

UNIVERZITET CRNE GORE
PRIRODNO-MATEMATIČKI FAKULTET

Andela Mijanović

**Višedimenzionalna logistička
raspodjela specifikovana svojom
karakterističnom funkcijom -
doprinos teoriji i praksi
DOKTORSKA DISERTACIJA**

Podgorica, 2023.

UNIVERSITY OF MONTENEGRO
FACULTY OF SCIENCES AND MATHEMATICS

Andela Mijanović

**Multivariate logistic distribution
specified by its characteristic
function
– contribution to the theory and
practice
PHD THESIS**

Podgorica, 2023.

Podaci i informacije o doktorandu

Ime i prezime: **Anđela Mijanović**

Datum i mjesto rođenja: **17.09.1994. godine, Nikšić, Crna Gora**

Naziv završenog postdiplomskog studijskog programa i godina završetka: **Matematika i računarske nauke, 2018.**

Podaci i informacije o mentoru

Ime i prezime: **Božidar Popović**

Titula: **doktor matematičkih nauka**

Zvanje: **vanredni profesor**

Naziv univerziteta i organizacione jedinice: **Univerzitet Crne Gore, Prirodno-matematički fakultet**

Članovi komisije:

dr Siniša Stamatović, redovni profesor Prirodno-matematičkog fakulteta, Univerziteta Crne Gore, predsjednik

dr Vladimir Božović, redovni profesor Prirodno-matematičkog fakulteta, Univerziteta Crne Gore

dr Bojana Milošević, vanredni profesor Matematičkog fakulteta, Univerziteta u Beogradu

dr Goran Popivoda, docent Prirodno-matematičkog fakulteta, Univerziteta Crne Gore

dr Božidar Popović, vanredni profesor Prirodno-matematičkog fakulteta, Univerziteta Crne Gore, mentor

Datum odbrane:

10.06.2023. godine

PODACI O DOKTORSKOJ DISERTACIJI

Naziv doktorskih studija: Matematika, Prirodno-matematički fakultet, Univerzitet Crne Gore

Naslov disertacije: Višedimenzionalna logistička raspodjela specifikovana svojom karakterističnom funkcijom - doprinos teoriji i praksi

Rezime: Cilj ove disertacije je dobijanje raspodjele višedimenzionalne logističke slučajne veličine na osnovu karakteristične funkcije. Neki od prikazanih rezultata su dobijeni numeričkim metodama koji se prvi put javljaju u literaturi.

Ključne riječi: Foksova H funkcija, Melinova transformacija, inverzna Melinova transformacija, linearna kombinacija, generisanje slučajnih brojeva, test saglasnosti, primjena u finansijama

Naučna oblast: Matematika

Uža naučna oblast: Matematička statistika

UDK broj: 519.2(043.3)

INFORMATION ON THE PHD THESIS

Name of the doctoral program: Mathematics, Faculty of Sciences and Mathematics, University of Montenegro

Thesis title: Multivariate logistic distribution specified by its characteristic function – contribution to the theory and practice

Summary: The goal of this dissertation is to obtain distribution of multivariate logistic random variable specified by its characteristic function. Some of the provided results obtained by numerical methods appear for the first time in the literature.

Key words: Fox-H function, Mellin transform, Inverse Mellin transform, Linear combination, Generation random numbers, Goodness of fit test, Application in finance

Scientific field: Mathematics

Scientific sub-field: Mathematical statistics

UDC: 519.2(043.3)

Predgovor

U osnovi ove disertacije prožimaju se teorija vjerovatnoće i matematičke statistike i rezultati dobijeni primjenom numeričke analize u statistici.

Sadržaj disertacije podijeljen je na sedam glava, a svaka glava podijeljena je na paragrafe. Paragrafi, definicije i tvrđenja, numerisani su sa dva broja, pri čemu je prvi od tih brojeva redni broj glave, a drugi je redni broj paragrafa, definicije ili tvrđenja.

Prva glava je uvodnog karaktera, tu smo naveli osnovne pojmove iz oblasti vjerovatnoće, statistike i numeričke analize. U prvom paragrafu je uveden prostor vjerovatnoće kao osnovni objekat proučavanja u Teoriji vjerovatnoća. Dati su osnovni pojmovi iz ove oblasti koju su neophodni za razumijevanje sadržaja disertacije. Drugi paragraf je posvećen kopula funkciji. Njihova primjena počela je pedesetih godina XX vijeka kada je Sklar objavio rad u kome je formulisao i dokazao teoremu kojom se višedimenzionalne apsolutno-neprekidne raspodjele na jedinstven način povezuju sa kopula funkcijama. Kopule su veoma važno oruđe u proučavanju zavisnih slučajnih veličina. Naime, mjere zavisnosti kao što su Spirmanov ρ , Kendalov tau ili Džinijev koeficijent se efektnije proučavaju ukoliko imamo poznatu kopulu. S druge strane poznavanje kopula omogućava i proučavanje zavisnosti na repu raspodjele (engl. tail dependence). Dakle, kopule omogućavaju fleksibilniji pristup proučavanju višedimenzionalnih raspodjela. S druge strane kopule nalaze i praktičnu primjenu u gotovo svim ostalim naukama. Četvrti paragraf je posvećen osnovnim elementima matematičke statistike i najvažnijim rezultatima koju su značajni za kasnija zaključivanja. U nastavku smo dali pregled specijalnih funkcija koje su nam bile potrebne u toku istaživanja. Takođe smo naveli definicije i osnovne karakteristike Furijeove i Melinove transformacije. Poseban paragraf je posvećen numeričkim metodama. Numerički proračun jednodimenzionalnih integrala ili kvadrature je jedna od najstarijih grana numeričke analize. Ovaj paragraf fokusira se na metodama baricentrične interpolacije i singularne dekompozicije. Pored osnovnih teorijskih pojmova, naveli smo i praktične primjere prikazane u softveru MATLAB. U posljednjem paragrafu u ovoj glavi opisana je jednodimenzionalna logistička raspodjela.

U drugoj glavi predstavljeni su neki od dosadašnjih rezultata koji su postignuti u okviru ove teme. Dat je kratak pregled najznačajnijih istraživanja koja su nam koristila u izradi distertacije. U daljim glavama prikazani su originalni rezultati.

Treća i četvrta glava su posvećene raspodjeli linearne kombinacije nezavisnih logistički raspodijeljenih slučajnih veličina. Cilj je da se nađu tačna funkcija raspodjele i funkcija gustine pomenute linearne kombinacije. Pokazaćemo da se raspodjela linearne kombinacije nezavisnih logistički raspodijeljenih slučajnih veličina može izraziti u obliku Foksove H funkcije. Poseban paragraf u trećoj glavi posvećen je opisu algoritma koji se koristi za numeričko izračunavanje vrijednosti Foksove H funkcije.

U četvrtoj glavi je opisana primjena metoda za numeričku inverziju karakteristične funkcije. Na osnovu algoritma koji se bazira na numeričkoj inverziji karakteristične funkcije provjerava se pouzdanost algoritma baziranog na Foksovoj H funkciji.

S obzirom da je u literaturi veoma malo radova posvećeno višedimenzionalnim logističkim raspodjelama (isključujući standardnu dvodimenzionalnu logističku raspodjelu) u petoj glavi opisali smo konstrukciju dvodimenzionalne logističke raspodjele koja će nam omogućiti proučavanje raspodjele linearne kombinacije dvije zavisne logistički raspodijeljene slučajne veličine. Poseban paragraf posvećen je opisu algoritma za numeričku inverziju dvodimenzionalne karakteristične funkcije. Nakon toga dat je opis algoritma za generisanje slučajnih brojeva, što ćemo dalje koristiti za izračunavanje mjera zavisnosti za dvodimenzionalnu logističku raspodjelu.

Poznato je da linearne kombinacije slučajnih veličina nalaze široku praktičnu primjenu. U glavi šest ilustrovana je njihova primjena koristeći odgovarajuće podatke iz finansija.

Tema sedme glave je test saglasnosti sa dvodimenzionalnom logističkom raspodjelom. Predložićemo dvije test statistike i primjenjujući Monte Karlo metodu izračunaćemo kritične vrijednosti za različite veličine uzoraka.

Koristim priliku da izrazim neizmjernu zahvalnost mentoru, profesoru Božidaru Popoviću, na prijateljskoj saradnji, bezrezervnoj podršci, veoma korisnim savjetima i sugestijama, kako u nastavnom tako i u naučnom radu i naravno pri izradi ove disertacije.

Beskrajno sam zahvalna svojim roditeljima Dragoju i Vesni i sestri Andrijani, na nemjerljivoj i svakodnevnoj pažnji i podršci. Ovaj rad posvećujem njima.

Izvod iz teze

Cilj istraživanja jeste proučavanje linearne kombinacije logističkih slučajnih veličina. Razmatramo metode za numeričko izračunavanje funkcije gustine, funkcije raspodjele i kvantilne funkcije linearne kombinacije n nezavisnih logistički raspođijeljenih slučajnih veličina. Korišćenjem Melinove i inverzne Melinove transformacije pokazali smo da su funkcija raspodjele i funkcija gustine izražene u obliku Foksove H funkcije. Takođe, implementiramo algoritam za numeričko izračunavanje Foksove H funkcije i upoređujemo dobijene rezultate sa algoritmom koji je zasnovan na metodama numeričke inverzije karakteristične funkcije. Ilustrovali smo primjenu razmatrane linearne kombinacije koristeći odgovarajuće podatke iz finansija.

Dalji rad podrazumijeva nalaženje raspodjele slučajne veličine $Z_1 = aX + bY$, $a, b \in \mathbb{R}$, u slučaju kada su X i Y zavisne logističke slučajne veličine. Cilj je da nađemo metod za konstrukciju dvodimenzionalne logističke raspodjele koja će nam omogućiti proučavanje raspodjele slučajne veličine Z_1 . U tom cilju posmatrali smo karakterističnu funkciju

$$\phi(t_1, t_2) = 2e^{i(t_1 m_1 + t_2 m_2)} B(1 - i\sigma_1 t_1, 1 - i\sigma_2 t_2, 1 + i\sigma_1 t_1 + i\sigma_2 t_2),$$

gdje je i imaginarna jedinica, $m_1, m_2, t_1, t_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$, B je beta funkcija sa tri promjenljive definisana sa

$$B(a_1, a_2, a_3) = \frac{\Gamma(a_1)\Gamma(a_2)\Gamma(a_3)}{\Gamma(a_1 + a_2 + a_3)},$$

gdje je

$$\Gamma(z) = \int_0^{+\infty} e^{-t} t^{z-1} dz,$$

standardna gama funkcija. Nakon toga slijedi primjena metoda za invertovanje odgovarajuće karakteristične funkcije u cilju dobijanja gustine i funkcije raspodjele slučajnog vektora. Gustina slučajnog vektora (X_1, X_2) , gdje $X_i \sim \text{Log}(m_i, \sigma_i)$, $m_i \in \mathbb{R}$, $\sigma_i > 0$,

$i = \overline{1, 2}$ data je izrazom

$$\frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-i(t_1 z_1 + t_2 z_2)} \phi(t_1, t_2) dt_1 dt_2.$$

Budući da problem nije moguće riješiti analitički, predlažemo numerički algoritam za inverziju dvodimenzionalne karakteristične funkcije. Na taj način, u nastavku ćemo formirati algoritam za numeričko izračunavanje vrijednosti kopula funkcije. Takođe ćemo predložiti algoritam za generisanje slučajnih brojeva za slučaj kada je dvodimenzionalna raspodjela određena svojom karakterističnom funkcijom. Ovaj algoritam će se zasnivati na uslovnoj karakterističnoj funkciji i taj pristup se prvi put javlja u literaturi.

Bavićemo se primjenom empirijskih karakterističnih funkcija u testovima saglasnosti sa dvodimenizionalnom logističkom raspodjelom koja je specifikovana svojom karakterističnom funkcijom. U tu svrhu, predložićemo dvije test statistike i primjenom Monte Karlo simulacije odredićemo kritične vrijednosti za različite veličine uzoraka. Takođe, sprovedena je simulaciona studija da bi se uporedila moć predloženih test statistika.

Abstract

The subject of our consideration is to study the linear combination of logistic random variables. We consider methods for numerical evaluation of the probability density function, cumulative distribution function and quantile function of a linear combination of n independent generalized logistic random variables. The exact cumulative distribution function and probability density function are derived by using Mellin and inverse Mellin transforms and they are given via the well known Fox H function. We will also implement the algorithm for the numerical calculation of the Fox-H function and compare the obtained results with the algorithm based on the methods of numerical inversion of the characteristic function. We study the application of the considered linear combination in the field of financial returns. Further work involves finding the random distribution quantities $Z_1 = aX + bY$, in the case when X and Y are dependent logistic random variables. The aim is to find a method for the construction of a bivariate logistic distribution that will enable us to study the distribution of the random variable Z_1 . To that end, we observed the characteristic function

$$\phi(t_1, t_2) = 2e^{i(t_1 m_1 + t_2 m_2)} B(1 - i\sigma_1 t_1, 1 - i\sigma_2 t_2, 1 + i\sigma_1 t_1 + i\sigma_2 t_2),$$

where i is imaginary unit, $m_1, m_2, t_1, t_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$, B is beta function of three variables defined by

$$B(a_1, a_2, a_3) = \frac{\Gamma(a_1)\Gamma(a_2)\Gamma(a_3)}{\Gamma(a_1 + a_2 + a_3)},$$

where

$$\Gamma(z) = \int_0^{+\infty} e^{-t} t^{z-1} dt,$$

is the standard gamma function. This is followed by the application of methods for numerical inverting the corresponding characteristic functions in order to obtain the density and distribution function of the random vector. Joint density function of random

vector (X_1, X_2) , $X_i \sim \text{Log}(m_i, \sigma_i^2)$, $m_i \in \mathbb{R}$, $\sigma_i > 0$, $i = \overline{1, 2}$ is given with

$$\frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-i(t_1 z_1 + t_2 z_2)} \phi(t_1, t_2) dt_1 dt_2.$$

Since the problem cannot be solved analytically, we propose a numerical algorithm for the inversion of the bivariate characteristic function. Subsequently, it will be possible to create numerical algorithms for a copula function. We will also propose an algorithm for generating random numbers for the case where the bivariate distribution is specified by its characteristic function. This algorithm will be based on the conditional characteristic function and this approach appears for the first time in the literature.

We will propose a goodness-of-fit test based on the empirical distribution function for the bivariate Logistic distribution. For this purpose, we will propose two test statistics and calculate critical values using Monte Carlo simulation. Also, a simulation study is conducted to compare the power values of the proposed test statistics.

Sadržaj

Predgovor	i
Izvod iz teze	iii
Abstract	v
1 Uvod	1
1.1 Slučajne veličine: osnovni pojmovi	1
1.1.1 Vjerovatnosni prostor, događaji, nezavisnost	1
1.1.2 Slučajne veličine i njihove raspodjele	2
1.1.3 Očekivanje	7
1.1.4 Disperzija	8
1.1.5 Kovarijansa, koeficijent korelacije	10
1.1.6 Funkcija generatriše momenata	11
1.1.7 Karakteristična funkcija	12
1.1.8 Vrste konvergencija u teoriji vjerovatnoća	15
1.2 Kopula funkcija	16
1.3 Mjerenje zavisnosti	19
1.3.1 Saglasnost mjera	19
1.3.2 Kendalov tau	19
1.3.3 Spirmanov koeficijent korelacije	20
1.4 Uvod u statistiku	21
1.4.1 Ocjenjivanje parametara	23
1.4.2 Metod momenata	24
1.4.3 Metod maksimalne vjerodostojnosti	26
1.4.4 Intervalno ocjenjivanje	27
1.4.5 Testiranje statističkih hipoteza	28
1.5 Primjena butstrap metode u testu Kolmogorov - Smirnova	31
1.5.1 Butstrap metoda	32
1.5.2 K - S test zasnovan na butstrap metodu	33
1.6 Specijalne funkcije	35
1.6.1 Gama funkcija	36
1.6.2 Gausova hipergeometrijska funkcija	38
1.6.3 H-funkcija	39
1.6.4 Specijalni slučajevi H-funkcije	41

1.7	Furijeova i Melinova transformacija	42
1.8	Numeričke metode	44
1.8.1	Numeričko diferenciranje: metod konačnih razlika	45
1.8.2	Metod kvadrature	47
1.8.3	Baricentrična interpolacija	48
1.8.3.1	Langranžova i Njutnova interpolacija	48
1.8.3.2	Poboljšana Langranžova formula	50
1.8.3.3	Baricentrična formula	52
1.8.3.4	Čebišovljeve tačke	52
1.9	Singularna dekompozicija	54
1.9.1	Singularne vrijednosti	55
1.9.2	Definicija singularne dekompozicije	56
1.10	Jednodimenzionalna logistička slučajna veličina	57
2	Dosadašnji rezultati	60
3	Raspodjela dvije nezavisne logistički raspodijeljene slučajne veličine	68
3.1	Algoritam za računanje vrijednosti Foksove H funkcije	74
4	Raspodjela linearne kombinacije ($n > 2$) nezavisnih logistički raspodijeljenih slučajnih veličina	79
4.1	Raspodjela linearne kombinacije n nezavisnih logistički raspodijeljenih slučajnih veličina primjenom Foksove H funkcije	80
4.2	Raspodjela linearne kombinacije n nezavisnih slučajnih veličina. Primjena metoda za numeričku inverziju karakteristične funkcije	83
4.3	Uprođivanje numeričke efikasnosti algoritama	85
4.4	Funkcija raspodjele, funkcija gustine i kvantilna funkcija linearne kombinacije nezavisnih logistički raspodijeljenih slučajnih veličina	88
5	Numerička inverzija dvodimenzionalne logističke raspodjele	90
5.1	Karakteristična funkcija dvodimenzionalne logističke raspodjele	90
5.2	Uslovna karakteristična funkcija	93
5.3	Algoritam za numeričku inverziju dvodimenzionalne karakteristične funkcije	96
5.4	Algoritam za generisanje slučajnih brojeva	100
5.5	Kopula funkcija, Spirmanov ρ i Kendalov tau za dvodimenzionalnu logističku raspodjelu	102
5.6	Dvije zavisne logistički raspodijeljene slučajne veličine	107
6	Primjena na realnim podacima	109
6.1	Primjer 1	110
6.2	Primjer 2	112
7	Novi testovi saglasnosti sa dvodimenzionalnom logističkom raspodjelom	116
7.1	Kritične vrijednosti	125

A	Algoritam za Foksovu H funkciju	130
B	Tablice	133
	Literatura	138
	Biografija	142

Glava 1

Uvod

1.1 Slučajne veličine: osnovni pojmovi

U ovoj glavi dajemo pregled teorije pojmova i rezultata iz teorije vjerovatnoće, statistike i numeričke analize, koje ćemo koristiti u radu. Ove teme su detaljno obrađene u veoma bogatoj literaturi, a posebno treba istaći udžbenike [40, 31, 24, 30, 38].

