ranked set sampling), rangirano uzorkovanje zasnovano na kvartilima (engl. quartile ranked set sampling) i uporediti dobijene rezultate sa rezultatima dobijenim primenom metoda opisanih u ovom radu.

LITERATURA

A

- Abdi, H. (2010). Coefficient of Variation. In: N. J. Salkind (Ed.), *Encyclopedia of Research Design* (pp. 169-171). Thousand Oaks, California: SAGE.
- Abu-Shawiesh, M. O. A., Banik, S., & Golam Kibria, B. M. (2011). A simulation study on some confidence intervals for the population standard deviation. *SORT*, 35(2), 83-102.
- Adrover, J., & Salibian-Barrera, M. (2010). Globally robust confidence intervals for simple linear regression. *Computational Statistics and Data Analysis*, 54(12), 2899-2913.
- Albatineh, A. N., Boubakari, I., & Golam Kibria, B. M. (2017). New confidence interval estimator of the signal-to-noise ratio based on asymptotic sampling distribution. *Communications in Statistics – Theory and methods*, 46(2), 574-590.
- Albatineh, A. N., Golam Kibria, B. M., Wilcox, M. L., & Zogheib, B. (2014). Confidence interval estimation for the population coefficient of variation using ranked set sampling: a simulation study. *Journal of Applied Statistics*, 41(4), 733-751.
- Albert, J. (2007). *Bayesian Computation with R*. New York: Springer-Verlag.
- Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 589-609.
- Altman, E. (1993). *Corporate Financial Distress and Bankruptcy* (2nded.). New York: John Wiley & Sons.
- Altman, E. (2005). An emerging market credit scoring system for corporate bonds. *Emerging Markets Review*, 6, 311-323.
- Altman, E., Hartzell, J., & Peck, M. (1995). *Corporate Financial Distress and Bankruptcy* (3rded.). New York: John Wiley & Sons.
- Amiri, S., Jafari Jozani, M., & Modarres, R. (2013). Resampling Unbalanced Ranked Set Samples With Applications in Testing Hypothesis About the Population Mean. *Journal of Agricultural, Biological and Environmental Statistics*, 19(1), 1-17.
- Antille, A., & Milasevic, P. (1988). A confidence interval for the slope of a truncated regression. *Probability Theory and Related Fields*, 78(1), 63-72.
- Antleman, G. (1996). *Elementary Bayesian Statistics*. Cheltenham: Edward Elgar Publishing.
- Aragon, G. A. (1989). *Financial Management*. Boston: Allyn & Bacon.
- Athreya, K. B. (1986). Bootstrap of the mean in the infinite variance case. *The Annals of Statistics*, 14, 724-731.

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12, 171–178.

Azzalini, A., & Capitanio, A. (2003). Distributions Generated by Perturbation of Symmetry with Emphasis on a Multivariate Skew t-Distribution. *Journal of the Royal Statistical Society, Series B*, 65, 367-389.

B

Banik, S., Albatineh, A. N., Abu-Shawiesh, M. O. A., & Golam Kibria, B. M. (2014). Estimating the population standard deviation with confidence interval: a simulation study under skewed and symmetric conditions. *Journal of Biometrics and Biostatistics*, 5(2), 1-9.

Barbe, P., & Bertail, P. (1995). *The Weighted Bootstrap*. New York: Springer-Verlag.

Barham, A. M., & Jeyaratnam, S. (1999). Robust confidence interval for the variance. *Journal of Statistical Computation and Simulation*, 62(3), 189-205.

Bartolucci, F., & Scaccia, L. (2005). The use of mixtures for dealing with non-normal regression errors. *Computational Statistics and Data Analysis*, 48(4), 821-834.

Bayarri, M. J., & Berger, J. O. (2004). The interplay of Bayesian and frequentist analysis. *Statistical Science*, 19(1), 58–80.

Berger, J. (1985). *Statistical Decision Theory and Bayesian Analysis*. New York: Springer-Verlag.

Bernard, G. A. (1963). Contribution to discussion. *Journal of the Royal Statistical Society, Series B*, 25, 294.

Besley, S., & Brigham, E. (2018). *Corporate Finance*. Boston: Cengage.

Bickel, P. J., & Freedman, D. (1981). Some asymptotic theory for the bootstrap. *The Annals of Statistics*, 9, 1196- 1217.

Binomial Distribution. (1999). Preuzeto 08.01.2021., sa sajta <https://mathworld.wolfram.com/BinomialDistribution.html>

Bodie, Z., Kane, A., & Marcus, A. (2009). *Osnovi investicija*, prevod. Beograd: Data Status.

Bolstad, W. M. (2007). *Introduction to Bayesian Statistics* (2nd ed.). New York: John Wiley & Sons.

Bonett, D. G. (2006). Approximate confidence interval for standard deviation of nonnormal distributions. *Computational Statistics and Data Analysis*, 50, 775-782.

Box, G. E. P., & Tiao, G. C. (1973). *Bayesian inference in statistical analysis*. Reading: Addison-Wesley.

Brealey, R. A., Myers, S. C., & Allen, F. (2011). *Principles of Corporate Finance* (10th ed.). New York: McGraw-Hill/Irwin.

Brooks, S. P. & Roberts, G. O. (1998). Assessing convergence of Markov chain Monte Carlo algorithms. *Statistics and Computing*, 8, 319-35.

Buchholz, A., Holländer, N., & Sauerbrei, W. (2008). On properties of predictors derived with a two-step bootstrap model averaging approach - A simulation study in the linear regression model. *Computational Statistics and Data Analysis*, 52(5), 2778-2793.

Buckland, S. T. (1985). Calculation of Monte Carlo confidence intervals. *Journal of the Royal Statistical Society, Series B*, 34, 296-301.

Burch, B. (2014). Estimating kurtosis and confidence intervals for the variance under nonnormality. *Journal of Statistical Computation and Simulation*, 84(12), 2710-2720.

C

Casella, G., & Berger, R. L. (2002). *Statistical Inference* (2nd ed.). Pacific Grove: Thomson Learning Academic Resource Center.

Carlin, B., & Louis, T. (2000). *Bayes and Empirical Bayes Methods for Data Analysis*. Boca Rotan: Chapman & Hall/CRC.

Carpenter, J. (1999). Test-inversion bootstrap confidence intervals. *Journal of the Royal Statistical Society, Series B*, 61, 159-172.

Carpenter, J., & Bithell, J. (2000). Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Statistics in Medicine*, 19(9), 1141-1264.

Chib, S., & Greenberg, E. (1995). Understanding the Metropolis-Hastings Algorithm. *The American Statistician*, 49(4), 327-335.

Chen, L. (1995). Testing the mean of skewed distributions. *Journal of the American Statistical Association*, 90, 767-770.

Chen, Z. (2007). Ranked set sampling: its essence and some new applications. *Environmental and Ecological Statistics*, 14, 355-363.

Chernick, M. R. (2008). *Bootstrap Methods: A Guide for Practitioners and Researchers* (2nd ed.). Hoboken: John Wiley & Sons.

Chernick, M. R., & LaBudde, R. A. (2011). *An Introduction to Bootstrap Methods with Applications to R*. Hoboken: John Wiley & Sons.

Cowles, M. K. (2013). *Applied Bayesian Statistics: With R and OpenBUGS Examples*. New York: Springer.

Cowles, M. K., & Carlin, B.P. (1996). Markov chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association*, 91, 883-904.

Curto, J. D., & Pinto, J. C. (2009). The coefficient of variation asymptotic distribution in the case of non-iid random variables. *Journal of Applied Statistics*, 36(1), 21-32.

Ćojbašić, V., & Lončar, D. (2011). One-sided confidence intervals for population variances of skewed distributions. *Journal of Statistical Planning and Inference*, 141(5), 1667-1672.

Ćojbašić Rajić, V., & Stanojević, J. (2013). Confidence intervals for the ratio of two variances. *Journal of Applied Statistics*, 40(10), 2181-2187.

Ćojbašić, V., & Tomović, A. (2007). Nonparametric confidence intervals for population variance of one sample and the difference of variances of two samples. *Computational Statistics and Data Analysis*, 51(12), 5562-5578.

D

Damodaran, A. (2015). *Applied Corporate Finance* (4th ed.). Hoboken: John Wiley & Sons.

Davis, C. (2015). *The Skewed Generalized T Distribution Tree Package Vignette*. Preuzeto sa sajta <https://cran.r-project.org/web/packages/sgt/vignettes/sgt.pdf>

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap Methods and their Application*. Cambridge: Cambridge University Press.

DiCiccio, T. J., & Efron, B. (1996). Bootstrap Confidence Intervals. *Statistical Science*, 11(3), 189-228.

DiCiccio, T. J. & Romano, J. P. (1988). A review of bootstrap confidence intervals (with discussions). *Journal of the Royal Statistical Society, Series B*, 50, 538-554.

Dell, T. R., & Clutter, J. L. (1972). Ranked set sampling theory with order statistics background. *Biometrics*, 28(2), 545-555.

Douglas, S. (1996). Bootstrap confidence intervals in a switching regressions model. *Economics Letters*, 53(1), 7-15.

Drikvandi, R., Modarres, R., & Jalilian, A. H. (2011). A bootstrap test for symmetry based on ranked set samples. *Computational Statistics and Data Analysis*, 55(41), 1807-1814.

Đorić, D., Mališić, J., Jevremović, V., & Nikolić-Đorić, E. (2007). *Atlas raspodela*. Beograd: Univerzitet u Beogradu, Građevinski fakultet.

E

Edgeworth, F. Y. (1898). On the Representation of Statistics by Mathematical Formulae (Part I.). *Journal of the Royal Statistical Society*, 61(4), 670-700.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7, 1-26.

Efron, B. (1981a). Non-parametric standard errors and confidence intervals (with discussion). *Canadian Journal of Statistics*, 9, 139-172.