1.1.1 Vjerovatnosni prostor, događaji, nezavisnost

Vjerovatnosni prostor je matematički model svakodnevnih situacija ili procesa sa poznatim skupom mogućih ishoda. Nepoznato nam je koji će se ishod desiti ali mi definišemo vjerovatnoće realizacija određene grupe ishoda.

Vjerovatnosni prostor čine tri komponente: prostor ishoda Ω , sigma - polje događaja \mathcal{A} i vjerovatnosna funkcija \mathbb{P} . Prostor ishoda Ω je skup ishoda koji se mogu desiti u procesu koji modelujemo. Određenim podskupovima od Ω , koje nazivamo događajima, dodjeljujemo vjerovatnoću da dobijemo odgovarajući ishod. Drugim riječima, za podskup $A \subset \Omega$, vjerovatnoća $\mathbb{P}(A)$ je broj koji opisuje šansu da dobijemo ishod iz A . Ako se ostvari ishod koji odgovara događaju A , konstatujemo da se događaj A realizovao, a u suprotnom da se nije realizovao.

Primijetimo da se vjerovatnoća ne dodjeljuje svim podskupovima prostora ishoda. Neka \mathcal{A} označava kolekciju svih događaja (podskupovi kojima je vjerovatnoća dodijeljena). Onda vjerovatnoća je funkcija $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$.

Neka je uređena trojka $(\Omega, \mathcal{A}, \mathbb{P})$ mjerljiv prostor. Tada vrijedi sljedeće.

Prvo, \mathcal{A} je sigma - polje događaja, odnosno a) cio prostor Ω je događaj; b) ako je A događaj, tada je i komplement događaja A , $\overline{A} := \Omega \setminus A$ događaj; c) bilo koja konačna ili prebrojiva unija događaja je događaj.

Drugo, \mathbb{P} je vjerovatnosna mjera. Osnovno svojstvo vjerovatnosne mjere je prebrojiva aditivnost - što znači da za bilo koji konačan ili prebrojiv skup disjunktih događaja imamo

$$\mathbb{P}\left(\bigcup_j A_j\right) = \sum_j \mathbb{P}(A_j).$$

Vjerovatnoća je normirana tj. $\mathbb{P}(\Omega) = 1$.

Događaji A i B su nezavisni ako je

$$\mathbb{P}(AB) = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

Definicija 1.1. Za dati prebrojivi ili konačni skup indeksa I , kažemo da je kolekcija događaja $C = \{A_i : i \in I\}$ u parovima nezavisna ako za svako $i, j \in I$, $\mathbb{P}(A_i A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$.

1.1.2 Slučajne veličine i njihove raspodjele

U stvarnom životu, često nas ne zanimaju toliko ishodi ω iz prostora Ω , koliko numerički rezultati koji zavise od ω . Na primjer, mnogi kockari su više zabrinuti za svoje gubitke nego za igre koje dovode do njih. Takva posljedica, vodi do prirodnog prelaska sa skupa Ω na skup realnih brojeva \mathbb{R} . Taj prelazak se ostvaruje uz posredovanje funkcije koja se naziva slučajna veličina.

Definicija 1.2. Vektorski prostor V je euklidski vektorski prostor ako je V vektorski prostor nad \mathbb{R} , i na njemu je definisano preslikavanje $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ takvo da je

1. za sve $u, v \in V$ važi $\langle u, v \rangle = \langle v, u \rangle$;
2. za sve $u, v, w \in V$ i $\alpha, \beta \in \mathbb{R}$ važi $\langle \alpha u + \beta v, w \rangle = \alpha \langle u, w \rangle + \beta \langle v, w \rangle$;
3. za sve $u \in V$ važi da je $\langle u, u \rangle \geq 0$ i $\langle u, u \rangle = 0$ ako i samo ako je $u = 0$.

Ovakvo preslikavanje zovemo skalarni proizvod.

Neka je \mathbb{R} skup realnih brojeva, dok \mathbb{R}^n označava Euklidov prostor dimenzije n . Sa \mathcal{B}^n označićemo Borelovo sigma polje, najmanje sigma polje podskupova od \mathbb{R}^n , koje sadrži sve otvorene skupove. Takođe pišemo $\mathcal{B} := \mathcal{B}^1$.

Neka su (Ω, \mathcal{F}) i (Λ, \mathcal{G}) mjerljivi prostori i f je funkcija iz Ω u Λ . Funkcija f se naziva mjerljiva funkcija iz (Ω, \mathcal{F}) u (Λ, \mathcal{G}) ili \mathcal{F} - mjerljiva funkcija ako i samo ako je $f^{-1}\mathcal{G} \subset \mathcal{F}$.

Ako je $\Lambda = \mathbb{R}$ i $\mathcal{G} = \mathcal{B}$ (Borelovo σ polje), tada se f naziva Borelova mjerljiva funkcija ili Borelova funkcija. Jasno je da je svaka neprekidna funkcija i Borelova funkcija.

Slučajna veličina je bilo koja mjerljiva funkcija $X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B})$. Ovdje mjerljiv znači da za bilo koji skup $B \in \mathcal{B}$ imamo da

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}.$$

Raspodjela od X je mjera na $(\mathbb{R}, \mathcal{B})$ definisana sa

$$P_X(B) := \mathbb{P}\{X \in B\}, \quad B \in \mathcal{B}.$$

Definicija 1.3. *Funkcija raspodjele slučajne veličine X (ili kraće funkcija raspodjele) je funkcija $F_X = F : \mathbb{R} \rightarrow [0, 1]$ zadana sa*

$$F_X(r) := P_X(-\infty, r] = \mathbb{P}\{X \leq r\}, \quad r \in \mathbb{R}.$$

Funkcija raspodjele može se definisati i kao $F_X(r) = P\{X < r\}$. Slučajne veličine su jednako raspodijeljene ako se njihove raspodjele poklapaju.

Slučajnoj veličini X pridružujemo sigma - polje

$$\mathcal{A}_X := \{X^{-1}(B), B \in \mathcal{B}\} \subset \mathcal{A}.$$

Znajući samo vrijednost od X , možemo utvrditi da li se realizovao neki događaj iz \mathcal{A}_X .

Slučajne veličine X_1, X_2, \dots, X_n su nezavisne ako su bilo koji događaji

$$A_1 \in \mathcal{A}_{X_1}, A_2 \in \mathcal{A}_{X_2}, \dots, A_n \in \mathcal{A}_{X_n}$$

nezavisni. Drugim riječima, za bilo koje skupove $B_1, B_2, \dots, B_n \in \mathcal{B}$ događaji $\{X_1 \in B_1\}, \{X_2 \in B_2\}, \dots, \{X_n \in B_n\}$ moraju biti nezavisni.

U vjerovatnoći je uobičajeno da ne pravimo razliku između slučajnih veličina X i Y takvih da $\mathbb{P}(\omega : X(\omega) \neq Y(\omega)) = 0$. Štaviše, prilikom definisanja slučajne veličine,

dozvoljeno je da se ostavi nedefinisana na skupu (događaju) čija je vjerovatnoća realizacije jednaka 0. Ovakva nepreciznost, nema uticaja na raspodjelu slučajne veličine niti na bilo koju njenu vjerovatnosnu osobinu.

U primjenama najčešće srećemo diskretne ili apsolutno neprekidne raspodjele.

Raspodjelu slučajne veličine X kao i samu slučajnu veličinu nazivamo diskretnom ako je njen skup realizacija (kodomen) konačan ili prebrojiv. Neka je $W = \{x_i : i \in I\}$, skup indeksa I je konačan ili prebrojiv, skup realizacija slučajne veličine X . U ovom slučaju raspodjela slučajne veličine X je data vjerovatnoćama $p_j := \mathbb{P}(X = x_j)$, $j \in I$, koje moraju zadovoljavati jednakost

$$\sum_{j \in I} p_j = 1.$$

Neka je X slučajna veličina diskretnog tipa koja ima realizacije a_1, a_2, \dots , tako da

$$p(a_i) > 0, \quad p(a_1) + p(a_2) + p(a_3) + \dots = 1,$$

tada

$$F(a) = \sum_{a_i < a} p(a_i).$$

Primijetimo da je raspodjela slučajne veličine X određena funkcijom raspodjele $F(a)$. Naime, tačke a_i u kojima funkcija $F(a)$ ima skok su vrijednosti slučajne veličine, a visina skoka jednaka je vjerovatnoći sa kojom se uzima odgovarajuća vrijednost, tj. jednaka je $p(a_i)$.

Navešćemo osobine funkcije raspodjele:

1. $F(x)$ je monotonu neopadajuća funkcija.
2. $F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0$, $F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1$.
3. $F(x)$ je neprekidna sa desne strane u svim tačkama, tj. $\forall x \in \mathbb{R}, F(x+0) = F(x)$.

Raspodjelu slučajne veličine X nazivamo apsolutno - neprekidnom ako postoji gustina raspodjele $f_X(\cdot)$, tj. nenegativna, integrabilna funkcija na \mathbb{R} , takva da

$$\mathbb{P}(X \in B) = P_X(B) = \int_B f_X(r) dr, \quad B \in \mathcal{B}.$$

Jasno je da

$$\int_{-\infty}^{+\infty} f_X(r) dr = \mathbb{P}(X \in \mathbb{R}) = \mathbb{P}(\Omega) = 1.$$

Postoji jednostavna veza između funkcije raspodjele i funkcije gustine kod slučajnih veličina apsolutno-neprekidnog tipa. Neka je X slučajna veličina apsolutno-neprekidnog tipa zadata sa funkcijom gustine $f_X(x)$. Tada iz integralnog računa slijedi da je

$$F(b) = P\{X \leq b\} = \int_{-\infty}^b f_X(x) dx \text{ i } f_X(x) = \frac{dF(x)}{dx}.$$

Pretpostavimo sada da je T slučajna veličina apsolutno neprekidnog tipa sa funkcijom gustine $f(t)$ i funkcijom raspodjele $F_T(t) = P\{T \leq t\}$. Često je za rad pogodnije posmatrati komplement od funkcije raspodjele, funkciju preživljavanja, definisanu sa

$$\overline{F}_T(t) = P\{T > t\} = 1 - F_T(t) = \int_t^{\infty} f(x) dx. \quad (1.1)$$

U daljem izlaganju važnu ulogu će imati uopšteni inverz monotonihi funkcija, koje nijesu obavezno strogo monotone. Ovdje ćemo navesti definiciju uopštenog inverza monotono neopadajuće funkcije.

Definicija 1.4. Neka je $F : \mathbb{R} \rightarrow \mathbb{R}$ proizvoljna neopadajuća funkcija i $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$. Uopšteni inverz funkcije F je funkcija $F^{-1} : \mathbb{R} \rightarrow \overline{\mathbb{R}}$, definisana pomoću nejednakosti

$$F^{-1}(y) = \inf\{x : F(x) \geq y\},$$

pri čemu je $\inf\{\} = +\infty$.

Za funkciju raspodjele F , inverzna funkcija raspodjele F^{-1} definisana je kao

$$F^{-1}(y) = \inf\{x : F(x) \geq y\}, \quad y \in [0, 1].$$

Funkcija F^{-1} zove se još i kvantilna funkcija.

Definicija 1.5. Neka je X slučajna veličina apsolutno - neprekidnog tipa i neka je p broj između 0 i 1. Kvantil reda p ili 100% p percentil od raspodjele slučajne veličine X je neki broj $q_p \in \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$ takav da $P\{X \leq q_p\} \geq p$ i $P\{X \geq q_p\} \geq 1 - p$.

Raspodjela vektora slučajih veličina X_1, X_2, \dots, X_n , koje su definisane na istom vjerovatnosnom prostoru je mjera P_{X_1, X_2, \dots, X_n} na $(\mathbb{R}^n, \mathcal{B}^n)$ data sa

$$P_{X_1, X_2, \dots, X_n}(A) = \mathbb{P}\{(X_1, X_2, \dots, X_n) \in A\}, \quad A \in \mathcal{B}^n.$$

Funkcija raspodjele kolekcije slučajnih veličina X_1, X_2, \dots, X_n definisana je sa

$$F(x_1, x_2, \dots, x_n) = P\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\}.$$

U mnogim primjenama teorije vjerovatnoće dešava se da je data raspodjela neke slučajne veličine X i da treba odrediti raspodjelu slučajne veličine $g(X)$, ili je potrebno odrediti raspodjelu slučajne veličine $g(X_1, X_2, \dots, X_n)$ kada je poznata raspodjela slučajnog vektora (X_1, X_2, \dots, X_n) , gdje je g neka funkcija.

Teorema 1.1 ([31]). *Neka su X_1, X_2, \dots, X_n slučajne veličine sa zajedničkom funkcijom gustine $f(x_1, \dots, x_n)$ i neka su slučajne veličine Y_1, Y_2, \dots, Y_n definisane sa*

$$Y_1 = g_1(X_1, \dots, X_n), Y_2 = g_2(X_1, \dots, X_n), \dots, Y_n = g_n(X_1, \dots, X_n),$$

gdje su g_1, g_2, \dots, g_n diferencijabilne funkcije n promjenljivih. Pretpostavimo da sistem jednačina

$$y_1 = g_1(x_1, \dots, x_n), y_2 = g_2(x_1, \dots, x_n), \dots, y_n = g_n(x_1, \dots, x_n),$$

ima jedno ili više, recimo m , rješenja po x_1, \dots, x_n ,

$$x_1 = h_{i_1}(y_1, \dots, y_n), x_2 = h_{i_2}(y_1, \dots, y_n), \dots, x_n = h_{i_n}(y_1, \dots, y_n),$$

gdje je $i = 1, 2, \dots, m$. Neka postoje neprekidni prvi parcijalni izvodi $\frac{\partial h_{i_j}}{\partial y_j}$, $j = 1, 2, \dots, n$. Označimo sa J_1, \dots, J_n Jakobijane ovih m transformacija,

$$J_i(y_1, \dots, y_n) = \begin{vmatrix} \frac{\partial h_{i_1}}{\partial y_1} & \frac{\partial h_{i_1}}{\partial y_2} & \dots & \frac{\partial h_{i_1}}{\partial y_n} \\ \vdots & \ddots & \dots & \vdots \\ \frac{\partial h_{i_n}}{\partial y_1} & \frac{\partial h_{i_n}}{\partial y_2} & \dots & \frac{\partial h_{i_n}}{\partial y_n} \end{vmatrix}.$$

Tada je funkcija gustine za (Y_1, \dots, Y_n) data sa

$$f_{Y_1, \dots, Y_n}(y_1, \dots, y_n) = \sum_{i=1}^m f(h_{i_1}(y_1, \dots, y_n), \dots, h_{i_n}(y_1, \dots, y_n)) |J_i(y_1, \dots, y_n)|.$$

1.1.3 Očekivanje

Neka je X slučajna veličina definisana na vjerovatnosnom prostoru $(\Omega, \mathcal{A}, \mathbb{P})$. Ako je integral $\int_{\Omega} |X(\omega)| \mathbb{P}(d\omega)$ konačan, onda očekivanjem od X nazivamo vrijednost

$$\mathbb{E}X := \int_{\Omega} X(\omega) \mathbb{P}(d\omega).$$

Konkretno, ako je P_X diskrentno, onda je

$$\mathbb{E}X := \sum_x x \cdot \mathbb{P}\{X = x\},$$

ako red apsolutno konvergira. Ako red ne konvergira apsolutno, tada kažemo da slučajna veličina X nema matematičko očekivanje.

Navešćemo neke osobine očekivanja.

- Očekivanje od konstante: $\mathbb{E}(c) = c$.
- Linearnost: $\mathbb{E}(c_1X + c_2Y) = c_1\mathbb{E}X + c_2\mathbb{E}Y$;
Pravilo sumiranja može se proširiti na konačan broj sabiraka. Specijalan slučaj je, $\mathbb{E}(X + c) = \mathbb{E}X + c$.
- Monotonost:

Ako je $\mathbb{P}(X \geq 0) = 1$, tada je $\mathbb{E}(X) \geq 0$.

Slijedi da

$$\mathbb{P}(X \geq Y) = 1 \implies \mathbb{E}X \geq \mathbb{E}Y$$

(ako su sva očekivanja dobro definisana).

- Neka su X_1, X_2, \dots, X_n slučajne veličine definisane na istom vjerovatnosnom prostoru i neka je $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^1$, $x \in \mathbb{R}^n$, mjerljiva funkcija. Onda

$$\mathbb{E}f(X_1, X_2, \dots, X_n) = \int_{\mathbb{R}^n} f(x) P_{X_1, X_2, \dots, X_n}(dx).$$

Konkretno, za jednu slučajnu veličinu $n = 1$ imamo

$$\mathbb{E}f(X) = \int_{\mathbb{R}^1} f(x) P_X(dx).$$

Stavljajući da je $f(x) = x$, imamo da

$$\mathbb{E}X = \int_{\mathbb{R}^1} x P_X(dx).$$

Konkretno, kada P_X ima gustinu $p(\cdot)$, tada je

$$\mathbb{E}X = \int_{-\infty}^{+\infty} x p(x) dx.$$

Primijetimo da $\mathbb{E}X$ je dobro definisano akko je $\int_{\mathbb{R}^1} |x| P_X(dx) < \infty$. Naglasimo da egzistencija i vrijednost od $\mathbb{E}X$ je u potpunosti određena raspodjelom od X . Konkretno, jednako raspodjeljene slučajne veličine imaju isto očekivanje i

- Pravilo proizvoda za nezavisne slučajne veličine. Pretpostavimo da slučajne veličine X_1, X_2, \dots, X_n su nezavisne i da su njihova očekivanja dobro definisana. Onda njihov proizvod takođe ima dobro definisano očekivanje i

$$\mathbb{E} \left(\prod_{j=1}^n X_j \right) = \prod_{j=1}^n (\mathbb{E}X_j).$$

1.1.4 Disperzija

Disperzija slučajne veličine opisuje mjeru njene raspršenosti oko očekivanja. Definisana je sa

$$\mathbb{D}X = \mathbb{E}(X - \mathbb{E}X)^2.$$

Alternativna formula za disperziju je:

$$\mathbb{D}X = \mathbb{E}(X^2) - (\mathbb{E}X)^2.$$

U nastavku ćemo navesti osobine disperzije.