Efron, B. (1981b). Non-parametric estimates of standard error: the jackknife, the bootstrap, and other methods. *Biometrika*, 68, 589-599.

Efron, B. (1982). *The Jackknife, the Bootstrap and other Resampling Plans*. Philadelphia: Society for Industrial and Applied Mathematics.

Efron, B. (1987). Better bootstrap confidence intervals (with discussions). *Journal of the American Statistical Association*, 82, 171-200.

Efron, B. (1992). Jackknife-after-bootstrap standard errors and influences functions (with discussions). *Journal of the Royal Statistical Society, Series B*, 54, 83-127.

Efron, B., & Stein, C. (1981). The jackknife estimate of variance. *The Annals of Statistics*, 9, 586-596.

Efron, B., & Tibshirani, R. J. (1986). Bootstrap measures for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, 1, 54-77.

Efron, B., & Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. New York: Chapman & Hall/CRC.

Eling, M. (2012). Fitting Insurance Claims to Skewed Distributions: Are the Skew-Normal and Skew-Student Good Models? *Insurance: Mathematics and Economics*, 51, 239-248.

F

Ford, P. (2018). *MCMC Algorithms*. Preuzeto sa sajta https://www.casact.org/sites/default/files/old/studynotes_masii_mcmc_algorithms_v0.5.pdf

G

Galindo, C. D., Liang, H., Kauermann, G., & Carroll, R. J. (2001). Bootstrap confidence intervals for local likelihood, local estimating equations and varying coefficient models. *Statistica Sinica*, 11, 121-134.

Gamerman, D., & Lopes, H. F. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference* (2nd ed.). Boca Raton: Chapman & Hall/CRC.

Ganeslingam, S., & Ganesh, S. (2006). Ranked set sampling versus simple random sampling in the estimation of the mean and the ratio. *Journal of Statistics and Management Systems*, 9(2), 459-472.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. (2014). *Bayesian Data Analysis* (3rd ed.). Boca Raton: CRC Press.

Gelman, A., & Rubin, D. (1992). Inference from Iterative Simulation using Multiple Sequences. *Statistical Science*, 7, 457-472.

George, M. R. W., Yang, N., Horn, M. L. V., Smith, J., Jaki, T., Feaster, D., ... Howe, G. (2013). Using regression mixture models with non-normal data: Examining an ordered polytomous approach. *Journal of Statistical Computation and Simulation*, 83(4), 759-772.

- Geyer, C. J. (1992). Practical Markov chain Monte Carlo (with discussion). *Statistical Science*, 7, 473-511.
- Ghosh, S., Chatterjee, A., & Balakrishnan, N. (2017). Nonparametric confidence intervals for ranked set samples. *Computational Statistics*, 32, 1689-1725.
- Givens, G., & Hoeting, J. (2005). *Computational Statistics*, Hoboken: John Wiley & Sons.
- Good, P. I. (2006). *Resampling Methods: A Practical Guide to Data Analysis* (3rd ed.). Boston: Birkhäuser.
- Green, P. J. (2001). A primer on Markov chain Monte Carlo. In O. E. Barndorff-Nielsen, D. R. Cox & C. Klüppelberg (Eds.), *Complex Stochastic Systems* (pp. 1–62). Boca Raton: Chapman & Hall/CRC.

H

- Hall, P. (1988a). On Symmetric Bootstrap Confidence Intervals. *Journal of the Royal Statistical Society, Series B*, 50(1), 35-45.
- Hall, P. (1988b). Theoretical comparison of bootstrap confidence intervals (with discussion). *The Annals of Statistics*, 16, 927-953.
- Hall, P. (1992a). On the removal of skewness by transformation. *Journal of the Royal Statistical Society, Series B*, 54(1), 221-228.
- Hall, P. (1992b). *The Bootstrap and Edgeworth Expansion*. New York: Springer.
- Hall, M., & Mayo, M. S. (2008). Bootstrap confidence intervals and coverage probabilities of regression parameter estimates using trimmed elemental estimation. *Journal of Modern Applied Statistical Methods*, 7(2), 514-525.
- Han, Y., Liu, W., Bretz, F., Wan, F., & Yang, P. (2016). Statistical calibration and exact one-sided simultaneous tolerance intervals for polynomial regression. *Journal of Statistical Planning and Inference*, 168, 90-96.
- Hartigan, J. A. (1971). Error analysis by replaced samples. *Journal of the Royal Statistical Society, Series B*, 33, 98-110.
- Hartigan, J. A. (1975). Necessary and sufficient conditions for asymptotic joint normality of a statistic and its subsample values. *The Annals of Statistics*, 3, 573-580.
- Hastings, W. (1970). Monte Carlo sampling methods using Markov chains and their application. *Biometrika*, 57, 97–109.
- Haq, A., Brown, J., Moltchanova, E., & Al-Omari, A.-I. (2013). Partial Ranked Set Sampling Design. *Environmetrics*, 24(3), 201-207.
- Held, L., & Bové D. S. (2014). *Applied Statistical Inference*. Berlin: Springer-Verlag.

Hillier, D., Grinblatt, M., & Titman, S. (2012). *Financial Markets and Corporate Strategy* (2nd ed.). London: McGraw-Hill Higher Education.

Hinkley, D. V. (1988). Bootstrap methods (with discussions). *Journal of the Royal Statistical Society, Series B*, 50, 321-337.

Hirschberg, J. G., & Lye, J. N. (2004). *Inferences for the extremum of quadratic regression models*. Preuzeto sa sajta https://minerva-access.unimelb.edu.au/bitstream/handle/11343/34385/66908_00002363_01_906.pdf?sequence=1&isAllowed=y

Hoaglin, D. C., (1985). Summarizing shape numerically: the g- and h-distributions. In D. C. Hoaglin, F. Mosteller & J. W. Tukey (Eds.), *Exploring Data Tables, Trends, and Shapes* (pp. 461-513). New York: John Wiley & Sons.

Hope, A. C. A. (1968). A simple Monte Carlo test procedure. *Journal of the Royal Statistical Society, Series B*, 30, 582-598.

Houdré, C. (1997). The iterated jackknife estimate of variance. *Statistics and Probability Letters*, 35(230), 197-201.

Horváth, L., & Reeder, R. (2013). A test of significance in functional quadratic regression. *Bernoulli*, 19(5A), 2120-2151.

Hummel, R., Banga, S., & Hettmansperger, T. P. (2005). *Better confidence intervals for the variance in a random sample*. Preuzeto sa sajta <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.482.4263&rep=rep1&type=pdf> .

Husby, C. E., Stasny, E. A., & Wolfe, D. A. (2005). An application of ranked set sampling for mean and median estimation using USDA crop production data. *Journal of Agricultural, Biological and Environmental Statistics*, 10(3), 354-373.

I

Ivković, I., & Rajić, V. (2018). Confidence intervals for the population standard deviation: simple random sampling vs. ranked set sampling. In N. Žarkić Joksimović & S. Marinković (Eds.), *Doing Business in the Digital Age: Challenges, Approaches and Solutions* (pp. 27-34). Beograd: Univerzitet u Beogradu, Fakultet organizacionih nauka.

Ivković, I., & Rajić, V. (2019). Better confidence intervals for the population coefficient of variation. *Communications in Statistics – Simulation and Computation*. Doi: 10.1080/03610918.2019.1642482.

Ivković, I., Rajić, V., & Stanojević, J. (2020). Coverage probabilities of confidence intervals for the slope parameter of linear regression model when the error term is not normally distributed. *Communications in Statistics – Simulation and Computation*, 49(1), 147-158.

J

Jackman, S. (2009). *Bayesian Analysis for the Social Sciences*. Chichester: John Wiley & Sons.

Jarque, C., & Bera, A. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, 55(2), 163-172.

Jarušková, D. (2001). Change-point estimator in continuous quadratic regression. *Commentationes Mathematicae Universitatis Carolinae*, 42(4), 741-752.

Johnson, N. (1978). Modified t tests and confidence intervals for asymmetrical populations. *Journal of American Statistical Association*, 73(363), 536-554.

Jones, J. A., & Waller, N. G. (2013). Computing confidence intervals for standardized regression coefficients. *Psychological Methods*, 18(4), 435-453.

K

Kass, R. E., Carlin, B. P., Gelman, A., & Neal, R. (1998). Markov Chain Monte Carlo in Practice: A Roundtable Discussion. *The American Statistician*, 52, 93–100.

Knight, K. (2000). *Mathematical Statistics*. New York: Chapman & Hall/CRC.

Košmelj, K., Blejec, A., & Cedilnik, A. (2005). Interval Estimate for Specific Points in Polynomial Regression. *Journal of Computing and Information Technology*, 13(4), 287-297.

Kruschke, J. K. (2015). *Doing Bayesian Data Analysis: A Tutorial with R, JUGS and STAN* (2nd ed.). London: Elsevier.

Kuč, V. R. (2017). *Analiza strukture finansiranja najvećih preduzeća u Srbiji u periodu velike recesije: preporuke za finansijsko restrukturiranje i poslovnu strategiju* (Doktorska disertacija), Univerzitet u Beogradu, Srbija.

Kuha, J., & Temple, J. (2003). Covariate Measurement Error in Quadratic Regression. *International Statistical Review*, 71(1), 131-150.

L

Lee, P. M. (2004). *Bayesian statistics: an introduction* (3rd ed.). London: Arnold.

Liu, H., & Wasserman, L. (2014). *Statistical Machine Learning*. Preuzeto sa sajta <http://www.stat.cmu.edu/~larry/=sml/Bayes.pdf>.

Looney, S., & Gullledge, T. (1985). Use of the correlation coefficient with normal probability plots. *The American Statistician*, 39(1), 75-79.

Lynch, S. M. (2007). *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. New York: Springer Science & Business Media.

M

Mallows, C. L., & Tukey, J. W. (1982). An overview of techniques of data analysis, emphasizing its exploratory aspects. In J. Tiago de Oliveira, & B. Epstein (Eds.), *Some Advances in Statistics* (pp. 111-172). New York: Academic Press.

- Mamman, E. (1992). *When does bootstrap work. Asymptotic results and simulations*. New York: Springer-Verlag.
- Marin, J. M., & Robert, C. P. (2014). *Bayesian Essentials with R* (2nd ed.). New York: Springer Science & Business Media.
- Maritz, J. S. & Jarrit, R. G. (1978). A note on estimating the variance of the sample median. *Journal of American Statistical Association*, 73, 194-196.
- Marriott, F. H. C. (1979). Barnard's Monte Carlo test: how many simulations? *Journal of the Royal Statistical Society, Series C*, 28(1), 75-77.
- McElreath, R. (2020). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan* (2nd ed.). Boca Raton: CRC Press.
- McIntyre, G. A. (1952). A method for unbiased selective sampling, using ranked sets. *Australian Journal of Agricultural Research*, 59(3), 385-390.
- McKay, A. T. (1932). Distribution of the coefficient of variation and the extended t distribution. *Journal of Royal Statistical Society*, 95, 695-698.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21, 1087-92.
- Miller, E. G. (1991). Asymptotic test statistics for coefficient of variation. *Communication in Statistic – Theory and Methods*, 20, 3351-3363.
- Modarres, R., Hui, T. P., & Zheng, G. (2006). Resampling methods for ranked set samples. *Computational Statistics and Data Analysis*, 51, 1039-1050.
- Multi parameter Metropolis-Hastings. (n.d.). Preuzeto 08.03.2021., sa sajta <https://stats.stackexchange.com/questions/224600/multi-parameter-metropolis-hastings>

N

- Nešić, A. (2010). *MCMC metoda u Bayes-ovskoj statistici* (Master rad, Univerzitet u Beogradu, Srbija). Preuzeto sa sajta <http://milanmerkle.etf.rs/wp-content/uploads/2016/01/Master-rad-Aleksandar-Nesic-septembar-2010.pdf>.
- Niwitpong, S., & Kirdwichai, P. (2008). Adjusted Bonett Confidence Interval for Standard Deviation of Non-normal Distributions. *Thailand Statistician*, 6(1), 1-6.

O

- Ozturk, O. (2018). Ratio estimators based on a ranked set sample in a finite population setting. *Journal of the Korean Statistical Society*, 47, 226-238.

P

Panichkitkosolkul, W. (2009). Improved confidence intervals for a coefficient of variation of a normal distribution. *Thailand Statistician*, 7(2), 193-99.

Pfaff, B., Zivot, E., McNeil, A., & Stephenson, A. (2018). *Evir*. R package version 1.7-4. Preuzeto sa sajta <https://cran.r-project.org/web/packages/evir/evir.pdf>

Piegorsch, W. W. (1986). Confidence Bands for Polynomial Regression With Fixed Intercepts. *Technometrics*, 28(3), 241-246.

Puig, P., & Stephens, M. A. (2007). Goodness of fit tests for the skew-Laplace distribution. *SORT*, 31(1), 45-54.

Q

Quenouille, M. H. (1949). Notes on bias in estimation. *Biometrika*, 43(3-4), 353-360.

R

Raftery, A. E., & Lewis, S. (1996). Implementing MCMC. In W. R. Gilks, S. Richardson & D. J. Spiegelhalter (Eds.), *Markov Chain Monte Carlo in Practice* (pp. 115-30). London: Chapman & Hall/CRC.

Rajan, R., & Zingales, L. (1995). What do we know about capital structure? Some evidence from international data. *The Journal of Finance*, 50, 1421-1460.

Rajić, V. M. (2007). *Statistički metodi ponovljenih uzoraka – analiza i primena u imovinskom osiguranju* (Doktorska disertacija), Univerzitet u Beogradu, Srbija.

Rajić, V., & Ivković, I. (2018). Confidence intervals for the variance in cases of some skewed distributions. In J. Kočović, J. Selimović, B. Boričić, V. Kaščelan & V. Rajić (Eds.), *Quantitative Models in Economics* (pp. 21-37). Beograd: Univerzitet u Beogradu, Ekonomski fakultet.

Robert, C. P., & Casella, G. (2004). *Monte Carlo statistical methods*. New York: Springer.

Robert, C. P., & Casella, G. (2010). *Introducing Monte Carlo methods with R*. New York: Springer.

Roberts, G. O. (1996). Markov chain concepts related to sampling algorithms. In W. R. Gilks, S. Richardson & D. J. Spiegelhalter (Eds.), *Markov Chain Monte Carlo in Practice* (pp. 45-57). London: Chapman & Hall/CRC.

Ross, S. A., Westerfield, R. W., & Jordan B. D. (2003). *Fundamentals of Corporate Finance* (6th ed.). New York: McGraw-Hill/Irwin.

S

Sahu, S. (2000). *Markov Chain Monte Carlo*. Preuzeto sa sajta <http://www.southampton.ac.uk/~sks/utrecht/mcmc.pdf>

Samawi, H. M. (1999). More efficient Monte Carlo methods obtained by using ranked set simulated samples. *Communications in Statistics – Simulation and Computation*, 28(3), 699-713.

Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of American Statistical Association*, 63(324), 1379-1389.

Shacham, M., Brauner, N., & Shore, H. (2007). A new procedure to identify linear and quadratic regression models based on signal-to-noise-ratio indicators. *Mathematical and Computer Modelling*, 46(1-2), 235-250.

Shao, J., & Tu, D. (1995). *The jackknife and bootstrap*. New York: Springer-Verlag.

Sharpe, W. F., Alexander, G. J., & Bailey J. V. (1999). *Investments* (6th ed.). Upper Saddle River: Prentice Hall.

Simon, J. L. (1969). *Basic Research Methods in Social Science*. New York: Random House.

Singh, K. & Xie, M. (n.d.). *Bootstrap: A Statistical Method*. Preuzeto sa sajta <https://stat.rutgers.edu/home/mxie/RCPapers/bootstrap.pdf>

Sobotka, F., Kauermann, G., Waltrup, L. S., & Kneib, T. (2013). On confidence intervals for semiparametric expectile regression. *Statistics and Computing*, 23(2), 135-148.

Soffritti, G., & Galimberti, G. (2011). Multivariate linear regression with non-normal errors: a solution based on mixture models. *Statistics and Computing*, 21(4), 523-536.

Steve, A. (1990). *Mathematical Statistics*. Prentice Hall.

Stokes, S. L. (1980). Estimation of variance using judgement ordered ranked set samples. *Biometrics*, 36(1), 35-42.

T

Terpstra, J. T., & Wang, P. (2008). Confidence intervals for a population proportion based on a ranked set sample. *Journal of Statistical Computation and Simulation*, 78(3-4), 351-366.

Tierney, L. (1994). Markov chain for exploring posterior distributions. *The Annals of Statistics*, 22, 1701–1762.

Todorović, M., & Ivanisević, M. (2018). *Poslovne finansije*. Beograd: Ekonomski fakultet.

Tóth, R., & Somorčík, J. (2017). On a non-parametric confidence interval for the regression slope. *METRON*, 75, 359-369.

Tukey, J. (1958). Bias and confidence in not quite large samples (Abstract). *Annals of Mathematical Statistics*, 29, 614.

V

Veljović, M. (2019). *MCMC*. Preuzeto 10.03.2021., sa sajta <http://www.matf.bg.ac.rs/p/files/95-MCMC.html>

W

Wang, B., Mishra, S. N., Mulekar, M. S., Mishra, N., & Huang, K. (2010). Comparison of bootstrap and generalized bootstrap methods for estimating high quantiles. *Journal of Statistical Planning and Inference*, 140(10), 2926-2935.

White, G., Sondhi, A., & Fried, D. (2003). *The Analysis and Use of Financial Statements* (3rd ed.). USA: John Wiley & Sons.

Wilcox, R. R. (1996). Confidence intervals for the slope of a regression line when the error term has nonconstant variance. *Computational Statistics and Data Analysis*, 22(1), 89-98.

Wilcox, R. R. (2005). Trimmed Means. In B. Everitt & D. Howell (Eds.), *Encyclopedia of Statistics in Behavioral Science* (pp. 2066-2067). Chichester, England: John Wiley & Sons.

Wolfe, D. (2004). Ranked set sampling: an approach to more efficient data collection. *Statistical Science*, 19(4), 636-643.

Wolfe, D. (2012). Ranked Set Sampling: Its Relevance and Impact on Statistical Inference. *International Scholarly Research Notices*, 2012, 1-32.

Wu, C. F. J. (1986). Jackknife, bootstrap and other resampling procedures (with discussions). *The Annals of Statistics*, 14, 1261-1350.

Y

Yao, F., & Müller, H-G. (2010). Functional quadratic regression. *Biometrika*, 97(1), 49-64.

Yazici, B., & Yolacan, S. (2007). A comparison of various tests of normality. *Journal of Statistical Computation and Simulation*, 77(2), 175-183.

Z

Zamanzade, E., & Mahdizadeh, M. (2017). A more efficient proportion estimator in ranked set sampling. *Statistics and Probability Letters*, 129, 28-33.

Zamanzade, E., & Vock, M. 2015. Variance estimation in ranked set sampling using a concomitant variable. *Statistics and Probability Letters*, 105, 1-5.

Zhang, Z., Liu, T., & Zhang, B. (2016). Jackknife empirical likelihood inferences for the population mean with ranked set samples. *Statistical and Probability Letters*, 108, 16-22.

Zhou, X. H., & Dinh, P. (2005). Nonparametric confidence intervals for the one- and two sample problems. *Biostatistics*, 6(2), 187-200.