- Disperzija slučajne veličine je određena njenom raspodjelom,

$$\mathbb{D}X = \int_{\mathbb{R}} x^2 P_X(dx) - \left(\int_{\mathbb{R}} x P_X(dx) \right)^2.$$

Konkretno, jednako raspodijeljene slučajne veličine imaju jednaku disperziju.

- Nenegativnost: $\mathbb{D}X \geq 0$.
- $\mathbb{D}(c) = 0$, $\mathbb{D}(X + c) = \mathbb{D}X$.
- 2-homogenost:

$$\mathbb{D}(cX) = c^2 \mathbb{D}X, \quad \forall c \in \mathbb{R}.$$

Konkretno, $\mathbb{D}(-X) = \mathbb{D}X$.

U cilju da dobijemo 1-homogenost mjere raspršenosti (disprzije) možemo razmatrati standardnu devijaciju definisanu sa

$$\sigma(X) := \sqrt{\mathbb{D}X}.$$

Tada je naravno, $\sigma(cX) = |c|\sigma(X)$. Jedna od prednosti standardne devijacije je ta što, zajedno sa očekivanjem, izražena je u istoj jedinici mjere kao i sami elementi raspodjele.

- Pravilo zbira. Slučajne veličine X, Y , koje zadovoljavaju

$$\mathbb{E}X^2 < \infty, \quad \mathbb{E}Y^2 < \infty$$

su nekorelisane, ako je $\mathbb{E}(XY) = \mathbb{E}X \cdot \mathbb{E}Y$. Slijedi da su nezavisne slučajne veličine nekorelisane.

Ako su slučajne veličine X, Y nekorelisane, onda imamo pravilo zbira

$$\mathbb{D}(X + Y) = \mathbb{D}X + \mathbb{D}Y.$$

Isto tvrđenje važi i za sumu konačnog broja u parovima nekorelisanih slučajnih veličina.

Konkretno, dobijamo pravilo sumiranja za disperziju nezavisnih slučajnih veličina: ako su X_1, X_2, \dots, X_n u parovima nezavisne slučajne veličine, $\mathbb{E}X_1^2 < \infty$, $\mathbb{E}X_2^2 < \infty, \dots, \mathbb{E}X_n^2 < \infty$, onda

$$\mathbb{D}\left(\sum_{j=1}^n X_j\right) = \sum_{j=1}^n \mathbb{D}(X_j).$$

Pogodnost pravila sumiranja daje disperziji odlučujuću prednost u odnosu na ostale mjere raspršenosti kao što su standardna devijacija ili $\mathbb{E}|X - \mathbb{E}X|$.

1.1.5 Kovarijansa, koeficijent korelacije

Kovarijansa dvije slučajne veličine X, Y je definisana sa

$$\text{cov}(X, Y) := \mathbb{E}((X - \mathbb{E}X)(Y - \mathbb{E}Y)) = \mathbb{E}(XY) - \mathbb{E}X\mathbb{E}Y,$$

ako su sva očekivanja dobro definisana. Konkretno, kovarijansa je dobro definisana kada je $\mathbb{D}X, \mathbb{D}Y < \infty$.

Navešćemo osnovna svojstva kovarijanse.

- Simetrija: $\text{cov}(X, Y) = \text{cov}(Y, X)$.
- Bilinearnost:

$$\begin{aligned}\text{cov}(cX, Y) &= c \cdot \text{cov}(X, Y), \\ \text{cov}(X_1 + X_2, Y) &= \text{cov}(X_1, Y) + \text{cov}(X_2, Y).\end{aligned}$$

- Povezanost sa disperzijom: $\text{cov}(X, X) = \mathbb{D}X$.
- Ako su slučajne veličine X, Y nekorelisane, onda $\text{cov}(X, Y) = 0$. Konkretno, ako je $\mathbb{E}X^2 < \infty, \mathbb{E}Y^2 < \infty$ i slučajne veličine X, Y su nezavisne, onda $\text{cov}(X, Y) = 0$.

Zbog posljednjeg tvrđenja, kovarijansa se koristi za mjerenje linearne zavisnosti između slučajnih veličina. Preciznije, odgovarajuća mjera za linearnu zavisnost je normirana veličina

$$\rho(X, Y) := \frac{\text{cov}(X, Y)}{\sqrt{\mathbb{D}X}\sqrt{\mathbb{D}Y}}, \quad (1.2)$$

koju zovemo koeficijent korelacije.

Prisjetimo se da za slučajne veličine X, Y uslov $\text{cov}(X, Y) = 0$ znači da su X i Y nekorelisane. U slučaju nenulte vrijednosti za kovarijansu slučajne veličine se nazivaju pozitivno, odnosno negativno korelisane, u zavisnosti da li kovarijansa ima pozitivnu ili negativnu vrijednost, respektivno. Osobine koeficijenta korelacije su sljedeće:

- normalizacija: za bilo koje slučajne veličine X, Y važi

$$-1 \leq \rho(X, Y) \leq 1.$$

- 0-homogenost:

$$\rho(cX + a, Y) = \begin{cases} \rho(X, Y), & c > 0, \\ -\rho(X, Y), & c < 0. \end{cases}$$

- Ekstremne vrijednosti: za bilo koju slučajnu veličinu X , takvu da $\mathbb{E}X^2 < \infty$, važi da je $\rho(X, X) = 1$, $\rho(X, -X) = -1$.
Posljedično, ako je $\rho(X, Y) = 1$, onda se X i Y razlikuju do na konstantu, tj. $Y = cX + a$, za neko $a \in \mathbb{R}$ i $c > 0$. Slično $\rho(X, Y) = -1$ vodi do $Y = cX + a$, za neko $a \in \mathbb{R}$ i $c < 0$.

Definicija 1.6. Neka je $X = (X_1, X_2, \dots, X_n)$ slučajni vektor takav da za sve $k \in \{1, 2, \dots, n\}$ važi $\mathbb{E}(X_k^2) < +\infty$. Matrica $C = [c_{ij}]_{n \times n}$, gdje je $c_{ij} = \text{cov}(X_i, X_j)$ naziva se kovarijaciona matrica slučajnog vektora X .

Jasno je da je kovarijaciona matrica C simetrična $n \times n$ matrica koja na dijagonali ima elemente $\mathbb{D}(X_k)$, $k \in \{1, 2, \dots, n\}$, a van dijagonale (za $i \neq j$) elemente $\text{cov}(X_i, X_j)$.

1.1.6 Funkcija generatriše momenata

Funkcija generatriše momenata je veoma moćan alat koji se koristi u računarskoj statistici. U nastavku dajemo kratak teorijski pregled koji se može pronaći u [38].

Funkcija generatriše momenata služi za određivanje momenata k -tog reda, $\mathbb{E}(X^k)$, $k \in \mathbb{N}$, slučajne veličine X pod uslovom da oni postoje. Funkcija generatriše momenata definiše se i za diskretne i za apsolutno neprekidne slučajne veličine.

Definicija 1.7. Neka je X slučajna veličina takva da postoji realan broj $h > 0$ takav da za svaki realan broj $t \in (-h, h)$ očekivanje $\mathbb{E}(e^{tX})$ postoji. Funkcija $M_X(t) = \mathbb{E}(e^{tX})$ zove se funkcija generatriše momenata slučajne veličine X .

Iz definicije 1.7 možemo zaključiti da funkcija generatriše momenata neće postojati ukoliko očekivanje $\mathbb{E}(e^{tX})$ ne postoji u okolini vrijednosti $t = 0$. S druge strane, vrijednost $M_X(0)$ uvijek postoji i jednaka je 1. Može se pokazati da funkcija raspodjele $F(x)$ slučajne veličine X jednoznačno određuje funkciju generatriše momenata $M_X(t)$ ako ova postoji, i obratno.

Momenti slučajne veličine određuju se pomoću funkcije generatriše na sljedeći način.

Teorema 1.2 ([38]). *Ako postoji funkcija generatriše momenata slučajne veličine X , tada je moment k -tog reda slučajne veličine X konačan i jednak*

$$\mathbb{E}(X^k) = \left. \frac{d^k M_X(t)}{dt^k} \right|_{t=0}.$$

Funkcija generatriše momenata ima neke interesantne osobine.

Teorema 1.3 ([38]). *Ako postoji funkcija generatriše momenata $M_X(t)$ slučajne veličine X i ako su $a \neq 0$ i b realne konstante, tada je funkcija generatriše momenata slučajne veličine $aX + b$ jednaka*

$$M_{aX+b}(t) = e^{bt} M_X(at).$$

Teorema 1.4 ([38]). *Ako postoje funkcije generatriše momenata $M_X(t)$ i $M_Y(t)$ nezavisnih slučajnih veličina X i Y , tada je*

$$M_{X+Y}(t) = M_X(t)M_Y(t).$$

Prethodna teorema može se uopštiti i za konačan broj nezavisnih slučajnih veličina.

1.1.7 Karakteristična funkcija

Definicija 1.8. *Neka je X slučajna veličina čija je funkcija raspodjele $F_X(x)$. Funkcija*

$$\phi_X(t) := \mathbb{E}e^{itX} = \int_{-\infty}^{+\infty} e^{itx} dF_X(x) \text{ za } t \in \mathbb{R}, \text{ gdje je } i^2 = -1$$

naziva se karakteristična funkcija slučajne veličine X .

Primijetimo, $\phi : \mathbb{R} \rightarrow \mathbb{C}$ i funkcija raspodjele postoji za svaku slučajnu veličinu. Integral sa kojim je definisana karakteristična funkcija apsolutno konvergira odakle slijedi konvergencija samog integrala.

Osobine karakteristične funkcije su:

- $\phi_X(0) = 1$.
- $|\phi_X(t)| \leq 1$.
- $\phi_X(-t) = \overline{\phi_X(t)}$.

- Ako je raspodjela slučajne veličine X diskretnog tipa tj. $P\{X = x_j\} = p_j$, $j \in J$, tada je

$$\phi_X(t) = \sum_{j \in J} p_j e^{itx_j}.$$

- Ako je slučajna veličina X apsolutno neprekidnog tipa sa gustinom $f(x)$ tada je

$$\phi_X(t) = \int_{\mathbb{R}} e^{itx} f(x) dx.$$

Dakle, do na konstantu, karakteristična funkcija je Furijeova transformacija od gustine f .

- Ako su X i Y nezavisne, onda je $\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t)$. Konkretno, ako su X_1, X_2, \dots, X_n nezavisne i jednako raspodijeljene slučajne veličine i $S := \sum_{1 \leq j \leq n} X_j$, tada je $\phi_S(t) = \left(\phi_{X_1}(t)\right)^n$.
- Za bilo koje $a, c \in \mathbb{R}$ važi da je

$$\phi_{cX+a}(t) = e^{iat} \phi_X(ct).$$

- Ako postoji funkcija generatriše momenata $M_X(t)$ slučajne veličine X , tada se domen karakteristične funkcije može proširiti na kompleksnu ravan i

$$\phi(-it) = M_X(t).$$

- $\phi_X(\cdot)$ je realno vrijednosna ako i samo ako je raspodjela od X simetrična, tj. $P_X(B) = P_X(-B)$ za svako $B \in \mathcal{B}^1$.
- Funkcija $\phi : \mathbb{R} \rightarrow \mathbb{C}$ je nenegativno definitna, ako za svako $n \in \mathbb{N}$, $t_1, t_2, \dots, t_n \in \mathbb{R}$, $c_1, c_2, \dots, c_n \in \mathbb{C}$

$$\sum_{j,k=1}^n \phi(t_j - t_k) c_j \overline{c_k} \geq 0.$$

Karakteristična funkcija je nenegativno definitna. Zaista, kako je

$$\sum_{j,k=1}^n \phi(t_j - t_k) c_j \overline{c_k} = \sum_{j,k=1}^n \mathbb{E} e^{i(t_j - t_k)X} c_j \overline{c_k} = \mathbb{E} \left(\sum_{j,k=1}^n e^{it_j X} c_j e^{\overline{it_k X} \overline{c_k}} \right) =$$

$$= \mathbb{E} \left(\sum_{j=1}^n e^{it_j X} c_j \overline{\sum_{k=1}^n e^{it_k X} c_k} \right) = \mathbb{E} \left| \sum_{j=1}^n e^{it_j X} c_j \right|^2 \geq 0.$$

Preciznije, poznata Bohner-Hinčinova teorema tvrdi da ako je $\phi : \mathbb{R} \rightarrow \mathbb{C}$ nenegativno definitna, neprekidna i $\phi(0) = 1$, onda postoji slučajna veličina X takva da je $\phi_X = \phi$.

Često je lakše manipulirati karakterističnim funkcijama nego funkcijama raspodjele. Ako je karakteristična funkcija poznata onda možemo izračunati funkciju raspodjele korišćenjem teoreme o inverziji. U radu [41] proučavana je inverzija karakterističnih funkcija u cilju izračunavanja funkcije raspodjele. Teorija se zasniva na poznatim idejama konvolucija, Lebegovih integrala i Furijeovih inverzija.

Na primjer, ako je karakteristična funkcija ϕ_X integrabilna, onda raspodjela F_X ima gustinu $f(\cdot)$ i možemo rekonstruisati gustinu pomoću inverzne Furijeove transformacije

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \phi_X(t) dt. \quad (1.3)$$

U opštem slučaju sljedeća teorema daje formulu inverzije.

Teorema 1.5 (Formula inverzije, [24]). *Ako su $a, b \in \mathbb{R}$ takvi da je $a < b$ i $\mathbb{P}(X = a) = \mathbb{P}(X = b) = 0$, onda je*

$$P_X[a, b] = \frac{1}{2\pi} \lim_{T \rightarrow \infty} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} \phi_X(t) dt.$$

Sljedeća teorema povezuje karakterističnu funkciju i momente od X .

Teorema 1.6 ([24]). **a)** *Ako postoji $\phi_X^{(k)}(0)$ onda*
$$\begin{cases} \mathbb{E}|X^k| < \infty, & \text{ako } k \text{ je parno,} \\ \mathbb{E}|X^{k-1}| < \infty, & \text{ako } k \text{ je neparno.} \end{cases}$$

b) *Ako je $\mathbb{E}|X^k| < \infty$ onda*

$$\phi_X(t) = \sum_{j=0}^k \frac{\mathbb{E}(X^j)}{j!} (it)^j + o(t^k),$$

$$i \phi_X^{(k)}(0) = i^k \mathbb{E}(X^k).$$

Definicija 1.9. Neka je (X, Y) slučajni vektor čija je funkcija gustine $f_{XY}(x, y)$. Karakteristična funkcija slučajnog vektora (X, Y) data je sa

$$\phi_{XY}(u, v) = \mathbb{E}e^{i(uX+vY)} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{i(ux+vy)} f_{XY}(x, y) dx dy, \quad u, v \in \mathbb{R}. \quad (1.4)$$

Primijetimo da je $|\phi_{XY}(u, v)| \leq \phi_{XY}(0, 0) = 1$.

Lako se pokazuje da je

$$E(XY) = \frac{1}{i^2} \frac{\partial^2 \phi(t_1, t_2)}{\partial t_1 \partial t_2} \Big|_{t_1=0, t_2=0}. \quad (1.5)$$

Ako su X i Y nezavisne slučajne veličine, tada iz (1.4) dobijamo

$$\phi_{XY}(u, v) = \mathbb{E}(e^{iuX})\mathbb{E}(e^{ivY}) = \phi_X(u)\phi_Y(v).$$

Takođe,

$$\phi_X(u) = \phi_{XY}(u, 0) \text{ i } \phi_Y(v) = \phi_{XY}(0, v).$$

1.1.8 Vrste konvergencija u teoriji vjerovatnoća

Neka su X, X_1, X_2, \dots slučajne veličine definisane na istom prostoru vjerovatnoća. U teoriji vjerovatnoća definišu se četiri vrste konvergencija niza slučajnih veličina X_1, X_2, \dots ka slučajnoj veličini X :

- konvergencija u vjerovatnoći,
- skoro sigurna konvergencija,
- konvergencija u srednjem reda p , $0 < p < +\infty$,
- konvergencija u raspodjeli.

Za slučajne veličine X, X_1, X_2, \dots pretpostavimo da su definisane na istom vjerovatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$. Navedene vrste konvergencije imaju veoma važnu ulogu u statistici i stohastičkoj analizi. U nastavku daćemo samo definiciju skoro sigurne konvergencije (za više detalja pogledati u [31]).

Definicija 1.10 ([31]). *Niz slučajnih veličina $\{X_n\}$ konvergira skoro sigurno (skoro svuda ili sa vjerovatnoćom 1) ka slučajnoj veličini X ako važi*

$$P\{\omega : \lim_{n \rightarrow +\infty} X_n(x) = X(\omega)\} = 1.$$

1.2 Kopula funkcija

Riječ kopula je prvi put, u statističkom smislu, uveo Sklar (1959). Kopulu možemo posmatrati kao funkciju koja spaja ili povezuje (eng. couple) višedimenzionalne funkcije raspodjele sa njenim jednodimenzionalnim marginalnim raspodjelama [30].

Kopula funkcija ima sposobnost da pokaže oblik zavisnosti između dvije ili više slučajnih veličina i takođe je praktičan i efikasan metod u modelovanju opšte zavisnosti višedimenzionalnih podataka. Navešćemo samo neke od prednosti korišćenja kopula u modelovanju višedimenzionalnih raspodjela:

- fleksibilnost u biranju proizvoljnih marginalnih raspodjela i oblika zavisnosti;
- mogućnost proširenja na više od dvije slučajne veličine;
- mogućnost odvojene analize marginalnih raspodjela i zavisnih komponenti.

Proučavanje kopula funkcije i njenih uloga u vjerovatnoći, statistici i stohastičkim procesima je i dalje u fazi istraživanja.

Definicija 1.11 ([30]). *Dvodimenzionalna kopula C je funkcija $C : [0, 1]^2 \rightarrow [0, 1]$ koja je neopadajuća, neprekidna sa desne strane i zadovoljava sljedeće uslove:*

- 1) $C(u, 0) = 0 = C(0, v)$, za svako $u, v \in [0, 1]$;
- 2) $C(u, 1) = u$ i $C(1, v) = v$, za svako $u, v \in [0, 1]$;
- 3) Za svako $u_1, u_2, v_1, v_2 \in [0, 1]$ takve da $u_1 \leq u_2$ i $v_1 \leq v_2$,

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0.$$

Sljedeća teorema ima centralnu ulogu u teoriji kopula i ona razjašnjava povezanost između višedimenzionalnih funkcija raspodjela i njenih jednodimenzionalnih marginalnih raspodjela. Poznata je kao Sklarova teorema.