PRILOZI

PRILOG 1. Podaci o indikatoru verovatnoće bankrotstva

6.822071	4.008055	10.97329	2.82734
1.401936	3.171158	4.032867	4.082203
4.686923	4.489225	3.67138	7.242526
4.922321	3.384795	7.125224	12.63594
6.925256	4.923921	3.905514	5.988496
6.585752	7.466853	3.904487	8.451346
2.900138	3.283633	7.468017	19.44854
4.343164	5.179604	4.441522	2.664383
10.18722	9.091079	7.522512	8.934593
4.141291	0.694993	3.131141	1.874173
4.964773	12.76363	4.295051	4.240902
8.05925	8.751821	6.311922	5.903974
5.964475	4.924529	22.12011	15.93642
7.685569	5.61517	4.25815	3.676555
5.955032	5.102958	3.726535	5.304518
7.899663	2.794789	4.359773	4.650719
6.928892	2.655165	7.305931	5.968181
7.529643	9.673797	6.933083	4.515827
4.806401	5.400419	7.530964	4.209291
3.152536	2.300468	18.38331	11.28008
1.408196	6.221388	5.789225	25.09482
5.289872	5.740707	4.933731	12.21247
5.838003	7.388388	8.255612	17.36164
5.05332	6.000385	6.226973	11.71749
3.962534	6.401195	4.771409	4.106647
2.754649	4.501212	4.119534	4.055197
5.322839	4.705771	9.917681	4.854671
7.992142	7.038981	6.668294	9.519146
4.208065	9.785035	10.80537	9.149429
6.526962	6.574886	4.349372	9.37969
7.967767	6.572211	6.177023	9.727456
7.904372	6.728092	3.815747	9.237723
8.883613	1.253617	7.78437	6.450398
4.651007	11.51734	3.657327	7.979669
5.850638	5.795581	4.80247	17.4203
5.732968	9.003017	3.579165	4.449385
4.526495	7.631092	15.16272	0.360949
6.248859	10.20164	6.230195	
5.245673	5.513713	4.420758	
12.48353	6.785481	3.184979	

PRILOG 2. Podaci o količniku ukupnog duga

0.486007	0.947337	0.351427	0.644829
0.825021	0.980201	0.753743	0.925689
0.391585	0.508326	0.970731	0.46919
0.426895	0.986188	0.529206	0.229262
0.282295	0.591856	0.956087	0.743764
0.351807	0.699914	0.672775	0.506596
0.627457	0.872835	0.256592	0.060396
0.430445	0.796028	0.679795	0.741471
0.246192	0.345726	0.537312	0.285187
0.827116	0.909034	0.909172	0.74367
0.788527	0.163511	0.893793	0.961271
0.206749	0.5446	0.681107	0.446222
0.44954	0.732107	0.052919	0.156828
0.22205	0.75595	0.915159	0.952203
0.270429	0.554791	0.654627	0.591414
0.34764	0.62192	0.421515	0.462843
0.391865	0.509561	0.581632	0.305255
0.430416	0.400228	0.685024	0.764867
0.886114	0.781442	0.463329	0.6425
0.742538	0.745238	0.142568	0.135679
0.793906	0.576869	0.676084	0.059119
0.354756	0.331683	0.644832	0.22886
0.232672	0.657935	0.570328	0.095543
0.585144	0.776646	0.724779	0.192971
0.882132	0.653444	0.780215	0.945795
0.807601	0.690349	0.678746	0.939962
0.274128	0.560913	0.19084	0.865071
0.434231	0.709609	0.660409	0.283014
0.858072	0.14779	0.288494	0.332406
0.695801	0.527305	0.873105	0.278565
0.410019	0.559795	0.647626	0.44836
0.530318	0.329282	0.955733	0.484083
0.263169	0.832082	0.276566	0.731185
0.612787	0.254149	0.532802	0.700092
0.440151	0.745957	0.736691	0.11523
0.52987	0.224635	0.946808	0.754406
0.427625	0.379959	0.096358	0.825157
0.335878	0.145425	0.482592	
0.755171	0.326765	0.704737	
0.178381	0.677627	0.738308	

PRILOG 3. Podaci o neto dobitku

196812	1218302	960817
62442	16316	767612
276168	11152	1272756
81519	148288	41169
101480	4177	1352711
92437	539229	1431417
42490	89632	9000
173821	24834	52911
10516	452180	57362
378944	67000	124941
344718	5142	17472
12839499	53043	55705
2217525	3686	981445
1858784	12657	1053990
182833	119773	13258
2875926	106424	140102
830991	1718871	10622
2673313	245565	147805
2285865	704610	263952
3535259	30292	256984
2389404	306327	76137
25114	299182	29804
21117	7578	244435
0	53309	73576
51999	495165	711006
88172	9445450	
60567	161414	
106776	422282	
70024	289093	
54175	39226	
17503	238862	
4279	204796	
523332	143953	
622744	213866	
219181	117868	
319854	48626	
86674	13058	
87333	859755	
483734	193221	
19561	65661	

PRILOG 4. Podaci o ukupnoj imovini

4132303	24835962	13015862
1809177	650788	5288090
6334506	594653	18044405
2718077	1557190	1440779
2782000	169396	19391743
2254451	11156032	25745168
895271	1374451	937429
4579282	595214	1603882
402147	12999992	1341950
8540594	1770643	3364425
5642512	129515	664532
197824373	531250	2215090
44119335	385044	14095694
39999983	361143	18826869
2148757	2569346	678225
50894179	2024232	2607129
13005071	31532023	265481
47091298	3790734	3559490
29748242	14131304	5354017
55157570	659991	5276727
39496647	4784612	3346202
1686227	4669117	1108021
233961	208567	3777215
19728	1642214	3217509
2349754	13529565	11592285
794220	1.64E+08	
691598	3662992	
3350170	10551371	
506984	3417454	
767536	2053773	
502911	5503576	
426035	4540162	
11510603	2843598	
9344772	4505899	
5091885	1555983	
5435010	14376719	
2437465	404951	
2293740	14671982	
4284316	2951157	
858547	829069	

PRILOG 5. Podaci o troškovima sirovina

890664	4443640	2897138
366053	61031	712190
3221054	30533	3748305
516640	199113	438362
1005774	6809	5397671
1019530	2891167	4827640
158914	415804	17844
320323	265589	320205
46497	2435315	522003
1389951	903326	722766
2160347	21344	164436
8357461	165667	555025
3949782	53852	3390341
3386442	115909	3830860
1078297	1199942	109648
1875196	439528	272752
1578651	6424060	24118
5817656	684543	1004248
1484842	2465120	1123241
7201903	130177	881511
4857676	703136	386717
683896	2147059	701763
28899	2866	882492
0	195584	798450
276459	1535232	1559880
126426	18830709	
135278	2724012	
812842	1319966	
100792	696560	
60574	241109	
153689	1928281	
120691	1506563	
1253496	350658	
2732832	833815	
675108	564271	
1249333	282969	
403638	84886	
604474	3164468	
3700077	379986	
186112	48600	

PRILOG 6. Podaci o broju zaposlenih

2155	3506	1916
456	147	793
800	55	3288
477	174	369
682	25	2494
724	3044	5916
168	299	198
331	111	444
104	2063	440
1357	330	524
750	33	171
14865	105	411
5716	52	2726
4191	84	2593
559	499	129
4414	630	527
1469	5885	110
5343	643	635
3185	2678	1996
7921	182	597
6231	941	262
301	938	344
81	47	1116
0	341	187
585	2928	2618
190	19379	
165	541	
989	1510	
212	1124	
186	245	
171	1036	
137	664	
1448	660	
1734	1098	
1098	434	
710	299	
1846	165	
863	2638	
434	167	
99	222	

PRILOG 7. Podaci o ceni leta

114.47	112.28	192.67	99.16	266.31	186.23	178.94	124.86	188.71	152.57
122.47	194.55	231.06	68.57	186.18	270.65	137.63	167.07	219.66	212.25
214.42	133.8	206.36	237	176.63	157.76	239.16	79.86	287.35	141.32
69.4	118.93	250.88	129.98	125.09	128.94	206.62	151.27	230.74	151.41
158.13	132	299.31	83.51	172.16	131.65	232.15	190.76	174.95	131.81
135.17	131.4	285.76	80.62	137.32	204.13	170.45	178.44	173.51	181.13
152.85	96.78	292	216.89	132.62	184.68	266.44	161.71	238.6	140.39
190.73	94.95	82.76	191.68	170.24	206.06	174.9	88.02	211.49	166.38
129.35	211.09	265.27	198.83	135.91	243.79	216.47	163.41	321.4	239.85
134.17	136.03	95.19	129.48	156.02	172.27	164.82	175.42	251.16	166.16
212.49	208.6	128.75	139.06	76.7	193.78	206.31	140.64	223.58	110.12
173.56	94.79	156.32	74.41	213.08	254.38	247.41	173.78	181.38	164.3
170.67	110.6	137.63	188.43	110.2	307.42	191.79	158.1	157.78	149.65
120.24	221.95	110.61	124	109.96	355.15	257.32	84.05	194.7	82.44
168.69	259.76	128.2	121.07	141.26	286.14	210.99	208.23	258.85	152.25
154.4	243.22	127.37	126.27	99.44	291.35	199.95	127.46	219.13	177.79
114.24	103.47	123.98	156.54	126.18	258.68	246.26	172.15	159.37	157.29
132.29	129.21	122.38	72.03	133.52	146.73	198.95	132.81	178.31	119.91
181.99	103.19	186.26	196.95	80.47	163.37	233.3	171.88	200.44	106.78
233.05	138.67	176.77	86.77	76.18	170.77	265.86	221.5	217.88	87.45
162.21	209.89	90.7	74.08	82.23	187.62	94.32	141.52	149.38	196.27
161.74	116.41	150.14	144.45	72.5	171.26	91.5	119.66	158.7	155.99
71.57	98.62	117.9	242.02	228.64	119.43	190.85	177.35	209.44	231.6
163.63	150.38	87.71	103.89	179.6	146.87	100.47	205.77	181.7	109.72
134.42	130.89	138.96	142.22	249.21	173.89	154.27	192.97	254.99	212.91
165.69	227.23	178.55	71.81	252.85	89.1	147.08	210.03	84.31	178.49
77.82	198.07	92.35	173.01	185.06	117.41	92.55	84.73	231.15	122.55
227.93	199.73	258.83	161.38	248.6	155.42	192.71	214.32	146.93	277.25
74.82	210.45	132.7	177.45	108.11	106.53	135.72	188.15	167.72	86.02
120.5	137.9	134.21	123.67	213.21	222.11	132.67	160.71	200.39	190.66
205.24	92.18	271.07	165.32	237.89	62.88	188.73	187.8	90.58	181.34
132.52	130.36	163.37	180.99	144.67	78.93	192.69	174.72	184.75	159.24
156.55	236.93	133.04	73.42	148.71	73.73	149.81	181.88	274.75	162.86
117.85	102.75	177.41	77.15	236.58	130.69	175.06	132.9	215.13	83.31
233.63	252.66	153.38	78.49	141.77	99.45	83.81	179.94	155.57	204.22
146.06	203.29	170.03	167.57	167.56	80.33	140.35	186.83	221.19	240.12
234.2	98.23	223.26	134.54	170.57	128.2	161.49	130.18	142.89	208.03
184.9	195.29	177.23	180.52	314.01	80.67	77.8	175.61	121.42	77.76
195.72	115.95	245.59	173.49	191.79	145.58	132.86	144.02	180.26	186.63
147.5	236.1	122.89	181.94	154.36	92.89	135.63	190.71	186.15	223.28

PRIOLOG 7. Podaci o ceni leta (nastavak)

78.14	157.93	149.29	124.43	79.19	222.34	78.41	164.21	242.45	327.45
184.36	141.29	183.75	187.61	122.15	162.08	112.46	157.28	71.93	220.24
235.76	180.55	222.8	137.12	200.89	202.23	195.63	76.3	231.87	102.5
224.98	270.89	182.28	133.46	148.82	142.84	179.29	96.44	128.46	76.06
107.28	176.27	203.99	181.77	178.54	184.89	75.47	166.32	174.32	206.3
248.32	143.91	262.81	137.4	276.58	85.81	125.96	87.86	189.3	180.5
191.54	162.72	124.75	134.72	125.15	199.03	150.22	146.45	93.2	190.63
269.76	209.7	179.02	127.56	128.83	313.22	142.08	75.76	80.35	124.68
83.35	187.48	145.26	127.32	123.12	209.62	159.93	94.67	211.29	159.48
206.8	141.59	145	168.55	161.54	249.06	122.11	99.57	97.4	185.57
214.23	161.33	139.48	77.28	156.97	225.27	171.31	70.48	165.91	120.93
189.56	159.98	210.02	134.27	185.2	161.85	71.12	104.42	110.66	160.42
226.96	189.68	174.32	166.14	182.11	120.65	150.04	125.35	76.98	226.77
193.73	245.67	225.83	136.96	125.98	255.5	154.67	152.93	139.4	213.25
223.85	165.91	223.35	174.48	165.28	93.47	164.27	74.55	134.25	275.67
238.67	239.88	165.56	132.92	134.75	122.8	161.36	153.56	205.37	197.35
207.45	178.1	89.36	126.77	160.15	156.11	163.57	211.33	76.46	171.09
227.12	175.06	194.5	174.87	82.7	218.63	129.12	71.48	361.57	224.33
224.54	146.57	199.25	179.74	148.08	170.97	172.27	183.61	216.04	169.68
152.58	178.05	234.28	220.8	148.42	104.43	199.6	61.02	123.41	211.95
223.86	174.44	219.31	173.5	157.84	199.52	50.52	135.45	70.1	211.17
84.77	163.12	253.52	77.23	118.32	122.32	65.98	163.45	128.45	212.02
250.47	187.65	232.82	132.36	212.37	171.73	158.48	121.99	137.34	211.46
283.9	174.49	139.08	126.07	163.93	154.45	158.3	126.6	126.82	216.8
288.99	194.68	207.91	119.26	104.53	148.94	201.61	52.99	158.67	197.02
215.22	197.52	148.33	130.35	190.44	109.86	158.86	123.99	133.71	205.13
254.29	223.35	142.63	178.22	178.65	140.07	151.03	130.11	253.85	242.45
194.21	182.01	105.83	150.93	115.08	104.9	153.65	165.99	200.22	221.07
210.98	186.93	135.62	149.71	273.66	140.87	159.74	268.96	164.94	247.48
80.26	137.86	69.84	133.03	141.06	185.14	176.94	330.28	194.34	211.65
256.12	243.55	116.15	152.88	103.82	105.36	174.34	189.9	166.74	182.01
201.21	161.86	123.21	154.74	271.02	106.82	97.56	226.37	199.87	236
173.97	145.25	129.56	173.42	209.13	173.47	161.93	193.66	210.96	108.13
163.7	139.31	123.94	168.8	116.26	97.89	128.36	186.88	171.97	83.24
132.14	147.64	127.9	139.98	159.96	144.93	58.45	317.2	159.93	229.24
122.52	218.1	169.86	152.08	188.55	162.25	164.19	197.65	173.49	124.05
137.78	211.6	121.3	110.96	160.37	130	164.41	75.02	193.9	242.28
209.57	216.8	75.44	154.55	244.07	161.5	69.97	178.12	163.37	172.95
151.25	170.4	130.1	205.8	163.21	156.87	204.31	151.38	141.97	133.87
140.31	168.21	171.41	81.36	231.91	239.95	109.45	232.41	212.64	94.97

PRILOG 7. Podaci o ceni leta (nastavak)

133.44	165.91	167.56	161.76	213.11
162	143.89	130.72	159.1	232.97
166.19	234	114.84	288.58	215.62
128.97	97.8	124.28	255.73	156.14
164.99	136.75	175.09	135.35	75.34
174.36	249.15	146.36	139.02	194.58
99.1	119.1	196.96	77.41	120.89
156.93	149.09	157.55	87.85	142.71
164.46	97.32	167.79	188.6	143.99
137.67	203.36	273.14	110.74	102.95
110.18	110.43	182.72	139.45	110.64
151.06	72.38	170.12	107.04	98.36
168.31	192.92	128.93	93.79	263.37
116.01	110.45	146.05	136.04	165.13
167.56	93.4	208.72	131.8	177.77
230.27	69.99	175.38	73.7	215.04
177.47	106.26	168.59	263.36	122.02
101.8	85.38	255.28	120.49	82.34
186.11	79.49	236.84	187.88	148.62
85.72	100.92	264.85	128.12	177.91
188.81	108.46	267.3	167.34	74.62
229.58	141.18	274.28	136.44	330.21
200.11	176.66	257.75	161.79	134.18
194.05	160.47	132.61	158.19	116.78
123.54	81.09	141.69	206.32	237.26
217.97	127.43	209.69	166.88	401.23
176.9	138.51	131.84	101.63	77.11
167.42	71.27	162.36	103.21	105.84
245.38	181.6	157.68	69.03	144.22
158.33	178.21	106.71	82.59	322.83
210.6	71.56	109.06	129.2	156.01
93.71	131.54	154.09	203.98	70.61
254.96	102.47	105.99	122.76	162.46
197.34	78.58	138.82	80.81	131.47
209.39	107.83	69.74	86.93	288.14
283.44	78.07	123.64	107.78	136.16
314.66	141.52	109.13	111.06	83.28
365.93	125.32	83.09	162.77	159.97
229.42	165.91	132.19	168.89	73.57
308.08	131.15	132.8	184.88	126.67

PRILOG 8. Podaci o udaljenosti između gradova

528	432	483	370	2457	860	641	773	1144	927
860	1199	1589	281	2438	1124	599	1750	738	1155
852	594	874	360	2335	1367	937	286	812	972
288	644	1891	621	842	200	1338	843	1069	930
723	581	2139	314	1976	1121	500	491	932	550
1204	515	2116	336	883	280	632	1197	879	1001
2237	306	1919	1217	584	2300	921	349	886	1772
2467	352	215	1491	597	496	1917	409	1678	1996
1073	859	2182	408	570	2537	2125	837	1900	627
1130	696	223	925	1618	612	651	740	596	483
1269	432	406	919	321	1046	930	717	589	802
1670	270	547	283	866	2105	543	1844	756	405
1121	692	545	1246	478	2588	468	693	773	1671
580	1747	781	515	460	2704	447	423	411	410
349	1946	977	663	649	2689	1774	1706	1021	1964
767	321	982	966	520	2592	575	1005	1201	2121
487	332	852	2106	674	2496	2296	678	1062	2017
677	595	914	220	735	1193	508	1444	1025	829
981	669	1340	2329	344	413	330	412	475	354
1825	906	972	495	291	1197	264	1751	1107	201
889	214	183	377	399	473	316	849	694	861
1552	425	768	946	287	1212	296	647	1825	650
328	508	525	936	727	1166	802	1680	2053	987
1111	761	160	587	867	1987	240	642	1080	562
628	516	1085	998	752	2217	895	528	622	1119
1180	2130	1238	159	563	301	306	1790	448	999
321	761	335	2446	640	1011	235	258	425	461
1650	821	1035	1180	1562	1912	1237	1258	895	1471
324	1900	753	2283	709	1054	1182	1041	1737	252
545	403	446	787	1754	1243	1120	1728	540	762
812	272	1519	1998	632	223	590	1854	487	919
576	665	995	2358	1237	325	594	1838	2026	461
946	1587	868	328	1249	369	783	1731	2161	1067
712	526	1161	255	1609	817	945	1050	2021	304
227	2172	717	277	1011	374	177	1733	932	1220
606	903	1161	2395	1256	358	677	607	310	1246
373	356	1471	737	2381	326	865	1012	603	733
554	481	1764	1864	2611	296	405	1440	604	293
446	749	1313	1407	1139	937	1521	585	833	432
732	2092	682	2295	1258	267	552	612	617	1121