Teorema 1.7 ([30]). *Neka je H dvodimenzionalna funkcija raspodjele sa marginalnim funkcijama raspodjele F i G . Tada postoji kopula C takva da je za svako $x, y \in \mathbb{R}$*

ispunjeno

$$H(x, y) = C(F(x), G(y)). \quad (1.6)$$

Ako su F i G neprekidne, tada je C jedinstveno; u suprotnom, C je jedinstveno određeno na $\text{Ran}F \times \text{Ran}G$, gdje je $\text{Ran}F$ kodomen funkcije F . Obratno, ako je C kopula i F i G su funkcije raspodjele, onda funkcija H definisana sa (1.6) je dvodimenzionalna funkcija raspodjele sa marginalnim funkcijama raspodjele F i G .

Nameće se pitanje da li se izraz (1.6) može invertovati tako da se kopula predstavi preko dvodimenzionalne funkcije raspodjele i inverza od marginala. Međutim, ako marginali nisu strogo rastuće funkcije, onda nemaju inverze. Zbog toga moramo prvo definisati kvazi-inverze od funkcija raspodjele.

Definicija 1.12 ([30]). Neka je F funkcija raspodjele. Tada kvazi-inverz od F je funkcija $F^{(-1)}$ sa domenom $[0, 1]$ takva da:

1) ako je $t \in \text{Ran}F$, tada $F^{(-1)}(t)$ je neki broj x iz \mathbb{R} takav da je $F(x) = t$, tj. za svako $t \in \text{Ran}F$,

$$F(F^{(-1)}(t)) = t;$$

2) ako $t \notin \text{Ran}F$, tada

$$F^{(-1)}(t) = \inf \{x | F(x) \geq t\} = \sup \{x | F(x) \leq t\}.$$

Ako je F strogo rastuća, tada ona ima jedan kvazi-inverz, koji je naravno uobičajen inverz, za koji koristimo standardnu oznaku F^{-1} .

Posljedica 1.1 ([30]). Neka je H dvodimenzionalna funkcija raspodjele sa neprekidnim marginalnim funkcijama raspodjele F i G , i neka su $F^{(-1)}$ i $G^{(-1)}$ kvazi-inverzi od F i G , respektivno. Tada za $(u, v) \in [0, 1] \times [0, 1]$,

$$C(u, v) = H(F^{(-1)}(u), G^{(-1)}(v)).$$

Sljedeća teorema pokazuje da proizvod kopula $\Pi(u, v) = uv$ određuje nezavisne slučajne veličine kada je funkcija raspodjele neprekidna.

Teorema 1.8 ([30]). Neka su X i Y neprekidne slučajne veličine. Tada X i Y su nezavisne ako i samo ako je $C_{XY} = \Pi$.

Sljedeće dvije teoreme se tiču parcijalnih izvoda kopule. Termin “skoro svuda” koristi se u Lebegovom smislu. ¹

Teorema 1.9 ([30]). *Neka je C kopula. Za svako $v \in [0, 1]$ parcijalni izvod $\frac{\partial C(u,v)}{\partial u}$ postoji skoro svuda i za svaki par tačaka (u, v) u kojima izvod postoji važi*

$$0 \leq \frac{\partial C(u, v)}{\partial u} \leq 1.$$

Analogno, za svako $u \in [0, 1]$ parcijalni izvod $\frac{\partial C(u,v)}{\partial v}$ postoji skoro svuda i za svaki par tačaka (u, v) u kojima izvod postoji važi

$$0 \leq \frac{\partial C(u, v)}{\partial v} \leq 1.$$

Štaviše, funkcije koje preslikavaju $u \mapsto \frac{\partial C(u,v)}{\partial v}$ i $v \mapsto \frac{\partial C(u,v)}{\partial u}$ su definisane i neopadajuće skoro svuda na $[0, 1]$.

Teorema 1.10 ([30]). *Neka je C kopula. Ako su $\frac{\partial C(u,v)}{\partial v}$ i $\frac{\partial^2 C(u,v)}{\partial u \partial v}$ neprekidni na $[0, 1] \times [0, 1]$ i $\frac{\partial C(u,v)}{\partial u}$ postoji za svako $u \in (0, 1)$ kada je $v = 0$, onda $\frac{\partial C(u,v)}{\partial u}$ i $\frac{\partial^2 C(u,v)}{\partial u \partial v}$ postoji na $(0, 1)^2$ i $\frac{\partial^2 C(u,v)}{\partial u \partial v} = \frac{\partial^2 C(u,v)}{\partial v \partial u}$.*

Sljedeća teorema tvrdi da pod određenim uslovima dvodimenzionalna funkcija gustine može biti zapisana kao proizvod marginalnih funkcija gustine i gustine kopule. Gustina kopule se često i naziva funkcija zavisnosti, jer uobličava sve informacije o zavisnosti između slučajnih veličina. Primjetimo da iz teoreme 1.10 slijedi da mješoviti parcijalni izvod $\frac{\partial^2 C(u,v)}{\partial u \partial v}$ postoji skoro svuda na $[0, 1]^2$.

Teorema 1.11 ([30]). *Ako su marginalne funkcije raspodjele F i G , sa gustinama f i g redom, neprekidne, tada zajednička dvodimenzionalna funkcija gustine može biti zapisana na sljedeći način*

$$h(x, y) = f(x)g(y)c(F(x), G(y)), \quad (x, y) \in \mathbb{R}^2,$$

gdje je c gustina kopule data sa

$$c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v}, \quad (u, v) \in [0, 1]^2.$$

¹Ako je (X, \mathcal{A}, μ) mjerljiv prostor, kažemo da svojstvo A važi skoro svuda na X ako postoji skup $N \in \mathcal{A}$ tako da $\mu(N) = 0$ i za svako $x \in X \setminus N$ imamo svojstvo A .

1.3 Mjerenje zavisnosti

Stohastička zavisnost između slučajnih mjerenja je jedan od važnih aspekata mnogih statističkih istraživanja kako sa teorijskog tako i primijenjenog stanovišta. Za mjerenje takve zavisnosti, nekoliko karakterističnih veličina je već uvedeno i široko se koristi, koji uključuju, na primjer, populacionu verziju Kendalovog tau (τ) i/ili Spearmanovog ro (ρ). Kopule su, s druge strane, dobro poznati alati za razumijevanje odnosa zavisnosti među slučajnim veličinama, a gornji τ i ρ su izraženi u terminima kopula funkcija i oba “mjere” oblik zavisnosti poznat kao saglasnost pa ćemo početi od toga. O kopulama i mjerama zavisnosti možemo više pronaći u [30].

1.3.1 Saglasnost mjera

Par slučajnih veličina je saglasan ako velike (male) vrijednosti jedne komponente su povezane sa velikim (malim) vrijednostima druge komponente. Preciznije, neka su (x_i, y_i) i (x_j, y_j) realizacije slučajnog vektora neprekidnog tipa (X, Y) . Kažemo da su (x_i, y_i) i (x_j, y_j) saglasni ako je $x_i < x_j$ i $y_i < y_j$, ili ako je $x_i > x_j$ i $y_i > y_j$. Slično kažemo da su (x_i, y_i) i (x_j, y_j) nesaglasni ako je $x_i < x_j$ i $y_i > y_j$ ili ako je $x_i > x_j$ i $y_i < y_j$. Primijetimo da je alternativna formulacija: (x_i, y_i) i (x_j, y_j) su saglasni ako je $(x_i - x_j)(y_i - y_j) > 0$ i nesaglasni ako je $(x_i - x_j)(y_i - y_j) < 0$.

1.3.2 Kendalov tau

Kendalov tau (τ) je mjera zavisnosti koje se definiše u terminima saglasnosti na sljedeći način.

Neka je $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ uzorak od n realizacija slučajnog vektora neprekidnog tipa (X, Y) . U uzorku postoji $\binom{n}{2}$ različitih parova (x_i, y_i) i (x_j, y_j) i svaki par je saglasan ili nesaglasan. Neka c označava broj saglasnih a d broj nesaglasnih parova. Tada se Kendalov tau za dati uzorak definiše kao

$$t = \frac{c - d}{c + d} = \frac{c - d}{\binom{n}{2}}.$$

Ekvivalentno, t je vjerovatnoća saglasnosti umanjena za vjerovatnoću nesaglasnosti para (x_i, y_i) i (x_j, y_j) koji je izabran slučajno iz uzorka. Kendalov tau na populaciji se definiše na sličan način. Neka su (X_1, Y_1) i (X_2, Y_2) nezavisni slučajni vektori neprekidnog tipa sa istom raspodjelom H . Tada Kendalov tau na populaciji se definiše kao razlika između

vjerovatnoća saglasnosti i nesaglasnosti:

$$\tau = \tau_{X,Y} = P\{(X_1 - Y_1)(X_2 - Y_2) > 0\} - P\{(X_1 - Y_1)(X_2 - Y_2) < 0\}.$$

U cilju da pokažemo povezanost kopula i Kendalovog tau, prvo ćemo definisati funkciju saglasnosti Q .

Teorema 1.12 ([30]). *Neka su (X_1, Y_1) i (X_2, Y_2) nezavisni vektori neprekidnog tipa sa zajedničkim funkcijama H_1 i H_2 , respektivno, i sa istim marginalnim funkcijama raspodjele F (za X_1 i X_2) i G (za Y_1 i Y_2). Neka su C_1 i C_2 kopule od (X_1, Y_1) i (X_2, Y_2) respektivno, tako da $H_1(x, y) = C_1(F(x), G(y))$ i $H_2(x, y) = C_2(F(x), G(y))$. Neka Q označava razliku između vjerovatnoća saglasnosti i nesaglasnosti od (X_1, Y_1) i (X_2, Y_2) , tj. neka je*

$$Q = P\{(X_1 - Y_1)(X_2 - Y_2) > 0\} - P\{(X_1 - Y_1)(X_2 - Y_2) < 0\}.$$

Tada

$$Q = Q(C_1, C_2) = 4 \int_0^1 \int_0^1 C_2(u, v) dC_1(u, v) - 1.$$

Teorema 1.13 ([30]). *Neka su X i Y neprekidne slučajne veličine čija je kopula C . Tada je Kendalov tau za X i Y dat sa*

$$\tau_{X,Y} = \tau_C = Q(C, C) = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1 = 4\mathbb{E}(C(U, V)) - 1,$$

gdje su U i V slučajne veličine uniformne na $(0, 1)$ čija je zajednička funkcija raspodjele C .

1.3.3 Spirmanov koeficijent korelacije

Kao i Kendalov tau, mjera povezanosti poznata pod nazivom Spirmanovo ro se zasniva na saglasnosti i nesaglasnosti. Da bi dobili populacionu verziju ove mjere, neka su (X_1, Y_1) , (X_2, Y_2) i (X_3, Y_3) tri nezavisna slučajna vektora sa istom zajedničkom raspodjelom H (čije marginalne funkcije raspodjele su F i G) i kopulom C . Spirmanovo ro $\rho_{X,Y}$ se definiše da je proporcionalno razlici vjerovatnoća saglasnosti i nesaglasnosti za dva vektora (X_1, Y_1) i (X_2, Y_3) , tj. paru vektora sa istim marginalima, ali jedan vektor ima funkciju raspodjele H , dok su komponente drugog vektora nezavisne:

$$\rho_{X,Y} = 3(P\{(X_1 - X_2)(Y_1 - Y_3) > 0\} - P\{(X_1 - X_2)(Y_1 - Y_3) < 0\})$$

(mogli smo uzeti i par (X_3, Y_2)). Primijetimo da funkcija raspodjele vektora (X_1, Y_1) je $H(x, y)$, a dok je funkcija raspodjele vektora (X_2, Y_3) je $F(x)G(y)$ (jer su X_2 i Y_3 nezavisne). Tada kopula od X_2 i Y_3 je Π , i korišćenjem teoreme 1.12 imamo

Teorema 1.14 ([30]). *Neka su X i Y neprekidne slučajne veličine čija je kopula C . Tada Spirmanovo ro za X i Y (u oznaci $\rho_{X,Y}$ ili ρ_C) je dato sa:*

$$\begin{aligned}\rho_{X,Y} = \rho_C &= 3Q(C, \Pi) = \\ &= 12 \int_0^1 \int_0^1 uv \, dC(u, v) - 3 = \\ &= 12 \int_0^1 \int_0^1 C(u, v) \, dudv - 3.\end{aligned}\tag{1.7}$$

Spirmanovo ro se često naziva i Spirmanov rang koeficijent korelacije. Ako su x i y realizacije dvije slučajne veličine X i Y sa funkcijama raspodjele F i G , respektivno, tada su rangovi od x i y dati sa $u = F(x)$ i $v = G(y)$. Primijetimo da su rangovi (u i v) realizacije iz uniformnih na $(0, 1)$ slučajnih veličina $U = F(X)$ i $V = G(Y)$ čija zajednička funkcija raspodjele je C . Pošto obje i U i V imaju očekivanje $\frac{1}{2}$ i disperziju $\frac{1}{12}$, izraz za ρ_C u (1.7) može se zapisati na sljedeći način:

$$\begin{aligned}\rho_{X,Y} = \rho_C &= 12 \int_0^1 \int_0^1 uv \, dC(u, v) - 3 = 12\mathbb{E}(UV) - 3 = \\ &= \frac{\mathbb{E}(UV) - 1/4}{1/12} = \frac{\mathbb{E}(UV) - \mathbb{E}(U)\mathbb{E}(V)}{\sqrt{\mathbb{D}(U)}\sqrt{\mathbb{D}(V)}}.\end{aligned}$$

Kao posljedica, Spirmanovo ro za par neprekidnih slučajnih veličina X i Y je identično Pirsonovom koeficijentu za rangove X i Y , tj. za slučajne veličine $U = F(X)$ i $V = G(Y)$.

1.4 Uvod u statistiku

U stvarnom životu radimo sa podacima koji su pod uticajem slučajnosti i mi moramo da izvučemo potrebne informacije iz tih podataka i donesemo zaključke. Slučajnost može biti prouzrokovana različitim aspektima. Razmotrimo sljedeća dva primjera:

1. Pretpostavimo da želimo da predvidimo ishod izbora. Pošto ne možemo da anketiramo cijelu populaciju, odabraćemo slučajni uzorak iz populacije i pitati ih za koga planiraju da glasaju. U ovom eksperimentu, slučajnost dolazi iz metoda izbora uzorka. Takođe treba imati na umu da anketu sprovodimo mjesec dana

prije izbora, još jedan izvor slučajnosti je da anketirane osobe mogu promijeniti svoje mišljenje tokom tog perioda od mjesec dana.

2. U sistemu bežične komunikacije, poruka se prenosi od predajnika ka prijemniku. Međutim, prijemnik dobija oštećenu verziju (verziju sa šumom) poslatog signala. Prijemnik treba da izvuče originalnu poruku iz primjene verzije sa šumom. Ovdje izvor slučajnosti predstavlja pojava šuma.

Primjeri koje smo sada naveli su iz oblasti statistike zaključivanja. Dakle, postupak kojim se rezultat dobijen na nivou uzorka uopštava na nivo populacije i kojim se mjeri pouzdanost uopštenog rezultata naziva se inferencijalna statistika.

Jasno je da koristimo znanje iz teorije vjerovatnoće kada radimo sa problemima iz oblasti statistike zaključivanja. Međutim, veliki dodatak je što ovdje moramo da radimo sa stvarnim podacima. Vjerovatnosni problemi su jasno definisani i poznati su nam vjerovatnosni modeli. Navedimo sljedeći primjer. Neka je X slučajna veličina koja ima normalnu raspodjelu sa očekivanjem $\mu = 100$ i disperzijom $\sigma^2 = 15$. Odrediti vjerovatnoću da je $X > 110$. U statističkom zadatku mi nećemo znati raspodjelu za X , već je potrebno da sakupimo podatke i da na osnovu podataka donesemo zaključak da li X ima normalnu raspodjelu. Rješavanje svih statističkih zadataka podrazumijeva sakupljanje podataka. Stoga se nerijetko može čuti da je statistika nauka koja razrađuje postupke zaključivanja na osnovu podataka.

Znamo da je sakupljanje podataka prvi korak u realizaciji mnogih naučnih projekata. Podaci se mogu prikupljati iz populacije i tako izučavati populaciju u cjelini. Međutim, ako populacija sadrži veliki broj elemenata, tada izučavanje cijele populacije može trajati dugo ili prouzrokovati veće materijalne troškove. U većini slučajeva je nemoguće ispitivati cijelu već samo dio populacije.

Definicija 1.13. *Dio populacije koji se proučava statističkim metodama zove se uzorak.*

Obilježje, označimo ga sa X , je slučajna veličina koja predstavlja numeričku karakteristiku nekog fenomena iz oblasti istraživanja. Navedimo nekoliko primjera: razred koje dijete pohađa, tjelesna masa, visina, stepen obrazovanja roditelja itd.

Neka se na nekoj populaciji posmatra obilježje X i neka se eksperiment sastoji u izboru jednog elementa populacije i registrovanju vrijednosti obilježja X na njemu. Statistički model koji odgovara ovom eksperimentu je $(\Omega, \mathcal{F}, \mathcal{P})$, gdje je Ω skup svih mogućih ishoda eksperimenta, \mathcal{F} je σ -algebra podskupova skupa Ω i \mathcal{P} je familija dopustivih raspodjela vjerovatnoća obilježja X . Na slučajan način biramo n elemenata w_1, w_2, \dots, w_n i neka su rezultati mjerenja vrijednosti obilježja X na tim elementima predstavljeni

slučajnim veličinama $X(w_1), X(w_2), \dots, X(w_n)$ ili skraćeno sa X_1, X_2, \dots, X_n . Tada se n -dimenzionalna slučajna veličina $X = (X_1, X_2, \dots, X_n)$ zove slučajan uzorak, a broj n predstavlja obim slučajnog uzorka. Statistički model koji odgovara ovom eksperimentu je uređena trojka $(\Omega^n, \mathcal{F}^n, \mathcal{P})$, gdje je $\Omega^n = \Omega \times \Omega \times \dots \times \Omega$, $\mathcal{F}^n = \mathcal{F} \times \dots \times \mathcal{F}$, a \mathcal{P} je familija dopustivih n -dimenzionalnih raspodjela vjerovatnoća slučajnog uzorka (X_1, X_2, \dots, X_n) definisanih na prostoru $(\Omega^n, \mathcal{F}^n)$. Postor $(\Omega^n, \mathcal{F}^n)$ nazivamo uzorački prostor.

Definicija 1.14. *Neka obilježje X ima funkciju raspodjele $F(x)$. Prost slučajni uzorak (eng. simple random sample) obima n iz raspodjele $F(x)$ je slučajni vektor (X_1, X_2, \dots, X_n) , gdje su X_1, X_2, \dots, X_n nezavisne slučajne veličine sa istom funkcijom raspodjele $F(x)$.*

Prije sprovođenja eksperimenta, vrijednosti koje ćemo dobiti su slučajne veličine X_1, X_2, \dots, X_n , registrovane vrijednosti zavise od slučaja, u različitim serijama dobijamo različite podatke-rezultate. Kada je eksperiment sproveden dobijamo vektor (x_1, x_2, \dots, x_n) -realizovani uzorak, brojevi x_1, \dots, x_n su registrovane vrijednosti.

Definicija 1.15. *Neka je $f : \mathbb{R}^n \rightarrow \mathbb{R}$ Borelova funkcija i (X_1, \dots, X_n) uzorak. Slučajna veličina $V = f(X_1, \dots, X_n)$ se naziva statistika ili uzoračka karakteristika.*

Ako je (x_1, \dots, x_n) realizovani uzorak, tada je $v = f(x_1, \dots, x_n)$ realizovana statistika V .

1.4.1 Ocjenjivanje parametara

Osnovni zadatak matematičke statistike jeste određivanje raspodjele vjerovatnoća posmatranog obilježja na elementima populacije. Znamo da je raspodjela vjerovatnoća nekog obilježja potpuno određena ako je poznat njen analitički oblik i tačne vrijednosti parametara koji se javljaju u formuli, kojom se definiše analitički oblik raspodjele vjerovatnoća. Međutim prilikom proučavanja obilježja populacije može se desiti da nam je poznat analitički oblik raspodjele vjerovatnoća ali da su nam nepoznate vrijednosti parametara koje figurišu u tom obliku.

Posmatrajmo statistički model $(\Omega^n, \mathcal{F}^n, \mathcal{P})$. Kod ocjenjivanja parametra pretpostavlja se da obilježje X ima raspodjelu vjerovatnoća koja pripada familiji dopustivih raspodjela vjerovatnoća $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$, gdje je Θ parametarski prostor i predstavlja skup svih mogućih vrijednosti parametra θ .

Do informacije o vrijednostima parametra se dolazi pomoću slučajnog uzorka. Neka je $X = (X_1, X_2, \dots, X_n)$ prost slučajni uzorak obima n iz populacije čije obilježje X ima raspodjelu vjerovatnoća koja pripada familiji dopustivih raspodjela vjerovatnoća $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$. Procedura ocjenjivanja nepoznatog parametra θ podrazumijeva izbor statistike $\hat{\theta}_n = U(X)$ čijom se realizacijom $u = U(x_1, \dots, x_n)$ aproksimira parametar θ . Statistika $\hat{\theta}_n = U(X)$ i njena realizacija $u = U(x_1, \dots, x_n)$ se nazivaju ocjena parametra θ .

Jedna od najpoželjnijih osobina ocjene parametra je nepristrasnost. Nepristrasnost nam govori o tome da je raspodjela ocjene parametra centrirana oko tačne vrijednosti nepoznatog parametra.

Definicija 1.16. *Neka je (X_1, X_2, \dots, X_n) prost slučajan uzorak obima n iz populacije čije obilježje X ima raspodjelu vjerovatnoća $P_\theta \in \mathcal{P}$. Kažemo da je ocjena $\hat{\theta}_n$ nepristrasna ocjena za funkciju parametara $\varphi(\theta)$ ako za svako $\theta \in \Theta$ važi*

$$\mathbb{E}(\hat{\theta}_n) = \varphi(\theta).$$

Prilikom rada sa ocjenama parametara, posmatra se veličina $b(\cdot)$, koja je za ocjenu $\hat{\theta}_n$ definisana kao $b(\hat{\theta}_n) = \mathbb{E}(\hat{\theta}_n) - \varphi(\theta)$. Ovako definisana veličina b naziva se pristrasnost ocjene $\hat{\theta}_n$ (eng. bias).

Ocjene parametara se dijele na tačkaste i intervalne. Tačkasta ocjena parametra $\hat{\theta}$ određena je jednim brojem $\theta = u(x_1, x_2, \dots, x_n)$. Intervalna ocjena se naziva ocjenom, koju određuju dva broja $\hat{\theta}_1$ i $\hat{\theta}_2$ - krajevi intervala, u čijim granicama se nalazi ocijenjeni parametar θ .

U nastavku dajemo osnovne metode za određivanje tačkastih ocjena nepoznatih parametara.

1.4.2 Metod momenata

Neka je (x_1, \dots, x_n) uzorak obilježja X čija funkcija raspodjele pripada familiji $\mathcal{P} = \{F(x, \theta), \theta \in \Theta\}$. Potrebno je ocijeniti nepoznati parametar θ date raspodjele.

Po analogiji sa momentima slučajne veličine X uvodi se pojam empirijskih momenata. Empirijski i njima odgovarajući teorijski momenti reda k dati su sljedećim formulama:

teorijski momenti

$$v_k = \begin{cases} \sum_i x_i^k p_i(x; \theta), & \text{za diskretnu slučajnu veličinu } X, \\ \int_{-\infty}^{+\infty} x^k f(x; \theta) dx, & \text{za apsolutno - neprekidnu slučajnu veličinu } X; \end{cases}$$

empirijski momenti

$$v_k^* = \overline{x^k} = \frac{1}{n} \sum_i x_i^k.$$

Posebno napominjemo da početni teorijski momenat prvog reda je matematičko očekivanje slučajne veličine X , a empirijski momenat prvog reda je aritmetička sredina posmatranih vrijednosti slučajne veličine X , tj.

$$v_1^* = \bar{x} = \frac{1}{n} \sum_i x_i.$$

Analogno se definišu empirijski i njima odgovarajući teorijski centralni momenti reda k :

teorijski centralni momenti

$$\mu_k = \begin{cases} \sum_i (x_i - m_x)^k p_i(x; \theta), & \text{za diskretnu slučajnu veličinu } X, \\ \int_{-\infty}^{+\infty} (x - m_x)^k f(x; \theta) dx, & \text{za apsolutno - neprekidnu slučajnu veličinu } X; \end{cases}$$

empirijski centralni momenti

$$\mu_k^* = \overline{(x - \bar{x})^k} = \frac{1}{n} \sum_i (x_i - \bar{x})^k.$$

Napomenimo da teorijski centralni moment drugog reda je disperzija slučajne veličine X .

Treba imati u vidu da empirijski momenti uzimaju slučajne vrijednosti, dok su teorijski momenti konstante koje karakterišu raspodjelu populacije.

Najprostiji metod za ocjenjivanje nepoznatog parametra je metod momenata, koji je predložio engleski statističar Karl Pearson 1894. godine. Metod momenata zasniva se na tome da se empirijski momenti (ili njihove funkcije) uzimaju za ocjene odgovarajućih teorijskih momenata (ili njihovih funkcija).

Dakle, izjednačavajući teorijski početni momenat prvog reda sa empirijskim početnim momentom prvog reda, zaključujemo da je ocjena matematičkog očekivanja

slučajne veličine X , iz proizvoljne raspodjele, aritmetička sredina posmatranih vrijednosti slučajne veličine X , tj.

$$\widehat{\mathbb{E}}(X) = \bar{x} = \frac{1}{n} \sum_i x_i.$$

Izjednačavajući teorijske i empirijske centralne momente drugog reda, dolazimo do toga da ocjena disperzije slučajne veličine X , iz proizvoljne raspodjele, nalazi se po formuli:

$$\widehat{\mathbb{D}}(X) = \hat{\sigma}_x^2 = \bar{s}^2 = \frac{\sum_i (x_i - \bar{x})^2}{n}.$$

Postupajući po analognom obrazcu, možemo odrediti ocjene teorijskih momenata proizvoljnog reda.

1.4.3 Metod maksimalne vjerodostojnosti

Metod maksimalne vjerodostojnosti je jedan od najpoznatijih i najčešće korišćenih metoda za ocjenjivanje nepoznatih parametara raspodjele obilježja.

Neka je $X = (X_1, \dots, X_n)$ uzorak obilježja X čija funkcija raspodjele pripada familiji $\mathcal{P} = \{F(x, \theta), \theta \in \Theta\}$.

Definisaćemo funkciju vjerodostojnosti u slučaju kada raspodjela obilježja pripada familiji raspodjela diskretnog tipa $\mathcal{P} = \{p(w_k, \theta), \theta \in \Theta\}$, $p(w_k, \theta) = P_\theta\{X = w_k\}$.

Definicija 1.17. *Funkcija vjerodostojnosti $L(\theta, x_1, x_2, \dots, x_n) = p(x_1, \theta) \dots p(x_n, \theta) = P_\theta(X_1 = x_1) \dots P_\theta(X_n = x_n) = P_\theta\{(X_1, X_2, \dots, X_n) = (x_1, x_2, \dots, x_n)\}$, $\theta \in \Theta$, je funkcija čiji je argument θ , (x_1, \dots, x_n) je neka realizacija uzorka obilježja X i brojevi x_1, \dots, x_n se tretiraju kao parametri.*

Definisaćimo sada funkciju vjerodostojnosti u slučaju kada raspodjela obilježja pripada familiji raspodjela apsolutno neprekidnog tipa $\mathcal{P} = \{g(x, \theta), \theta \in \Theta\}$, g je funkcija gustine.

Definicija 1.18. *Funkcija vjerodostojnosti*

$$L(\theta) = L(\theta, x_1, x_2, \dots, x_n) = g(x_1, \theta)g(x_2, \theta) \dots g(x_n, \theta), \quad \theta \in \Theta,$$

je funkcija čiji je argument θ , (x_1, \dots, x_n) je neka realizacija uzorka obilježja X i brojevi x_1, \dots, x_n se tretiraju kao parametri.

Znači, funkcija vjerodostojnosti je definisana kao funkcija parametra θ .

Princip metoda maksimalne vjerodostojnosti sastoji se u pronalaženju ocjene za nepoznati parametar $\theta = (\theta_1, \theta_2, \dots, \theta_k)$, koji maksimizuje funkciju vjerodostojnosti $L(\theta)$ za različite vrijednosti parametara, tj. želimo da odredimo $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k)$ tako da

$$L(\hat{\theta}) > L(\theta), \forall \theta \in \Theta,$$

tj., $L(\hat{\theta}) = \sup_{\theta \in \Theta} L(\theta)$. Dakle, ako postoji funkcija $\hat{\theta} = \hat{\theta}(x_1, x_2, \dots, x_n)$ koja maksimizuje $L(\theta)$ za različite vrijednosti parametra θ , tada je $\hat{\theta}$ ocjena maksimalne vjerodostojnosti parametra θ . Dakle, $\hat{\theta}$ je rješenje, ako je

$$\frac{\partial L(\theta)}{\partial \theta} = 0 \text{ i } \frac{\partial^2 L(\theta)}{\partial \theta^2} < 0. \quad (1.8)$$

Kako je funkcija vjerodostojnosti $L(\theta)$ nenegativna funkcija, to maksimum možemo tražiti u onim vrijednostima u kojima je $L(\theta) > 0$. Poznato nam je da funkcije $L(\theta) > 0$ i $\ln L(\theta)$ imaju ekstremume u istim vrijednostima tako da umjesto određivanja maksimuma funkcije $L(\theta)$ najčešće određujemo maksimum funkcije $\ln L(\theta)$ koja je poznata pod nazivom logartiamska funkcija vjerodostojnosti. Tada prva od dvije jednačine u (1.8) može se zapisati kao

$$\frac{1}{L(\theta)} \cdot \frac{\partial L(\theta)}{\partial \theta} = 0 \implies \frac{\partial \ln L(\theta)}{\partial \theta} = 0.$$

1.4.4 Intervalno ocjenjivanje

Tačkasta ocjena nepoznatnog parametra θ , dobijena na uzorku obima n iz opšte populacije sa funkcijom raspodjele $F(x, \theta)$, ne može nam direktno odgovoriti na pitanje, koliku grešku pravimo uzimajući umjesto tačne vrijednosti za parametar θ neku od njenih ocjena $\hat{\theta} = u(x_1, x_2, \dots, x_n)$. Iz tog razloga, u mnogim slučajevima korisnije je koristiti intervalnu ocjenu, zasnovanu na određivanju nekog intervala, unutar koga se sa određenom vjerovatnoćom nalazi nepoznata vrijednost parametra θ .

Neka je statistička karakteristika $\hat{\theta} = u(x_1, x_2, \dots, x_n)$, koja je određena na osnovu uzorak obima n , tačna ocjena nepoznatog parametra θ . Što je manja razlika $|\theta - \hat{\theta}|$, to je bolji kvalitet ocjene, odnosno bolja je ocjena. Dakle, pozitivan broj ε karakteriše tačnost ocjene

$$|\theta - \hat{\theta}| < \varepsilon. \quad (1.9)$$

Jasno je da tačnost ocjene zavisi od obima uzorka n . Kolika bi trebalo da bude veličina uzorka n , da bi se obezbjedila zadata tačnost ε ili kako da odredimo tačnost na osnovu zadatog obima uzorka? Na ova pitanja ne možemo direktno da dobijemo odgovor koristeći nejednakost (1.9). Može se samo govoriti o vjerovatnoći $1 - \alpha$, sa kojom je ova nejednakost zadovoljena.

Definicija 1.19. *Nivo povjerenja ocjene je vjerovatnoća $p = 1 - \alpha$, $0 < \alpha < 1$, za koju je ispunjena nejednakost $|\theta - \hat{\theta}| < \varepsilon$.*

Nivo povjerenja za tačkastu ocjenu pokazuje da pri izboru uzorka obima n iz jedne te iste populacije sa funkcijom raspodjele $F(x, \theta)$ u $(1 - \alpha)100\%$ slučajeva parametar θ biće pokriven ovim intervalom.

Neka je nejednakost $|\theta - \hat{\theta}| < \varepsilon$ ispunjena sa vjerovatnoćom $1 - \alpha$:

$$P \left\{ |\theta - \hat{\theta}| < \varepsilon \right\} = 1 - \alpha,$$

odnosno

$$P \left\{ \hat{\theta} - \varepsilon < \theta < \hat{\theta} + \varepsilon \right\} = 1 - \alpha. \quad (1.10)$$

Formula (1.10) pokazuje da se nepoznati parametar θ nalazi unutar intervala $(\hat{\theta} - \varepsilon, \hat{\theta} + \varepsilon)$.

Definicija 1.20. *Interval $(\hat{\theta} - \varepsilon, \hat{\theta} + \varepsilon)$ naziva se interval povjerenja za nepoznati parametar θ sa nivoom povjerenja $p = 1 - \alpha$, ($0 < \alpha < 1$).*

U praktičnim primjenama važnu ulogu ima dužina intervala povjerenja. Što je dužina intervala povjerenja $(\hat{\theta} - \varepsilon, \hat{\theta} + \varepsilon)$ manja to je ocjena preciznija. Ako je dužina intervala povjerenja velika, tada ocjena nije praktično upotrebljiva.

1.4.5 Testiranje statističkih hipoteza

U prethodim poglavljima vidjeli smo metode koje se koriste za ocjenjivanje nepoznatog parametra. Pronalaženje tačkastih ili intervalnih ocjena je po pravilu neka preliminarna faza statističkog istraživanja. Krajnji cilj studije može, na primjer, biti uporedna analiza različitih tehnoloških procesa u pogledu njihove produktivnosti, tačnosti ili ekonomičnosti, u poređenju karakteristika uređaja, proizvoda i sl.

Na matematičkom jeziku zadaci uporedne analize formulišu se kao zadaci statističkog testiranja hipoteza o parametrima zakona raspodjele, pošto promjena parametara karakteriše razlike u tehnološkim procesima, konstrukcijama, uređajima itd.

Pretpostavimo da se za rješavanje problema uporedne analize iz opšte populacije izdvaja uzorak $(x_1, x_2, x_3, \dots, x_n)$ obima n . Neka smo dalje na osnovu oblika histograma ili poligona relativnih frekvencija, ili iz bilo kog drugog razmatranja postavili hipotezu o proučavanom zakonu raspodjele slučajne veličine X i ocjenivši parametre ovog zakona, izgradili vjerovatnosni model raspodjele vjerovatnoće ove slučajne veličine, koji po našem mišljenju, objašnjava glavne karakteristike obilježja X .

Ako nije poznat tip raspodjele F obilježja X , u statističkom zaključivanju se primjenjuju postupci koji se nazivaju neparametarski metodi ili metodi nezavisni od raspodjele. Termin “metod nezavisan od raspodjele” koristi se jer se o raspodjeli F ništa ne pretpostavlja, osim možda da je neprekidnog tipa. Tada se postavljaju hipoteze o obliku funkcije raspodjele koje se nazivaju neparametarske hipoteze.

Dalji zadatak je da testiramo predloženu hipotezu, odnosno da objasnimo koji model je vjerovatnije odabran. Za provjeru ove hipoteze koriste se različiti statistički kriterijumi koje možete pronaći u [38].

Pretpostavimo da smo se uz pomoć ovih kriterijuma uvjerali da je model “dobar”, tj. eksperimentalni podaci nijesu u suprotnosti sa ovim modelom.

Ako sada, na primjer, ocjenimo parametar θ vjerovatnosnog modela $F(x, \theta)$ na osnovu dva nezavisna statistička uzorka, uzetih, recimo, prije uvođenja nekog poboljšanja u tehnološki proces i poslije uvođenja, onda dobijamo dvije ocjene $\hat{\theta}_1$ i $\hat{\theta}_2$, koje zbog slučajnosti neće biti jednake jedna drugoj. Postavlja se pitanje (postavlja se hipoteza), da li su ove ocjene, ocjene jednog te istog parametra θ vjerovatnosnog modela, ili je ovaj parametar promjenjen usljed uvođenja poboljšanja u tehnološki proces? Problemi ove vrste su tipični problemi uporedne analize.

Definicija 1.21. *Statistička hipoteza se naziva parametarskom ako sadrži pretpostavke o vrijednostima parametara funkcije raspodjele poznatog oblika.*

Uz predloženu hipotezu razmatra se jedna ili više alternativnih (konkurentskih) hipoteza. Ako se predložena hipoteza odbaci, onda na njeno mjesto dolazi alternativna hipoteza. Sa ove tačke gledišta, statističke hipoteze se dijele na nulte i alternativne.

Definicija 1.22. *Nulta hipoteza se naziva glavna (predložena) hipoteza. Nulta hipoteza se označava simbolom H_0 . Tipično, nulte hipoteze navode da ne postoji razlika*

između upoređenih vrijednosti (parametara ili funkcija raspodjele), a uočena odstupanja se objašnjavaju samo slučajnim fluktuacijama uzorka.