PRILOG 8. Podaci o udaljenosti između gradova (nastavak)

319	533	1750	1093	295	1850	491	2363	2401	2585
853	629	1979	2342	851	1050	690	2027	370	2724
853	862	610	1263	1519	1632	1113	345	2136	403
631	1709	1145	1243	2298	687	1489	397	834	204
448	895	238	1487	2527	1220	313	1372	2592	921
1391	680	528	793	1050	199	1072	368	2239	297
1213	1062	457	673	852	1320	1038	1069	390	746
1468	1626	926	1076	1279	1656	1044	258	373	1066
181	957	508	781	1050	1629	1481	414	1593	1460
585	495	2068	2575	2213	1365	904	386	590	1845
1200	472	957	178	532	1900	1442	226	1210	1077
984	819	453	992	957	787	237	866	337	634
1302	1545	1671	1972	2502	456	919	806	308	695
879	1557	1953	994	1111	1218	706	1984	954	1040
1068	602	614	2693	326	451	1333	365	945	1028
1626	1290	501	1188	1133	1591	1506	1076	2158	1494
1062	992	440	680	862	1814	1501	2089	451	1310
1158	1436	1215	1056	359	503	1048	296	2311	980
1442	804	1957	2268	817	664	950	2465	662	1276
551	910	2079	2583	668	828	647	353	718	726
999	770	2063	2717	1243	587	231	965	254	1426
248	391	1960	197	393	1489	236	2300	727	981
1182	794	1932	901	1397	1783	1624	803	820	1518
1476	853	983	1180	803	1944	2356	972	1142	449
1450	967	405	946	482	1866	2175	108	2565	991
1216	948	1087	1183	484	840	1524	1034	1204	1532
1670	846	684	1416	964	873	1300	1052	700	1589
929	1024	584	1080	449	421	1588	2587	964	1576
824	1506	714	992	1059	987	1501	1619	683	1522
237	639	347	966	670	2153	2248	2342	874	1399
1192	1476	496	892	305	484	2155	1756	656	1307
1103	1038	636	1643	1428	513	407	1536	762	931
1925	1149	748	408	1649	834	987	1797	1501	563
554	1127	1011	626	423	543	1099	1671	674	471
808	1085	1173	1023	807	742	197	2475	1097	766
1123	548	965	461	1353	634	2039	2371	1013	616
1703	1093	1005	534	853	1140	2176	337	1972	675
1671	231	318	365	1335	1363	256	1187	1013	1448
889	814	1242	610	1028	436	1910	1330	700	907
977	629	2174	429	1131	393	762	2217	1068	443

PRILOG 8. Podaci o udaljenosti između gradova (nastavak)

822	1047	2534	1788	1736
1751	722	1073	935	1570
1977	2421	534	1979	1710
612	209	1033	2253	869
1183	1019	2407	873	351
1903	2136	880	952	719
550	229	1929	444	626
1088	1035	1040	479	599
1301	655	2148	1708	585
918	2330	2445	630	588
779	717	2419	933	689
1879	361	2553	550	546
604	2437	1053	569	1851
495	646	759	859	1129
1599	543	2075	129	1774
1911	180	267	279	1381
2086	588	336	2350	447
487	446	813	566	417
969	371	1495	2567	1050
288	671	2369	1137	2087
2576	723	2521	372	367
1345	747	2380	1148	2276
1155	2432	2378	667	372
2430	833	928	2193	678
950	462	951	2379	2392
2153	1210	1813	2354	2442
340	1037	1009	587	342
284	342	2277	488	697
2454	1012	1891	188	721
431	2171	601	564	2424
292	325	647	1079	978
264	838	1262	1679	224
2521	394	507	532	2520
892	389	843	480	1216
1989	558	304	404	2329
1587	333	651	605	1104
2446	956	621	649	200
2586	861	338	1156	814
2569	1848	1107	786	174
2454	834	1020	1558	859

PRILOG 9. Podaci o meri sistematskog rizika

0.95	1.76	0.55	1.5	1.46	1.63	1.07	1.14	0.66	0.53
1.54	0.86	0.67	0.31	1.62	3.04	0.59	0.96	1.19	1.11
1.43	1.65	0.78	0.73	1.07	0.5	0.9	0.93	1.44	1.5
1.08	1.39	0.13	1.28	0.12	0.78	1.8	0.66	0.29	1.51
0.79	1.03	1.23	-0.01	0.84	1.12	1.73	0.67	2	0.56
1.74	0.73	0.89	0.08	0.84	0.89	1.47	1.68	0.33	0.88
0.99	1.06	0.32	0.7	1.25	1.11	1.26	2.71	0.75	0.98
3.09	1.24	1.67	1.24	0.63	0.76	1.33	1.64	0.4	1.52
0.78	0.69	0.92	0.96	0.32	0.54	0.8	1.35	0.22	0.36
1.48	1.43	1.64	0.07	0.9	1.14	1.92	1.03	0.93	0.47
0.59	1.82	2.19	0.24	0.74	0.96	0.56	1.38	0.66	1.09
1.74	1.05	1.52	1.37	0.99	1.25	1.61	1.47	1.48	0.55
1.07	1.3	1.47	0.95	0.72	1.39	1.16	1.17	1.08	0.74
1.56	0.64	1.02	0.6	0.57	1.78	0.41	0.74	0.65	1.53
1.01	0.44	0.57	1.16	0.6	0.47	0.68	0.48	1.29	0.98
0.35	2.34	0.82	1.03	0.7	0.84	1.4	0.8	1.73	0.93
0.71	0.8	0.88	-0.12	0.25	0.97	1.21	0.85	1.59	1.83
2.33	1.08	0.86	1.07	0.61	1.22	1.44	0.97	0.82	0.74
1.07	0.56	1.79	0.72	0.38	0.78	0.64	1.42	0.26	0.18
1.26	0.53	1.4	1.77	1.03	0.75	0.36	1.3	1.25	1.45
1.5	1.27	2.33	1.07	0.63	0.11	0.91	-0.49	1.2	0.49
1.22	0.59	1.2	1.44	1.08	0.64	0.83	1.26	1.12	1.64
1.25	0.99	0.1	2.44	0.42	1.2	0.53	2	1.17	1.46
1.63	1.61	0.88	0.24	0.65	1.1	0.74	1.4	1.25	0.57
0.39	1.07	0.39	1.56	1.06	0.14	1.76	0.46	1.01	0.33
1	0.36	0.48	1.6	0.45	1.5	0.98	1.46	0.88	1.31
0.92	0.78	0.92	1.39	0.8	1.33	0.3	0.63	1.49	1.18
0.94	1.12	0.84	1.17	0.32	0.53	0.31	0.74	0.95	0.78
0.51	0.98	0.85	0.74	1.48	1.09	1.39	1.04	0.7	1.34
1.42	0.76	1.4	0.45	0.81	1	1.16	1.31	1	1.07
0.31	0.36	1.58	0.26	1.43	0.34	1.6	1.08	0.61	1.11
1.33	0.45	1.39	1.21	0.24	1.23	1.04	1.5	1.07	1.62
0.15	1.73	0.25	1.44	0.81	1.79	1.41	1.43	0.15	0.14
1.13	-0.12	0.74	0.54	0.6	0.78	1.04	0.22	0.93	0.82
1.34	1.22	0.13	0.14	1.35	1.3	1.19	2.54	0.89	1.36
1	1.9	0.58	1.85	1.34	0.49	0.17	1.47	0.23	1.71
0.16	0.94	0.67	0.1	1.91	1.15	2.1	0.61	0.73	1.11
1.92	0.94	0.98	1.39	0.31	1.27	1.06	0.89	0.97	0.63
0.89	0.99	0.75	1.3	1.33	1.09	0.59	2.45	0.27	1.44
1.2	1.11	1.26	1.3	1.24	1.36	0.67	1.5	1.68	0.55

PRILOG 9. Podaci o meri sistematskog rizika (nastavak)

0.63	0.7	2.03
0.71	0.93	0.09
1.37	-0.05	0.88
1.04	1.03	1.05
0.06	1.03	0.87
1.21	0.59	1.12
1.39	0.86	0.8
1.24	2.87	1.14
1.55	1.08	1.61
0.65	0.83	1.24
1.28	1.52	
0	0.66	
0.59	1.03	
1.42	0.86	
1.02	0.29	
1.01	1.06	
1.2	0.7	
0.41	0.58	
1.15	1.92	
0.27	1.7	
1.14	1.08	
0.89	1.17	
1.22	0.98	
1.26	0.07	
1.54	0.85	
0.98	1.3	
1.81	0.65	
0.85	1.24	
0.56	0.02	
0.92	1	
1.14	0.49	
1.16	1.14	
0.64	0.97	
1.63	1.73	
1.19	1.62	
1.97	1.75	
1.46	0.63	
1.4	1.64	
0.3	1	
0.42	1.35	