Definicija 1.23. Alternativna hipoteza je pretpostavka koja je suprotna nultoj hipotezi u tom smislu da ako se nulta hipoteza odbaci, onda se prihvata alternativna. Alternativnu hipotezu označavamo simbolom H_1 .

Termini “odbacivanje hipoteze” i “prihvatanje hipoteze” ne znače da se neka hipoteza u potpunosti odbacuje, već da uzorak govori u korist jedne od te dvije hipoteze. Za druge uzorke se mogu donijeti drugačiji zaključci o prihvatnosti ili odbacivanju hipoteze.

Navešćemo primjer nulte i alternativne hipoteze paramterarskog tipa. Pretpostavka oblika $H_0(\theta \in \Theta_0)$, $\Theta_0 \subset \Theta$ ili malo drugačije rečeno pretpostavka koja se sastoji u tome da obilježje X ima raspodjelu koja pripada užoj familiji određenoj parametarskim skupom Θ_0 , predstavlja nultu hipotezu. Ako je Θ_0 jednočlani skup, tada kažemo da je hipoteza H_0 prosta. U protivnom govorimo o složenoj hipotezi. Hipotezi H_0 ćemo suprotstaviti alternativnu hipotezu $H_1(\theta \in \Theta \setminus \Theta_0)$.

Testiranje statističkih hipoteza vrši se na osnovu realizovanog uzorka. Za ovo se koristi posebno odabrana slučajna veličina (statistika uzorka), koja je funkcija posmatranih vrijednosti, čija tačna ili približna raspodjela je poznata.

Definicija 1.24. Statistički kriterijum (test) je slučajna veličina T , uz pomoć koje se donose odluke o prihvatanju ili odbijanju predložene nulte hipoteze.

Za neke realizovane uzorke možemo da donesemo odluku o prihvatanju nulte hipoteze, dok za druge realizovane uzorke donosimo odluku o odbacivanju nulte hipoteze i prihvatanju alternativne hipoteze. Dakle, uzorački prostor Ω^n dijeli se na dva disjunkt-na skupa, na skup realizovanih uzoraka na kome prihvatamo nultu hipotezu i na skup realizovanih uzoraka na kojima se nulta hipoteza odbacuje.

Definicija 1.25. Skup svih realizovanih uzoraka $(x_1, x_2, \dots, x_n) \in \Omega^n$ na osnovu kojih se odbacuje nulta hipoteza zove se kritična oblast i označava se sa C .

Definicija 1.26. Skup svih realizovanih uzoraka $(x_1, x_2, \dots, x_n) \in \Omega^n$ na osnovu kojih se prihvata nulta hipoteza zove se oblast prihvatnja i označava se sa C^c .

Prilikom odlučivanja, postoji mogućnost da se napravi greška.

1. Grešku prve vrste pravimo kada odbacimo faktički tačnu hipotezu H_0 .
2. Grešku druge vrste pravimo kada odbacimo faktički tačnu hipotezu H_1 .

Kada su hipoteze H_0 i H_1 proste, tada se ove vjerovatnoće definišu na sljedeći način. Vjerovatnoća greške prve vrste je broj α definisan kao

$$\alpha = P_{H_0}\{(X_1, X_2, \dots, X_n) \in C\}.$$

Vjerovatnoća α se zove još i prag značajnosti. Za α uzimamo male vrijednosti iz intervala $(0, 1)$, uglavnom 0,05 i 0,01. Ako kritična oblast C ima vjerovatnoću greške prve vrste jednaku α , tada kažemo da je C kritična oblast veličine α . Vjerovatnoća greške druge vrste se označava sa β i jednaka je

$$\beta = P_{H_1}\{(X_1, X_2, \dots, X_n) \in C^c\}.$$

1.5 Primjena butstrap metode u testu Kolmogorov - Smirnova

Kolmogorov - Smirnov (K - S) test koristi maksimalnu apsolutnu razliku D_n između empirijske i teorijske funkcije raspodjele i jedan je od najčešće korišćenih testova. Generalno, K - S se smatra veoma efikasnim, jer koristi svaki pojedinačni element u uzorku. U K - S, D_n je mjera testa maksimalne vjerodostojnosti (eng. goodness of fit). S druge strane D_n je takođe stohastička veličina koja u velikoj mjeri zavisi od parametara hipotetičkog modela i raspodjelu od D_n je veoma teško izvesti. Iako postoje brojni radovi na tu temu i dalje ne postoji metod kojim se mogu lako ocijeniti parametri raspodjele od D_n .

Princip butstrap metodologije uveo je američki statističar Efron (1979). To je takođe efikasan statistički alat koji se može koristiti u skoro svakoj primjeni za ocjenu nulte hipoteze. Posebno je korisna u situacijama gdje je asimptotska aproksimacija previše teška ili uopšte nije izvodljiva. Bez ikakvih hipoteza o teorijskim raspodjelama, butstrap metoda procjenjuje nepoznatu raspodjelu koristeći informacije iz konačnog broja uzorka, koji su dobijeni računarskom simulacijom. Drugim riječima, zaključiti o nepoznatim populacijama (koje su predstavljene statističkim modelima) se donose na osnovu podataka iz graničnog uzorka. Butstrap metode su fleksibilnije u odnosu na klasične metode koje mogu biti analitički neatraktivne jer ne postoje odgovarajuće pretpostavke koje moraju biti zadovoljene. S druge strane ako se mogu koristiti klasične metode, butstrap metode će dati prilično slične rezultate. U stvari, butstrap metode su toliko široko korišćene da ih je Efron opisao u cilju stvaranja "butstrap svijeta", gdje analitičar podataka zna sve. Odnosno, u butstrap svijetu, analitičar podataka može dobiti bilo koju informaciju primjenom simulacije.

1.5.1 Butstrap metoda

Ako je poznata teorijska funkcija raspodjele $F(x)$, statistički parametar θ može se dobiti iz sljedeće formule:

$$\theta = \int t(x)dF(x),$$

gdje je $t(x)$ statistička funkcija. Pretpostavimo da je $\{x_i : i = 1, 2, \dots, n\}$ posmatrani uzorak. Funkcija raspodjele je obično nepoznata zbog toga što se karakteristike uzorka tokom vremena mijenjaju. Ako je obim uzorka dovoljno veliki, vrijednost statističkog parametra može se dobiti pomoću empirijske raspodjele.

Butstrap metoda prvo proširuje veličinu uzorka pomoću metode ponovnog uzorkovanja (reuzorkovanja) sa ponavljanjem. Pravilo metode je da je mogućnost pojave svakog elementa $\frac{1}{n}$, gdje je n obim uzorka. To znači da će se u procesu reuzorkovanja neki elementi možda pojaviti više od jednog puta, dok se neki drugi elementi možda neće uopšte pojaviti. Onda se funkcija raspodjele $F(x)$ zamjenjuje sa empirijskom funkcijom raspodjele $\hat{F}(x)$:

$$\hat{F}(x) = \frac{N\{x_i \leq x\}}{n}, \quad i = 1, 2, \dots, n,$$

gdje je $N\{x_i \leq x\}$ broj onih vrijednosti iz uzorka koji su manje ili jednake x . Kada je dobijena empirijska funkcija $\hat{F}(x)$, tada je statistički parametar uzorka $\hat{\theta}$ dat sa:

$$\hat{\theta} = \int t(x)d\hat{F}(x).$$

Novi uzorci X_i^* , gdje indeks i označava i -to reuzorkovanje, se dobijaju kada se M puta uzastopno vrši reuzorkovanje iz originalnog uzorka. Na taj način dobijamo različite odgovarajuće empirijske funkcije $\hat{F}(x)$ i statistike θ_i^* .

Uticaj koliko puta vršimo reuzorkovanje na rezultat i empirijsku funkciju nije veliki. U opštem slučaju broj M možemo odrediti pomoću empirijske formule [54]:

$$M = 40n,$$

gdje je n obim originalnog uzorka. Broj ponavljanja uzoraka treba da bude proporcionalan originalnoj veličini uzorka. Odnosno, što je veća veličina uzorka, to je više kombinacija za novo uzorkovanje.

Najvažnija ideja butstrap metode je da na osnovu empirijske funkcije raspodjele $\hat{F}(x)$ vrijednost za θ_i^* je približno $\hat{\theta}$, slično na osnovu teorijske funkcije raspodjele $F(x)$ vrijednost od $\hat{\theta}$ je približno θ .

1.5.2 K - S test zasnovan na butstrap metodu

Kompatibilnost slučajnih uzoraka u odnosu na neku teorijsku funkciju raspodjele može se mjeriti testovima saglasnosti. Test Kolmogorov - Smirnova (K - S), koji je pogodan za skoro sve vrste raspodjela, je neparametarska statistika za upoređivanje dvije empirijske raspodjele koja se definiše kao najveća apsolutna razlika između funkcija raspodjela što predstavlja mjeru nepoklapanja. Prema teoriji uzorkovanja, što je veći obim uzorka, ne samo da je preciznija ocjena parametara, već je i pouzdanost veća. Butstrap metoda koristi informacije iz uzorka ne narušavajući teorijsku raspodjelu, analizirajući reuzorkovane podatke uzastopno dobijamo preciznije ocjene parametara nego što su ocijenjeni rezultati na osnovu jednog originalnog uzorka. Test Kolmogorov - Smirnova zasnovan na butstrap metodu možemo opisati sljedećim koracima.

Korak 1: Pretpostavimo da je konačni uzorak $\{x_i : i = 1, 2, \dots, n\}$ dat u rastućem poretku $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$. Tada kumulativne vjerovatnoće F_i mogu se izračunati korišćenjem proste ocjene na sljedeći način:

$$F_i = \frac{i}{n}.$$

U ovom slučaju, empirijska funkcija raspodjele uzorka se ocjenjuje sa

$$\hat{F}_n(x) = \begin{cases} 0, & x < x_{(1)}, \\ \frac{i}{n}, & x_{(i)} \leq x < x_{(i+1)}, \quad i \in \{1, 2, \dots, n-1\}, \\ 1, & x \geq x_{(n)}. \end{cases} \quad (1.11)$$

Po metodu testa Kolmogorov - Smirnova, maksimalna razlika između empirijske funkcije raspodjele i teorijske funkcije raspodjele može se dobiti iz sljedeće formule

$$D_n = \sup_{x \in \mathbb{R}} |\hat{F}_n(x) - F(x)|. \quad (1.12)$$

Korak 2: Reuzorkovani skup podataka $X^{(1)} = \{x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)}\}$ može se dobiti primjenom butstrap metode, ponovnim uzorkovanjem iz originalnog uzorka $X = \{x_i : i = 1, 2, \dots, n\}$.

Korak 3: Ponoviti korak 1 uzimajući reuzorkovani skup podataka $X^{(1)} = \{x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)}\}$ kao posmatrani uzorak. To znači da se K - S test primjenjuje na reuzorkovani skup podataka. Odnosno, empirijska funkcija raspodjele $\hat{F}_n^{(1)}(x)$ može se izračunati primjenom formule (1.11). Dok maksimalna razlika D_n , koju označavamo sa $D_n^{(1)}$, može se dobiti primjenom formule (1.12).

Korak 4: Ponavljajući M puta korak 2 i korak 3, dobijamo M razlika označenih kao $\{D_n^{(1)}, D_n^{(2)}, \dots, D_n^{(M)}\}$.

Korak 5: Određivanjem odgovarajuće teorijske raspodjele od $\{D_n^{(1)}, D_n^{(2)}, \dots, D_n^{(M)}\}$, ocjenom parametara raspodjele i izvođenjem testa maksimalne vjerodostojnosti, možemo dobiti neke karakteristične vrijednosti u statistici kao što su aritmetička sredina, standardna devijacija i interval povjerenja, koje se smatraju mjerom K - S testa.

Razmotrimo sada familiju raspodjela sa parametrom lokacije $\mu \in \mathbb{R}$ i parametrom skaliranja $\sigma > 0$, tj. familiju raspodjela za koje važi

$$F_{\mu, \sigma}(x) = F_{0,1}\left(\frac{x - \mu}{\sigma}\right). \quad (1.13)$$

Pogodnost raspodjela iz ove porodice je ta što možemo odabrati ocjene $\hat{\mu}$ i $\hat{\sigma}$ koje su respektivno lakacijski-skalaro invarijante i skalaro invarijantne. Ocjena T je lakacijski-skalaro invarijanta ako za $a > 0$ i $b \in \mathbb{R}$

$$T(aX_1 + b, aX_2 + b, \dots, aX_n + b) = aT(X_1, X_2, \dots, X_n) + b.$$

Odnosno, ocjena T je skalaro invarijantna ako

$$T(aX_1 + b, aX_2 + b, \dots, aX_n + b) = aT(X_1, X_2, \dots, X_n).$$

Ocjene dobijene metodom maksimalne vjerodostojnosti zadovoljavaju tražene invarijantnosti (za više detalja pogledati u [4]).

Ako imamo lokacijski-skalaro invarijantu ocjenu $\hat{\mu}$ i skalaro invarijantnu ocjenu $\hat{\sigma}$ pokazuje se da tada raspodjele $\frac{\hat{\mu} - \mu}{\hat{\sigma}}$ i $\frac{\hat{\sigma}}{\sigma}$ ne zavise od μ i σ .

Pretpostavimo da želimo da testiramo da li slučajni uzorak X_1, X_2, \dots, X_n dolazi iz familije raspodjela $\mathcal{F} = \{F_{\mu, \sigma}, \mu \in \mathbb{R}, \sigma > 0\}$ koje zadovoljavaju (1.13). Odnosno, razmatramo sljedeći problem

$$H_0 : F_{\mu, \sigma} \in \mathcal{F} \text{ protiv } H_1 : F_{\mu, \sigma} \notin \mathcal{F}.$$

Primjenom testa Kolmogorov - Smirnova sa ocijenjenim parametrima dobijamo

$$\begin{aligned} D_n(t) &= \sup_{t \in \mathbb{R}} |\hat{F}_n(t) - F_{\hat{\mu}, \hat{\sigma}}(t)| = \sup_{t \in \mathbb{R}} \left| \frac{1}{n} \sum_{i=1}^n \mathbf{I}\{X_i \leq t\} - F_{\hat{\mu}, \hat{\sigma}}(t) \right| = \\ &= \sup_{t \in \mathbb{R}} \left| \frac{1}{n} \sum_{i=1}^n \mathbf{I}\{Z_i \leq F_{\hat{\mu}, \hat{\sigma}}(t)\} - F_{\hat{\mu}, \hat{\sigma}}(t) \right|, \end{aligned}$$

gdje je $Z_i = F_{\hat{\mu}, \hat{\sigma}}(X_i) = F_{0,1}\left(\frac{X_i - \hat{\mu}}{\hat{\sigma}}\right) = F_{0,1}\left(\frac{\frac{X_i - \mu}{\sigma} - \frac{\hat{\mu} - \mu}{\sigma}}{\frac{\hat{\sigma}}{\sigma}}\right)$ i \mathbf{I} je indikator funkcija.

Dakle, raspodjela Z_i neće zavisiti od nepoznatih parametara μ i σ . Raspodjela test statistike pod hipotezom H_0 rijetko se može izračunati analitički, ali može se aproksimirati pomoću Monte Karlove simulacije na sljedeći način (pogledati u [10]):

korak 1: generisati slučajni izorak X_1, X_2, \dots, X_n iz raspodjele $F_{0,1}$;

korak 2: izračunati vrijednost test statistike za generisani uzorak, tj. odrediti vrijednosti za $\hat{\mu} = \hat{\mu}(X_1, X_2, \dots, X_n)$ i $\hat{\sigma} = \hat{\sigma}(X_1, X_2, \dots, X_n)$, zatim izračunati $F_{0,1}\left(\frac{X_i - \hat{\mu}}{\hat{\sigma}}\right)$ i potom izračunati D_n ;

korak 3: ponovljiti korake 1 i 2 ogroman broj puta.

1.6 Specijalne funkcije

Funkcije koje su dovoljno značajne da dobiju sopstveno ime su poznate kao specijalne funkcije. Tu spadaju dobro poznate logaritamske, eksponencijalne i trigonometrijske funkcije i proširuje se na gama, beta, zeta funkciju, potom na klasu ortogonalnih polinoma i još na mnoge druge. Matematičari, inženjeri i fizičari koriste brojne formule i indentitete ovih funkcija. Specijalne funkcije imaju ogromnu primjenu u teorijskoj matematici, kako i u primjenjenim oblastima, poput elektrotehnike, dinamike fluida, momenta inercije, kvantne mehanike i sl.

U srcu teorije specijalnih funkcija leži hipergeometrijska funkcija, jer se sve klasične specijalne funkcije mogu izraziti preko ove moćne funkcije. Hipergeometrijska funkcija pored toga što se može zapisati u obliku reda ima i integralnu reprezentaciju, na taj način pruža idealan alat za brojne transformacije.

Ako je hipergeometrijska funkcija u srcu teorije specijalnih funkcija, onda je gama funkcija centralna u teoriji hipergeometrijskih funkcija.

1.6.1 Gama funkcija

Ojler se bavio problemom pronalaska neprekidne funkcije od promjenljive x koja je jednaka $n!$ kada je $x = n$. Da bi smo došli do Ojlerove generalizacije faktoriijela, pretpostavimo da su $x \geq 0$ i $n \geq 0$ cijeli brojevi. Tada je

$$x! = \frac{(x+n)!}{(x+1)_n}, \quad (1.14)$$

gdje je $(a)_n$ Pochhammerov simbol definisan sa

$$(a)_n = a(a+1)(a+2)\dots(a+n-1), \text{ za } n > 0, (a)_0 = 1,$$

gdje je a bilo koji kompleksan ili realan broj.

Definicija 1.27. *Za sve kompleksne brojeve $x \neq 0, -1, -2, \dots$ gama funkcija $\Gamma(x)$ definisana je sa*

$$\Gamma(x) = \lim_{k \rightarrow \infty} \frac{k! k^{x-1}}{(x)_k}.$$

Posljedice definicije 1.27 su:

$$x\Gamma(x) = \Gamma(x+1), \Gamma(n+1) = n! \text{ i } \Gamma(1) = 1.$$

Definicija 1.28. *Beta integral definisan je za $\operatorname{Re}(x) > 0$ i $\operatorname{Re}(y) > 0$ sa*

$$B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt. \quad (1.15)$$

Teorema 1.15 ([3]).

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

Kao posljedicu prethodne teoreme dobijamo integralnu reprezentaciju gama funkcije.

Posljedica 1.2 ([3]). *Za $\operatorname{Re}(x) > 0$*

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt. \quad (1.16)$$

Beta integral ima nekoliko korisnih formi koje se dobijaju zamjenom promjenljivih u (1.15). Tako zamjenom $t = \frac{s}{s+1}$ dobijamo beta integral na polu-ravni

$$\int_0^{\infty} \frac{s^{x-1}}{(s+1)^{x+y}} ds = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}. \quad (1.17)$$

Uzmimo sada da je $t = \sin^2 \theta$, tada je

$$\int_0^{\frac{\pi}{2}} \sin^{2x-1} \theta \cos^{2y-1} \theta d\theta = \frac{\Gamma(x)\Gamma(y)}{2\Gamma(x+y)}.$$

Zamjenom $x = y = \frac{1}{2}$ dobijamo sljedeći rezultat

$$\frac{\Gamma(\frac{1}{2})^2}{2\Gamma(1)} = \frac{\pi}{2}$$

ili

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

Sljedeća formula povezuje gama funkciju sa sinusnom funkcijom.

Teorema 1.16 (Ojlerova formula refleksije, ([3])).

$$\Gamma(x)\Gamma(1-x) = \frac{\pi}{\sin \pi x}. \quad (1.18)$$

Da bi se izvela korisna svojstva gama funkcije često se koristi logaritam od gama funkcije. Zbog toga često možemo sresti izvod logaritma od gama funkcije, tj. popularnu psi ili digama funkciju.

Definicija 1.29. *Digama funkcija je definisana za $x \neq 0, -1, -2, \dots$ sa*

$$\psi(x) = \frac{d}{dx} \log \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}.$$

Proširićemo sada definiciju beta funkcije na tri ili više promjenljivih.

Definicija 1.30. *Neka je $a = (a_1, a_2, \dots, a_k) \in \mathbb{C}^k$, $a_i \neq \{0, -1, -2, \dots\}$, $i = \overline{1, k}$, $k \in \mathbb{N}$ i $k \geq 2$. Beta funkcija sa k promjenljivih definisana je sa*

$$B(a_1, a_2, \dots, a_k) = \frac{\Gamma(a_1)\Gamma(a_2) \cdots \Gamma(a_k)}{\Gamma(a_1 + a_2 + \dots + a_k)}. \quad (1.19)$$

Napomena 1. Za $k = 2$ dobijamo standardnu beta funkciju definisanu sa (1.15).

1.6.2 Gausova hipergeometrijska funkcija

Kažemo da je red $\sum c_n$ hipergeometrijski red ako je $\frac{c_{n+1}}{c_n}$ racionalna funkcija od promjenljive n .

Definicija 1.31. Za $z \in \mathbb{C}$, $|z| < 1$, gdje su a, b i c realni ili kompleksni parametri, takvi da $c \neq 0, -1, -2, \dots$, Gausov hipergeometrijski red je beskonačan red (u oznaci ${}_2F_1(a, b; c; z)$) definisan sa

$${}_2F_1(a, b; c; z) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{k! (c)_k} z^k. \quad (1.20)$$

Kada je $a = c$ i $b = 1$ (1.20) postaje osnovni geometrijski red, što objašnjava terminologiju, ali (1.20) nije definisano kada je parametar u imeniocu negativan cio broj, jer ovo dovodi do dijeljenja sa nulom. Proučimo konvergenciju Gausovog hipergeometrijskog reda.

Teorema 1.17 ([3]). Gausov hipergeometrijski red ${}_2F_1(a, b; c; z)$ je apsolutno konvergentan kada je $|z| < 1$, divergentan kada je $|z| > 1$ i konvergentan na granici $|z| = 1$ za $\operatorname{Re}(c - a - b) > 0$.

Dokaz. Razmotrimo ${}_2F_1(a, b; c; z) = \sum_{k=0}^{\infty} t_k$. Kako je $\frac{(\alpha)_{k+1}}{(\alpha)_k} = \alpha + k$, po definciji Gausove funkcije imamo da je

$$\left| \frac{t_{k+1}}{t_k} \right| = \frac{|z|(1 + \frac{|a|}{k})(1 + \frac{|b|}{k})}{(1 + \frac{|c|}{k})(1 + \frac{1}{k})}.$$

Dakle, $|z| = \lim_{k \rightarrow \infty} \left| \frac{t_{k+1}}{t_k} \right|$, tako da red konvergira za $|z| < 1$, a divergira za $|z| > 1$. Da bi ispitali slučaj kada je $|z| = 1$, neka je $\delta = \frac{1}{2}\operatorname{Re}(c - a - b) > 0$ i uporedimo članove reda $1 + \sum_{k=1}^{\infty} \left| \frac{(a)_k (b)_k z^k}{k! (c)_k} \right|$ sa članovima konvergentnog reda $\sum_{k=1}^{\infty} \frac{1}{k^{1+\delta}}$.

Ako je $|z| = 1$, imamo da je

$$\begin{aligned} \lim_{k \rightarrow \infty} \left| \frac{k^{1+\delta} (a)_k (b)_k}{(c)_k k!} \right| &= \lim_{k \rightarrow \infty} \left| \frac{(a)_k}{(k-1)! k^a} \frac{(b)_k}{(k-1)! k^b} \frac{(k-1)! k^c}{(c)_k} \frac{(k-1)! k^{1+\delta}}{k! k^{c-a-b}} \right| = \\ &= \left| \frac{1}{\Gamma(a)} \frac{1}{\Gamma(b)} \Gamma(c) \right| \lim_{k \rightarrow \infty} \left| \frac{1}{k^{c-a-b-\delta}} \right|. \end{aligned}$$

Kako je $\operatorname{Re}(c - a - b - \delta) = 2\delta - \delta = \delta > 0$, ova granična vrijednost je nula, i dakle iz konvergencije reda $\sum_{k=1}^{\infty} \frac{1}{k^{1+\delta}}$ slijedi konvergencija hipergeometrijskog reda ${}_2F_1(a, b; c; z)$ na granici $|z| = 1$ kada je $\operatorname{Re}(c - a - b) > 0$. \square

Slater [43] daje sljedeće dodatne kriterijume konvergencije. Za $|z| = 1$, $z \neq 1$, Gausov red divergira kada je $\operatorname{Re}(c - a - b) < -1$ i konvergira (ali ne apsolutno) kada je $-1 < \operatorname{Re}(c - a - b) \leq 0$. Takođe, ako je $\operatorname{Re}(c - a - b) = -1$ onda red konvergira ako je $\operatorname{Re}(a + b) > \operatorname{Re}(ab)$ u suprotnom divergira.

Teorema 1.18 (Ojlerova integralna reprezentacija, [3]). *Neka je $\operatorname{Re}(c) > \operatorname{Re}(b) > 0$ i $|x| < 1$. Tada je*

$${}_2F_1(a, b; c; x) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 t^{b-1} (1-t)^{c-b-1} (1-xt)^{-a} dt.$$

Pomoću konvergentnog Gausovog reda (1.20) definišemo hipergeometrijsku funkciju.

Definicija 1.32. *Za $c \neq 0, -1, -2, \dots$ Gausov hipergeometrijski red naziva se hipergeometrijska funkcija, definisana sa*

$${}_2F_1(a, b; c; z) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{k! (c)_k} z^k,$$

za $|z| < 1$, i produživanjem svuda drugdje.

1.6.3 H-funkcija

H-funkcija u literaturi je poznata pod nazivom Foksova H funkcija pronašla je primjenu u velikom broju problema povezanih sa inženjeringom, komunikacijama, integralnim jednačinama, u mnogim oblastima teorijske fizike, statistike i sl. U ovom paragrafu damo kratak pregled osnovnih njenih karakteristika, koje su preuzete iz [26].

Definicija 1.33. *Foksova H funkcija se definiše na sljedeći način*

$$\begin{aligned} H_{p,q}^{m,n}(z) &= H_{p,q}^{m,n} \left[z \left| \begin{matrix} (a_1, A_1) & (a_2, A_2) & \dots & (a_p, A_p) \\ (b_1, B_1) & (b_2, B_2) & \dots & (b_q, B_q) \end{matrix} \right. \right] = \\ &= \frac{1}{2\pi i} \int_L \frac{\prod_{j=1}^m \Gamma(b_j + B_j s) \prod_{j=1}^n \Gamma(1 - a_j - A_j s)}{\prod_{j=m+1}^q \Gamma(1 - b_j - B_j s) \prod_{j=n+1}^p \Gamma(a_j + A_j s)} z^{-s} ds, \end{aligned} \quad (1.21)$$

gdje je $i = \sqrt{-1}$, $z \neq 0$ i gdje su $m, n, p, q \in \mathbb{N}_0$, $0 \leq n \leq p$, $1 \leq m \leq q$, $A_i, B_j \in \mathbb{R}_+$, $a_i, b_j \in \mathbb{R}$, $i = 1, \dots, p$, $j = 1, \dots, q$. L je odgovarajuća kontura koja razdvaja polove

$$\xi_{jv} = -\left(\frac{b_j + v}{B_j}\right), \quad j = 1, 2, \dots, m; \quad v = 0, 1, 2, \dots \quad (1.22)$$

od gama funkcije $\Gamma(b_j + sB_j)$ od polova

$$w_{\lambda k} = \left(\frac{1 - a_\lambda + k}{A_\lambda}\right), \quad \lambda = 1, 2, \dots, n; \quad k = 0, 1, 2, \dots \quad (1.23)$$

od gama funkcije $\Gamma(1 - a_\lambda - sA_\lambda)$, tako da

$$A_\lambda(b_j + v) \neq B_j(a_\lambda - k - 1), \quad j = 1, 2, \dots, m; \quad \lambda = 1, 2, \dots, n; \quad v, k = 0, 1, 2, \dots$$

Teorema 1.19 ([26]). Foksova H funkcija je analitička funkcija po z i postoji u sljedećim slučajevima:

1. $q \geq 1$, $\mu > 0$, Foksova H funkcija postoji za svako $z \neq 0$;
2. $q \geq 1$, $\mu = 0$, Foksova H funkcija postoji za svako $0 < |z| < \beta$;
3. $q \geq 1$, $\mu = 0$, $\text{Re}(\delta) < -1$, Foksova H funkcija postoji za svako $|z| = \beta$;
4. $p \geq 1$, $\mu < 0$, Foksova H funkcija postoji za svako $z \neq 0$;
5. $p \geq 1$, $\mu = 0$, Foksova H funkcija postoji za svako $|z| > \beta$;
6. $p \geq 1$, $\mu = 0$ i $\text{Re}(\delta) < -1$ Foksova H funkcija postoji za svako $|z| = \beta$;
7. $\alpha > 0$, $|\arg z| < \frac{1}{2}\pi\alpha$, Foksova H funkcija postoji za svako $z \neq 0$;
8. $\alpha = 0$, $\gamma\mu + \text{Re}(\delta) < -1$, Foksova H funkcija postoji za $\arg z = 0$ i $z \neq 0$, gdje su

$$\begin{aligned} \beta &= \left\{ \prod_{j=1}^p (A_j)^{-A_j} \right\} \left\{ \prod_{j=1}^q (B_j)^{B_j} \right\}; \\ \mu &= \sum_{j=1}^q B_j - \sum_{j=1}^p A_j; \\ \delta &= \sum_{j=1}^q b_j - \sum_{j=1}^p a_j + \frac{p-q}{2}; \\ \alpha &= \sum_{j=1}^n A_j - \sum_{j=n+1}^p A_j + \sum_{j=1}^m B_j - \sum_{j=m+1}^q B_j. \end{aligned}$$

1.6.4 Specijalni slučajevi H-funkcije

Definicija 1.34. Meijerova G ili G -funkcija se definiše na sljedeći način

$$G_{p,q}^{m,n}(z) = G_{p,q}^{m,n} \left(z \left| \begin{matrix} a_1 & a_2 & \dots & a_p \\ b_1 & b_2 & \dots & b_q \end{matrix} \right. \right) = \frac{1}{2\pi i} \int_L \frac{\prod_{j=1}^m \Gamma(b_j + s) \prod_{j=1}^n \Gamma(1 - a_j - s)}{\prod_{j=m+1}^q \Gamma(1 - b_j - s) \prod_{j=n+1}^p \Gamma(a_j + s)} z^{-s} ds, \quad (1.24)$$

gdje su $0 \leq m \leq q$, $0 \leq n \leq p$, a_j , $j = 1, 2, \dots, p$ i b_j , $j = 1, 2, \dots, q$ su kompleksni brojevi takvi da

$$a_j - b_h \neq 0, 1, \dots; \quad j = 1, 2, \dots, n; \quad h = 1, 2, \dots, m.$$

Parametri su takvi da razdvajaju tačke

$$s = -(b_j + v), \quad j = 1, 2, \dots, m, \quad v \in \mathbb{N}_0$$

i

$$s = -(a_j - v - 1), \quad j = 1, 2, \dots, n, \quad v \in \mathbb{N}_0.$$

Ovdje L je ista kontura kao u definiciji H-funkcije 1.33.

U slučaju kada je $A_j = B_k = C$, $C > 0$, $j = 1, 2, \dots, p$, $k = 1, 2, \dots, q$ u (1.21) Foksova H funkcija se svodi na Meijerovu G funkciju, imamo sljedeću relaciju:

$$H_{p,q}^{m,n} \left[z \left| \begin{matrix} (a_1, C) & (a_2, C) & \dots & (a_p, C) \\ (b_1, C) & (b_2, C) & \dots & (b_q, C) \end{matrix} \right. \right] = \frac{1}{C} G_{p,q}^{m,n} \left(z^{\frac{1}{C}} \left| \begin{matrix} a_1 & a_2 & \dots & a_p \\ b_1 & b_2 & \dots & b_q \end{matrix} \right. \right). \quad (1.25)$$

Brojne analitičke funkcije su specijalni slučajevi Foksove H funkcije. Ove analitičke funkcije pojavljuju se u raznim problemima koju se javljaju u teorijskim i primjenjenim granama matematike, statistike i inženjerskih nauka. Mi ćemo ovdje predstaviti samo one specijalne slučajeve Foksove H funkcije koji su bile značajni u našem istaživanju.

$$H_{1,1}^{1,1} \left[z \left| \begin{matrix} (1-u, 1) \\ (0, 1) \end{matrix} \right. \right] = \Gamma(u)(1+z)^{-u}, \quad |z| < 1, \quad (1.26)$$

$$H_{2,2}^{1,2} \left[z \left| \begin{matrix} (1-a, 1) & (1-b, 1) \\ (0, 1) & (1-c, 1) \end{matrix} \right. \right] = \frac{\Gamma(a)\Gamma(b)}{\Gamma(c)} {}_2F_1(b, a; c; -z), \quad (1.27)$$

$$H_{1,2}^{1,1} \left[-z \left| \begin{array}{cc} (0, 1) & -- \\ (0, 1) & (1 - \beta, \alpha) \end{array} \right. \right] = E_{\alpha, \beta}(z), \quad (1.28)$$

$$H_{1,3}^{2,0} \left[\frac{z^2}{4} \left| \begin{array}{ccc} (\frac{a-k-1}{2}, 1) & -- & -- \\ (\frac{a-k}{2}, 1) & (\frac{a+k}{2}, 1) & (\frac{a-k-1}{2}, 1) \end{array} \right. \right] = \left(\frac{z}{2} \right)^a Y_v(z), \quad (1.29)$$

gdje je $E_{\alpha, \beta}$ Mittag - Lefflerova funkcija koja se definiše na sljedeći način:

$$E_{\alpha, \beta}(z) = \sum_{k=0}^{+\infty} \frac{z^k}{\Gamma(\alpha k + \beta)}, \quad \alpha, \beta, z \in \mathbb{C},$$

i $Y_v(z)$ je modifikovana Beselova funkcija druge vrste ili Neumanova funkcija definisana sa:

$$Y_v(z) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{\Gamma(s - \frac{v}{2})\Gamma(s + \frac{v}{2})}{\Gamma(s - \frac{v+1}{2})\Gamma(\frac{3+v}{2} - s)} \left(\frac{z^2}{4} \right)^{-s} ds, \quad -3 < \operatorname{Re}(v) < -1.$$

1.7 Furijeova i Melinova transformacija

Furijeova transformacija \mathcal{F} je preslikavanje iz prostora funkcija kompleksnih vrijednosti u funkcije kompleksnih vrijednosti.

Definicija 1.35. *Neka je $f : \mathbb{R} \rightarrow \mathbb{C}$ integrabilna funkcija. Furijeova transformacija od f se označava sa $\mathcal{F}[f] = \hat{f}$ gdje je*

$$\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-ikx} dx.$$

Slično, inverzna Furijeova transformacija od f se označava sa $\mathcal{F}^{-1}[f] = \check{f}$ gdje je

$$\check{f}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(k) e^{ikx} dk. \quad (1.30)$$

Teorema 1.20 (Inverzna Furijeova transformacija, [45]). *Ako su f i f' dio po dio neprekidne funkcije, tada je $\mathcal{F}^{-1}[\mathcal{F}f] = f$ i $\mathcal{F}[\mathcal{F}^{-1}f] = f$. Odnosno,*

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \hat{f}(k) e^{ikx} dk \text{ i } f(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \check{f}(x) e^{-ikx} dx.$$

Teorema 1.21 (Planšerelova teorema, [45]). *Ako je f integrabilna i kvadratno integrabilna funkcija, tada je*

$$\int_{-\infty}^{+\infty} |f(x)|^2 dx = \int_{-\infty}^{+\infty} |\hat{f}(k)|^2 dk = \int_{-\infty}^{+\infty} |\check{f}(x)|^2 dx.$$

Definicija 1.36. *Konvolucija od f i g je funkcija $f * g$ definisana sa*

$$(f * g)(x) = \int_{-\infty}^{+\infty} f(x-y)g(y)dy = \int_{-\infty}^{+\infty} f(y)g(x-y)dy.$$

Teorema 1.22 (Osobine Furijeove transformacije, [45]). *Ako su f i g integrabilne, onda*

- | | |
|---|--|
| 1. $\mathcal{F}[f(x-a)] = e^{-ika} \hat{f}(k);$ | 5. $\mathcal{F}[f'(x)] = ik \hat{f}(k);$ |
| 2. $\mathcal{F}[f(x)e^{ibx}] = \hat{f}(k-b);$ | 6. $\mathcal{F}[xf(x)] = i \hat{f}'(k);$ |
| 3. $\mathcal{F}[f(\lambda x)] = \lambda ^{-1} \hat{f}(\lambda^{-1}k);$ | 7. $\mathcal{F}[(f * g)(x)] = \sqrt{2\pi} \hat{f}(k) \hat{g}(k);$ |
| 4. $\mathcal{F}[\hat{f}(x)] = f(-k);$ | 8. $\mathcal{F}[f(x)g(x)] = \frac{1}{\sqrt{2\pi}} (\hat{f} * \hat{g})(k).$ |

Za razliku od Furijeovih transformacija koje su uvedene za rješavanje fizičkih problema, Melinova transformacija nastala je u matematičkom kontekstu.