PRILOG 10. Podaci o dividendama

4.7	1	1.07	2.4	3.92	3.56	0.16	0.8	1.84	2.32	2.8
1.06	1.44	2.16	1.16	2	2.66	0.25	0.92	2.76	0.48	0
2.56	2.52	1.2	0.81	1.16	0.68	2.08	0.2	0.68	1	3.1
2.42	0.4	0.32	0.35	0.56	0.8	1.16	0.2	1.76	1.52	0
0.3	1.28	0	0.24	1.44	0.53	7.28	3.93	3	1.7	2.31
0.52	0.24	4.32	3.72	0.6	1.38	0.25	1.36	1.72	1.04	0.8
0	1.56	2.84	1.2	0.76	0.32	1.64	0.72	8	0.3	0.66
0	2.12	0.76	0	0.84	2.6	3.6	0.7	0.36	0.52	2.84
0.24	1.96	0.04	1.04	0.66	1.6	3	0.4	0.15	1.32	1
0.48	0	0.32	3.02	1.24	1.09	2.84	1.48	2.28	2.28	2.04
2	2.28	2	1.76	0.6	0.8	1.51	2.44	1.8	2.48	1.5
0.8	5.68	1.33	1.84	0.28	6	0.2	1.52	0.08	1.6	1.56
1.72	1.8	1.16	2.32	0.72	1.85	1.44	4	0.88	0	1.7
0.53	0.68	0.64	3.3	0.8	0.72	1.32	0.12	3.19	2.2	0
3.8	0.4	0.56	1.52	0	2.56	1.5	1.51	2.54	2	2.08
1.2	0.3	3.36	3.42	0.92	1.36	1.68	0.56	2.12	0.08	1.52
1.28	0.76	2.64	0.72	2.04	1.16	0.4	3.04	0.28	0.6	3.48
3.44	1.04	1.33	0	0	2.2	0.88	2.2	1.28	2	2
0.64	0.64	1.35	2.04	3.36	0.6	0.6	2.46	2.04	1.61	0.7
2.8	1.2	1.48	2.4	0.96	0.92	1.88	0.76	0.96	1.25	1.6
2.08	2.92	0.6	1.48	0.88	3	3.76	1	3.04	0.6	1.24
1.26	0.6	1.6	2.17	1.96	3.36	1.84	2.64	1.32	0.4	4.4
1.48	1.24	0.63	1.92	2.7	1	1.88	1.04	1.4	1.08	0.72
2.44	10	1.04	3.48	2.08	2	1.6	1.84	0.64	0	1.2
1.76	0.96	0.8	0.67	0.05	0.4	1.45	1.38	1.92	0	2.12
0.4	5.68	1.06	0.12	3	1.32	0	0.69	2.45	2.88	2.32
2.36	0.56	2.76	1.56	0.4	2.08	1.56	3.22	2	0.36	2
1.28	3	2.08	8	5.12	0.38	3.48	0.28	1.2	0.36	1.44
1.28	1.56	0.06	2.02	0.72	3.88	1.64	0.64	2.52	0.9	0.25
2.56	4.08	0.62	7	1.46	1.08	0.76	1.28	0.64	1.24	1.4
1.66	1.8	2	1.36	2.12	0.5	2.16	2.12	3.29	0	0.88
3.32	1.02	0.5	1.9	0.92	2.16	1.52	4.16	3.4	1.12	0.72
1.46	0.2	3.8	5	2.28	2.2	0.8	2.8	1.24	2.42	1.2
0.36	1.4	0.4	1.31	0	2.4	0.6	2.62	7	0	0.96
4.6	1.6	0.8	1.12	1.48	0.5	1.88	0.08	1.12	3	0.32
0.64	1.85	4.1	0.84	2.8	2.4	1	2.2	3.1	3.32	0.42
0.2	1	2	0	2.47	3	0	2	2.84	2.8	
1.8	1.6	0.4	3.12	1	0	1.52	1.6	2.32	0.4	
2.6	3.12	0.56	3.08	0.26	1.8	0.2	1.58	0.4	0.8	
1.44	0.72	2.52	1.28	0.6	1.44	0.64	3.15	1.64	1.68	

SPISAK TABELA

Tabela 1.1: Balansirani rangirani uzorak sa m ciklusa i k jedinica u svakom ciklusu.....	15
Tabela 2.1: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent kada slučajna greška ima eksponencijalnu raspodelu.....	49
Tabela 2.2: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent kada slučajna greška ima Weibull-ovu raspodelu.....	50
Tabela 2.3: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_1 kada slučajna greška ima gama raspodelu.....	51
Tabela 2.4: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_2 kada slučajna greška ima gama raspodelu.....	52
Tabela 2.5: Proporcije simuliranih 90% dvostranih intervala za regresioni koeficijent β_1 kada slučajna greška ima gama raspodelu.....	53
Tabela 2.6: Proporcije simuliranih 90% dvostranih intervala za regresioni koeficijent β_2 kada slučajna greška ima gama raspodelu.....	54
Tabela 2.7: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_1 kada slučajna greška ima Weibull-ovu raspodelu.....	55
Tabela 2.8: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_2 kada slučajna greška ima Weibull-ovu raspodelu.....	56
Tabela 2.9: Proporcije simuliranih 90% dvostranih intervala za regresioni koeficijent β_1 kada slučajna greška ima Weibull-ovu raspodelu.....	57
Tabela 2.10: Proporcije simuliranih 90% dvostranih intervala za regresioni koeficijent β_2 kada slučajna greška ima Weibull-ovu raspodelu.....	58
Tabela 2.11: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_1 kada slučajna greška ima eksponencijalnu raspodelu.....	59
Tabela 2.12: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_2 kada slučajna greška ima eksponencijalnu raspodelu.....	60
Tabela 2.13: Proporcija simuliranih 90% dvostranih intervala za regresioni koeficijent β_1 kada slučajna greška ima eksponencijalnu raspodelu.....	61
Tabela 2.14: Proporcija simuliranih 90% dvostranih intervala za regresioni koeficijent β_2 kada slučajna greška ima eksponencijalnu raspodelu.....	62

Tabela 2.15: EM-skor i kreditni rejting kompanije.....	67
Tabela 2.16: Deskriptivne mere za indikator verovatnoće bankrotstva i količnik ukupnog duga.....	68
Tabela 2.17: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent kada indikator verovatnoće bankrotstva zavisi od količnika ukupnog duga.....	70
Tabela 2.18: Deskriptivne mere za zbirne finansijske indikatore.....	71
Tabela 2.19: Širine 95% dvostranih intervala poverenja za regresioni koeficijent kada neto dobitak zavisi od troškova sirovina.....	74
Tabela 2.20: Širine 95% dvostranih intervala poverenja za regresioni koeficijent kada neto dobitak zavisi od broja zaposlenih.....	75
Tabela 2.21: Deskriptivne mere za prosečnu cenu i udaljenost.....	75
Tabela 2.22: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent kada prosečna cena zavisi od udaljenosti.....	77
Tabela 2.23: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_1 kada neto dobitak zavisi od ukupne imovine.....	79
Tabela 2.24: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_2 kada neto dobitak zavisi od ukupne imovine.....	79
Tabela 2.25: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_1 kada indikator verovatnoće bankrotstva zavisi od količnika ukupnog duga.....	81
Tabela 2.26: Proporcije simuliranih 95% dvostranih intervala za regresioni koeficijent β_2 kada indikator verovatnoće bankrotstva zavisi od količnika ukupnog duga.....	81
Tabela 2.27: Kodovi za t interval i <i>bootstrap-t</i> interval za regresioni koeficijent u prostom linearnom regresionom modelu.....	82
Tabela 2.28: Kodovi za T_1 interval i T_2 interval za regresioni koeficijent u prostom linearnom regresionom modelu.....	83
Tabela 2.29: Kodovi za percentil interval i <i>jackknife</i> interval za regresione koeficijente u kvadratnom regresionom modelu.....	84
Tabela 2.30: Kodovi za BCa interval i T_2^* interval za regresione koeficijente u kvadratnom regresionom modelu.....	85
Tabela 3.1: Proporcije simuliranih 95% dvostranih intervala za varijansu gama raspodele.....	96

Tabela 3.2: Proporcije simuliranih 95% dvostranih intervala za varijansu log-normalne raspodele.....	97
Tabela 3.3: Proporcije simuliranih 95% dvostranih intervala za varijansu eksponencijalne raspodele.....	97
Tabela 3.4: Proporcije simuliranih 95% dvostranih intervala za varijansu <i>skew</i> -normalne raspodele.....	98
Tabela 3.5: Proporcije simuliranih 95% dvostranih intervala za varijansu <i>skew</i> -Student-ove raspodele.....	99
Tabela 3.6: Proporcije simuliranih 95% dvostranih intervala za varijansu <i>skew</i> -Laplace-ove raspodele.....	100
Tabela 3.7: Proporcije simuliranih 95% dvostranih intervala za varijansu <i>skew</i> generalizovane <i>t</i> raspodele.....	101
Tabela 3.8: Proporcije simuliranih 95% dvostranih intervala za koeficijent varijacije gama raspodele.....	103
Tabela 3.9: Proporcije 95% dvostranih simuliranih intervala za koeficijent varijacije Weibull-ove raspodele.....	105
Tabela 3.10: Proporcije 95% dvostranih simuliranih intervala za koeficijent varijacije log-normalne raspodele.....	107
Tabela 3.11: Proporcije 95% dvostranih simuliranih intervala za koeficijent varijacije <i>skew</i> -normalne raspodele.....	109
Tabela 3.12: Prosečna širina 95% dvostranog intervala za koeficijent varijacije gama raspodele.....	111
Tabela 3.13: Prosečna širina 95% dvostranog intervala za koeficijent varijacije Weibull-ove raspodele.....	113
Tabela 3.14: Prosečna širina 95% dvostranog intervala za koeficijent varijacije log-normalne raspodele.....	115
Tabela 3.15: Prosečna širina 95% dvostranog intervala za koeficijent varijacije <i>skew</i> -normalne raspodele.....	117
Tabela 3.16: Deskriptivne mere za meru sistematskog rizika.....	121
Tabela 3.17: Deskriptivne mere za dividendu.....	123
Tabela 3.18: Proporcije simuliranih 95% dvostranih intervala za varijansu mere sistematskog rizika.....	124
Tabela 3.19: Proporcije simuliranih 95% dvostranih intervala za varijansu dividende.....	125

Tabela 3.20: Deskriptivne mere za logaritam iznosa šteta.....	126
Tabela 3.21: Proporcije simuliranih 95% dvostranih intervala za varijansu logaritama iznosa šteta.....	126
Tabela 3.22: Proporcije simuliranih 95% dvostranih intervala za koeficijent varijacije mere sistematskog rizika.....	128
Tabela 3.23: Prosečna širina 95% dvostranog intervala poverenja za koeficijent varijacije mere sistematskog rizika.....	129
Tabela 3.24: Proporcije simuliranih 95% dvostranih intervala za koeficijent varijacije dividende.....	130
Tabela 3.25: Kodovi za χ^2 interval i Bonett-ov interval za varijansu.....	131
Tabela 3.26: Kodovi za interval poverenja sa korigovanim stepenima slobode i RSS Bonett-ov interval za varijansu.....	132
Tabela 3.27: Kodovi za Steve-ov interval zasnovan na normalnoj aproksimaciji u slučaju velikih uzoraka i interval zasnovan na logaritamskoj transformaciji za varijansu.....	133
Tabela 3.28: Kodovi za t interval i percentil interval za koeficijent varijacije.....	134
Tabela 3.29: Kodovi za <i>jackknife</i> interval i <i>bootstrap-t</i> interval za koeficijent varijacije.....	135
Tabela 3.30: Kodovi za BC interval i BCa interval za koeficijent varijacije.....	136
Tabela 3.31: Kodovi za RSS BCa interval i RSS percentil interval za koeficijent varijacije.....	137
Tabela 3.32: Kodovi za Miller-ov interval i Miller-ov interval zasnovan na Bonett-ovoj odsečenoj sredini za koeficijent varijacije.....	138
Tabela 3.33: Kodovi za Miller-ov interval zasnovan na odsečenoj sredini koja eliminiše 10% vrednosti i Miller-ov <i>bootstrap-t</i> interval za koeficijent varijacije.....	138
Tabela 3.34: Kodovi za RSS Miller-ov interval i RSS Miller-ov interval zasnovan na Bonett-ovoj odsečenoj sredini za koeficijent varijacije.....	139
Tabela 3.35: Kodovi za RSS Miller-ov interval zasnovan na odsečenoj sredini koja eliminiše 10% vrednosti i PRSS Miller-ov interval za koeficijent varijacije.....	140
Tabela 3.36: Kodovi za Curto-ov & Pinto-ov interval i McKay-ev interval za koeficijent varijacije.....	141
Tabela 5.1: Rezultati analize lanaca Markova za koeficijent varijacije log-normalne raspodele ako je apriorna raspodela $B(3;0,4)$	157

Tabela 5.2: Rezultati analize lanaca Markova za koeficijent varijacije log-normalne raspodele ako je apriorna raspodela $U(0,4)$	158
Tabela 5.3: Rezultati analize lanaca Markova za koeficijent varijacije log-normalne raspodele ako je apriorna raspodela $B_4(5;0,7;0;3)$	159
Tabela 5.4: Širine 95% dvostranog intervala poverenja za koeficijent varijacije log-normalne raspodele.....	160
Tabela 5.5: Kolmogorov-Smirnov i hi-kvadrat statistike testa.....	160
Tabela 5.6: Rezultati analize lanaca Markova za koeficijent varijacije neto dobitka ako je apriorna raspodela $U(3,6)$	161
Tabela 5.7: Rezultati analize lanaca Markova za koeficijent varijacije neto dobitka ako je apriorna raspodela $B_4(5;0,7;3;6)$	162
Tabela 5.8: Širine 95% dvostranog intervala poverenja za koeficijent varijacije neto dobitka.....	163

SPISAK SLIKA

Slika 1.1: Histogram <i>bootstrap</i> raspodele statistike \bar{X}^*	7
Slika 2.1: Funkcija gustine Weibull-ove raspodele.....	18
Slika 2.2: Funkcija gustine eksponencijalne raspodele.....	19
Slika 2.3: Funkcija gustine gama raspodele.....	20
Slika 2.4: Funkcija gustine log-normalne raspodele.....	21
Slika 2.5: Funkcija gustine <i>skew</i> -normalne raspodele.....	22
Slika 2.6: Funkcija gustine <i>skew</i> -Student-ove raspodele.....	23
Slika 2.7: Funkcija gustine <i>skew</i> generalizovane <i>t</i> raspodele.....	25
Slika 2.8: Funkcija gustine <i>skew</i> -Laplace-ove raspodele.....	26
Slika 2.9: Histogram raspodele indikatora verovatnoće bankrotstva.....	68
Slika 2.10: Histogram raspodele količnika ukupnog duga.....	69
Slika 2.11: Dijagram raspršenosti na osnovu podataka o količniku ukupnog duga i indikatoru verovatnoće bankrotstva.....	69
Slika 2.12: Histogram raspodele slučajne greške.....	70
Slika 2.13: Histogram raspodele neto dobitka.....	71
Slika 2.14: Histogram raspodele ukupne imovine.....	72
Slika 2.15: Histogram raspodele troškova sirovina.....	72
Slika 2.16: Histogram raspodele broja zaposlenih.....	73
Slika 2.17: Dijagram raspršenosti na osnovu podataka o troškovima sirovina i neto dobitku.....	73
Slika 2.18: Dijagram raspršenosti na osnovu podataka o broju zaposlenih i neto dobitku.....	74
Slika 2.19: Histogram raspodele prosečne cene.....	75
Slika 2.20: Histogram raspodele udaljenosti.....	76
Slika 2.21: Histogram raspodele slučajne greške.....	76
Slika 2.22: Dijagram raspršenosti na osnovu podataka o neto dobitku i ukupnoj imovini.....	78
Slika 2.23: Histogram raspodele slučajne greške.....	78

Slika 2.24: Dijagram raspršenosti na osnovu podataka o indikatoru verovatnoće bankrotstva i količniku ukupnog duga.....	80
Slika 2.25: Histogram raspodele slučajne greške.....	80
Slika 3.1: Histogram raspodele mere sistematskog rizika.....	121
Slika 3.2: Histogram raspodele dividende.....	124
Slika 3.3: Histogram raspodele logaritama iznosa šteta.....	126
Slika 4.1: Lanac Markova koji brzo konvergira ka ciljnoj raspodeli.....	150
Slika 4.2: Lanac Markova koji sporo konvergira ka ciljnoj raspodeli.....	150
Slika 4.3: Lanac Markova koji se zadržava u jednom stanju veliki broj iteracija.....	151
Slika 4.4: Autokorelacija brzo iščezava.....	151
Slika 4.5: Autokorelacija sporo iščezava.....	151

BIOGRAFIJA AUTORA

Ivana Ivković je rođena 1988. godine u Tuzli. Osnovnu školu i gimnaziju završila je u Šapcu sa odličnim uspehom. Diplomirala je na Ekonomskom fakultetu u Beogradu 2011. godine, na smeru Statistika, informatika i kvantitativne finansije, opciona grupa Statistika, sa prosečnom ocenom 9,55. Tokom studija bila je dobitnik većeg broja stipendija.

Master studije na Ekonomskom fakultetu u Beogradu (studijski program Kvantitativna analiza, podmodul Ekonometrija) završila je sa prosečnom ocenom 9,57. Master tezu pod nazivom „Primena log-linearnih i logit modela u analizi tabela kontingencije“ odbranila je 2014. godine, nakon čega je upisala doktorske studije na Ekonomskom fakultetu u Beogradu (modul Statistika). Sve ispite položila je sa prosečnom ocenom 10,00.

Rad na Ekonomskom fakultetu u Beogradu započela je 2012. godine kao demonstrator na predmetu Osnovi statističke analize. U zvanje saradnika u nastavi za užu naučnu oblast Statistika i Matematika - Osnovi statističke analize izabrana je 2013. godine, dok je u zvanje asistenta unapređenja 2015. godine. Autor je većeg broja radova na temu statističkog zaključivanja o različitim parametrima skupa koristeći programski jezik *R*.

Прилог 1.

Изјава о ауторству

Потписана **Ивана Ивковић**

број индекса _____

Изјављујем

да је докторска дисертација под насловом

“Непараметарске статистичке технике оцењивања регресионих коефицијената и коефицијента варијације у корпоративним финансијама”

- резултат сопственог истраживачког рада,
- да предложена дисертација у целини ни у деловима није била предложена за добијање било које дипломе према студијским програмима других високошколских установа,
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио интелектуалну својину других лица.

Потпис докторанда

У Београду, _____

Прилог 2.

Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора **Ивана Ивковић**

Број индекса _____

Студијски програм **Статистика**

Наслов рада **“Непараметарске статистичке технике оцењивања регресионих
коэффицијената и коэффицијента варијације у корпоративним финансијама”**

Ментор **др Весна Рајић, редовни професор**

Потписана **Ивана Ивковић**

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла за објављивање на порталу **Дигиталног репозиторијума Универзитета у Београду**.

Дозвољавам да се објаве моји лични подаци везани за добијање академског звања доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

Потпис докторанда

У Београду, _____

Прилог 3.

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

“Непараметарске статистичке технике оцењивања регресионих коефицијената и коефицијента варијације у корпоративним финансијама”

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигитални репозиторијум Универзитета у Београду могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство

2. Ауторство - некомерцијално

3. Ауторство – некомерцијално – без прераде

4. Ауторство – некомерцијално – делити под истим условима

5. Ауторство – без прераде

6. Ауторство – делити под истим условима

Потпис докторанда

У Београду, _____