Podsjetimo se da se $L^c(0, \infty)$ sastoji od realno ili kompleksno-vrijednosnih funkcija od realne promjenljive koje su neprekidne na intervalu $(0, \infty)$, osim na prebrojivom skupu izolovanih tačaka gdje ove funkcije mogu uzeti vrijednost beskonačno i zbog toga Rimanov nesvojstveni integral apsolutno konvergira na $(0, \infty)$.

Definicija 1.37. *Neka je $f(t)$ funkcija definisana na pozitivnoj realnoj osi $0 < t < \infty$. Melinova transformacija \mathcal{M} je preslikavanje funkcije f na funkciju F definisano u kompleksnoj ravni relacijom*

$$\mathcal{M}[f; s] \equiv F(s) = \int_0^{\infty} f(t) t^{s-1} dt, \quad (1.31)$$

i inverzna Melinova transformacija se definiše kao

$$f(t) = \mathcal{M}^{-1}[F(s); t] = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} F(s)t^{-s} ds, \quad t > 0, \quad a = \operatorname{Re}(s). \quad (1.32)$$

Veza Melinove i Furijeove transformacije data je relacijom

$$\mathcal{M}[f(t); s] = \int_0^{\infty} f(t)t^{s-1} dt = \int_{-\infty}^{+\infty} f(e^t)e^{it(-is)} dt = \mathcal{F}[f(e^t); -is].$$

Funkcija F je Melinova transformacija funkcije f . Vidimo da je integral u (1.31) dobro definisan, na primjer za funkcije $f \in L^c(a_1, a_2)$, $0 < a_1 < a_2 < \infty$, neprekidne na intervalima $(0, a_1]$, $[a_2, \infty)$ i koje zadovoljavaju ocjene $|f(t)| \leq Mt^{-\gamma_1}$, za $0 < t < a_1$, i $|f(t)| \leq Mt^{-\gamma_2}$, za $t > a_2$, gdje je M konstanta i $\gamma_1 < \gamma_2$. Kada su zadovoljeni ovi uslovi Melinova transformacija F postoji i ona je analitička funkcija u oblasti $a_1 < \operatorname{Re}(s) < a_2$.

Ako je f dio po dio neprekidna, $f(t)t^{a-1} \in L^c(0, \infty)$ i njena Melinova transformacija F data je sa (1.31), tada formula (1.32) važi u svim tačkama neprekidnosti funkcije f .

Melinova konvolucija definiše se kao

$$(f *_M g)(x) = \int_0^{\infty} f\left(\frac{x}{t}\right)g(t)\frac{dt}{t}. \quad (1.33)$$

Teorema 1.23 ([36]). *Neka je $f(t)t^{a-1} \in L^c(0, \infty)$ i $g(t)t^{a-1} \in L^c(0, \infty)$. Melinova konvolucija $h = (f *_M g)$ data sa (1.33) je dobro definisana, $h(x) = hx^{a-1} \in L^c(0, \infty)$ i*

$$\mathcal{M}[(f *_M g)(x); s] = \mathcal{M}[f(t); s] \times \mathcal{M}[g(t); s],$$

tj. Melinova konvolucija funkcija f i g jednaka je proizvodu njihovih Melinovih transformacija. Štaviše, važi Parsevalova jednakost

$$\int_0^{\infty} f\left(\frac{x}{t}\right)g(t)\frac{dt}{t} = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} F(s)G(s)x^{-s} ds. \quad (1.34)$$

1.8 Numeričke metode

U ovom poglavlju dajemo pregled pojmova i rezultata iz numeričke analize, koje ćemo koristiti u radu. Ove teme su detaljno obrađene u veoma bogatoj literaturi [16, 22, 29].

Prvo ćemo se pozabaviti tehnikom konačnih razlika za numeričko diferenciranje diskretnih podataka. Razmatramo formule za izračunavanje izvoda glatke funkcije, ali samo na diskretnom skupu međusobno različitih tačaka $x_0, x_1, x_2, \dots, x_N$ na realnoj osi. Fokusiraćemo se na tehnike konačnih razlika za dobijanje numeričkih vrijednosti izvoda u tačkama mreže.

1.8.1 Numeričko diferenciranje: metod konačnih razlika

Formule konačnih razlika mogu se lako izvesti iz Tejlorovog reda. Počnimo sa najjednostavnijom aproksimacijom izvoda funkcije $f(x)$ u tačkama x_j , koristeći Tejlorov red

$$f(x_{j+1}) = f(x_j) + (x_{j+1} - x_j)f'(x_j) + \frac{(x_{j+1} - x_j)^2}{2}f''(x_j) + \dots$$

Preuređenjem dobijamo

$$f'(x_j) = \frac{f(x_{j+1}) - f(x_j)}{\Delta x_j} - \frac{\Delta x_j}{2}f''(x_j) + \dots \quad (1.35)$$

gdje je $\Delta x_j = x_{j+1} - x_j$ rastojanje između uzastopnih čvorova. Prvi član na desnoj strani u (1.35) je aproksimacija izvoda konačnom razlikom. Sljedeći član je odgovarajuća greška aproksimacije. Označimo sa h rastojanje između uzastopnih čvorova. Ekvidistantna mreža čvorova definiše se pomoću početnog čvora $x_0 \in \mathbb{R}$ i svog koraka $h > 0$. Čvorovi su $x_j = x_0 + jh$, $j = 0, \dots, N$, gdje je $N \geq 1$.

Formula (1.35) se obično zapisuje u sljedećem obliku za ekvidistantnu mrežu čvorova sa korakom h

$$f'_j = \frac{f_{j+1} - f_j}{h} + O(h), \quad (1.36)$$

što se naziva metoda prednje razlike prvog reda, gdje $O(h)$ označava da je greška manja od ch za neku konstantu c . Ovo je isti izraz koji se u analizi koristi za definisanje izvoda, s tim što u analizi imamo da $h \rightarrow 0$, a ovdje h je konačno.

Eksponent od h u $O(h^\alpha)$ je red tačnosti metoda, pa kažemo da je red preciznosti ili tačnosti metoda jednak α . On nam ukazuje koliko brzo možemo da poboljšamo tačnost u zavisnosti od rastojanja između uzastopnih čvorova. Na primjer, ako u formuli prvog reda (1.36) umjesto h stavimo $\frac{h}{2}$, tada se greška (nazvana greška skraćivanja) smanjuje aproksimativno na polovinu prethodne greške. Primjetimo da kada pričamo o grešci skraćivanja metoda konačnih razlika, uvijek se pozivamo na vodeći član greške sa implikacijom da su članovi većeg reda u Tejlorovom razvoju mnogo manji od vodećeg člana.

Razmotrimo druge popularne formule metoda konačnih razlika. Razvojem f_{j-1} oko x_j dobijamo

$$f'_j = \frac{f_j - f_{j-1}}{h} + O(h),$$

što je takođe šema prvog reda, koja se naziva metod zadnje razlike. Šeme većeg reda (koje su preciznije) mogu se izvesti iz Tejlorovog razvoja funkcije f u različitim tačkama oko tačke x_j . Na primjer, simetrični metod dobijamo oduzimanjem dva Tejlorova razvoja. Pretpostavljajući ekvidistantnu mrežu imamo

$$f_{j+1} = f_j + hf'_j + \frac{h^2}{2}f''_j + \frac{h^3}{6}f'''_j + \dots \quad (1.37)$$

$$f_{j-1} = f_j - hf'_j + \frac{h^2}{2}f''_j - \frac{h^3}{6}f'''_j + \dots, \quad (1.38)$$

što vodi do

$$f'_j = \frac{f_{j+1} - f_{j-1}}{2h} - \frac{h^2}{6}f'''_j + \dots$$

Ovo je naravno formula drugog reda. Odnosno, ako korak mreže smanjimo dva puta očekujemo da će se greška skraćivanja aproksimativno smanjiti 4 puta. U opštem slučaju, možemo dobiti veću preciznost ako uključimo više tačaka. Formula četvrtog reda je

$$f'_j = \frac{f_{j-2} - 8f_{j-1} + 8f_{j+1} - f_{j+2}}{12h} + O(h^4). \quad (1.39)$$

Glavna poteškoća sa formulama višeg reda javlja se blizu granica domena. One zahtijevaju vrijednosti funkcije u tačkama izvan domena, koje su nepoznate.

Na primjer, ako su poznate vrijednosti funkcije f u tačkama x_0, x_1, \dots, x_N i tražimo izvod funkcije f u x_{-1} , formula (1.36) zahtijeva vrijednost od f u x_{-1} (pored x_0, x_1, x_2 i x_3) koja nije poznata. U praksi da bismo izbjegli ovaj problem koristimo formule nižeg reda ili nesimetrične formule blizu granice. Slično možemo izvesti formule za izvode višeg reda. Na primjer, formula drugog reda za simetrični metod može se dobiti sabiranjem izraza (1.37) i (1.38), dva f'_j člana se poništavaju i nakon manjeg sređivanja dobijamo

$$f''_j = \frac{f_{j+1} - 2f_j + f_{j-1}}{h^2} + O(h^2).$$

1.8.2 Metod kvadrature

Numeričko izračunavanje jednodimenzionalnih integrala, odnosno kvadrature, jedna je od najstarijih grana numeričke analize. U savremenoj primjenjenoj matematici i statistici, kvadratura je toliko rasprostranjena da su čak i ručni kalkulatori rutinski programirani da je izvede.

U analizi, integral funkcije $I(f) = \int f(x)dx$ se definiše kao granična vrijednost gornje i donje Rimanove sume. Za integraciju funkcije $f(x)$ nad intervalom $[a, b]$, potrebna pretpostavka je neprekidnost funkcije f , tako da granična vrijednost lijeve i desne Rimanove sume

$$R_n^-(f) = \sum_{i=0}^{n-1} hf(a + ih) \text{ i } R_n^+(f) = \sum_{i=1}^n hf(a + ih),$$

gdje je $h = \frac{b-a}{n}$, konvergira ka integralu $I(f)$ kada $n \rightarrow \infty$.

Kao metod numeričke integracije može se koristiti ili lijeva ili desna Rimanova suma. U ovom slučaju greška prilikom aproksimacije integrala je:

$$|I(f) - R_n(f)| \leq (b-a)w(h), \text{ gdje je } w(h) = \sup_{|x-y|<h} |f(x) - f(y)|.$$

Uzimajući srednju vrijednost između lijeve i desne Rimanove sume, dobijamo jedan od najčešćih metoda numeričke integracije trapezno pravilo

$$T_n(h) = \frac{h}{2} \left[f(a) + 2 \sum_{i=1}^{n-1} f(a + ih) + f(b) \right]. \quad (1.40)$$

Naziv trapez potiče od linearne aproksimacije funkcije f na svakom podintervalu $[a + (i-1)h, a + ih]$, formirajući trapez čije su dvije strane u krajnjim tačkama i intervalom u osnovi. Gledajući na trapezno pravilo kao na srednju vrijednost Rimanovih suma pokazuje se njegova sposobnost da integrali funkciju samo sa pretpostavkom neprekidnosti. S druge strane, ako gledamo na trapezno pravilo kao linearnu aproksimaciju funkcije f na svakom podintervalu dolazimo do sljedeće analize greške kada funkcija f ima neprekidan drugi izvod

$$I(f) - T_n(f) = -(b-a)^3 f''(\xi)/(12n^2) = -nh^3 f''(\xi)/12, \text{ za } \xi \in [a, b]. \quad (1.41)$$

Trapezna formula (1.40) ima grešku reda veličine h^2 , odnosno ima red veličine $\alpha = 2$. Ovaj rezultat pokazuje da kada se broj tačaka u kojima se procjenjuje funkcija udvostruči greška se svodi (grubo govoreći) na četvrtinu prethodne greške.

Jednostavna alternativa evaluaciji funkcije u krajnjim tačkama je procjena funkcije f u centru svakog podintervala $[a + (i-1)h, a + ih]$. Ovo vodi do pravila središnje tačke (eng. midpoint rule),

$$M_n(f) = h \sum_{i=1}^n f\left(a + (b-a)\frac{2i-1}{2n}\right).$$

Pravilo središnje tačke takođe samo aproksimira funkciju f pomoću stepene funkcije na svakom podintervalu, ali njegova brzina konvergencije je konkurentna trapeznom pravilu:

$$I(f) - M_n(f) = (b-a)^3 f''(\xi)/(24n^2) = nh^3 f''(\xi)/24, \text{ za } \xi \in [a, b]. \quad (1.42)$$

Sličnost u izrazima (1.41) i (1.42) poziva nas da uporedimo ove kvadrature formule. Razlika u znaku kod rezultata greški dovodi do interesantne formule za konveksne funkcije

$$M_n(f) \leq I(f) \leq T_n(f), \text{ ako je } f''(x) > 0 \text{ na } [a, b].$$

1.8.3 Baricentrična interpolacija

U ovom paragrafu govorićemo o zadatku interpolacije. Dat je skup tačaka i želimo da odredimo glatku funkciju koja prolazi kroz sve te tačke. Prvo ćemo pokazati metod Langranžove i Njutnove interpolacije, a potom ćemo govoriti o baricentričnoj interpolaciji.

1.8.3.1 Langranžova i Njutnova interpolacija

Razmotrimo funkciju $f : \mathbb{R} \rightarrow \mathbb{R}$. Neka je dato $n+1$, $n \geq 0$, međusobno različitih tačaka x_0, x_1, \dots, x_n na realnoj osi. Pretpostavimo da su date vrijednosti $f_j = f(x_j)$, $0 \leq j \leq n$. Neka Π_n označava vektorski prostor svih polinoma stepena najviše n . Želimo da odredimo polinom $p \in \Pi_n$, koji zadovoljava uslov

$$p(x_j) = f_j, \quad j = 0, 1, 2, \dots, n.$$

Za polinom p se kaže da je interpolaciona funkcija. Kažemo da želimo da riješimo zadatak o interpolaciji za podatke (x_j, f_j) , $0 \leq j \leq n$. Problem je dobro postavljen, odnosno ima jedinstveno rješenje u razmatranoj klasi Π_n . Štaviše, kao što je objašnjeno u gotovo

svakom uvodnom tekstu numeričke analize, rješenje se može napisati u Lagranžovom obliku [21]

$$p(x) = \sum_{j=0}^n f_j l_j(x), \quad l_j(x) = \frac{\prod_{k=0, k \neq j}^n (x - x_k)}{\prod_{k=0, k \neq j}^n (x_j - x_k)}. \quad (1.43)$$

Langranžov polinom l_j koji odgovara čvoru x_j ima svojstvo

$$l_j(x_k) = \begin{cases} 1, & j = k, \\ 0, & \text{ina}\check{\text{c}}e, \end{cases} \quad j, k = 0, 1, 2, \dots, n. \quad (1.44)$$

Neki nedostaci Langražove interpolacije su:

1. svaka evaluacija polinoma $p(x)$ zahtjeva $O(n^2)$ sabiranja i množenja;
2. dodavanje para novih podataka (x_{n+1}, f_{n+1}) zahtjeva ponovno izračunavanje od početka;
3. proračun je numerički nestabilan.

Uobičajeno se Lagranžov oblik za polinom $p(x)$ koristi u teoriji. Dok za proračune generalno se preporučuje da se koristi Njutnova formula, koja zahtjeva samo $O(n)$ flopova za svaku evaluaciju polinoma $p(x)$ nakon što se izračunaju neki brojevi, koji su nezavisni od tačke x . Definišemo flop, tj. operaciju sa pokretnim zarezom, kao operaciju množenja ili dijeljenja plus operacija sabiranja ili oduzimanja.

Njutnov pristup se sastoji od dva koraka. Prvo, potrebno je izračunati Njutnovu tabelu podijeljenih razlika

$$\begin{array}{ccccccc}
f[x_0] & & & & & & \\
& f[x_0, x_1] & & & & & \\
f[x_1] & & f[x_0, x_1, x_2] & & & & \\
& f[x_1, x_2] & & f[x_0, x_1, x_2, x_3] & & & \\
f[x_2] & & f[x_1, x_2, x_3] & & \vdots & \ddots & \\
& f[x_2, x_3] & & & & & f[x_0, x_1, \dots, x_n] \\
f[x_3] & & \vdots & & f[x_{n-3}, \dots, x_n] & \ddots & \\
& \vdots & & & & & \\
& & \vdots & & f[x_{n-2}, x_{n-1}, x_n] & & \\
& & & & & & \\
& & & & & & f[x_{n-1}, x_n] \\
& & & & & & \\
& & & & & & \\
& & & & & & f[x_n]
\end{array}$$

pomoću rekurzivne formule

$$f[x_j, x_{j+1}, \dots, x_{k-1}, x_k] = \frac{f[x_{j+1}, \dots, x_k] - f[x_j, \dots, x_{k-1}]}{x_k - x_j}$$

sa početnim uslovom $f[x_j] = f_j$. Potrebno nam je oko n^2 operacija oduzimanja i oko $\frac{n^2}{2}$ operacija dijeljenja.

Drugi korak je da za svako x izračunamo $p(x)$ pomoću Njutnove interpolacione formule

$$p(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) + \dots + f[x_0, x_1, \dots, x_n](x - x_0)(x - x_1) \cdots (x - x_{n-1}). \quad (1.45)$$

Ovo zahtjeva samo n flopova kada se izvrši množenje proizvoda prostih polinoma, što je mnogo manje od $O(n^2)$ što je potrebno za direktnu primjenu (1.43).

1.8.3.2 Poboljšana Langranžova formula

Potrebno je naglasiti da se Langražova formula (1.43) može modifikovati tako da su za evaluaciju potrebne $O(n)$ operacije kao što je slučaj kod Njutnove formule (1.45). Da bismo to odradili, dovoljno je primijetiti da se brojilac od l_j u (1.43) može zapisati kao

$$\prod_{k=0, k \neq j}^n (x - x_k) = \frac{l(x)}{x - x_j},$$

gdje je

$$l(x) = (x - x_0)(x - x_1) \cdots (x - x_n). \quad (1.46)$$

Ako definišemo baricentrične težine sa

$$\omega_j = \frac{1}{\prod_{k \neq j} (x_j - x_k)}, \quad j = 0, 1, 2, \dots, n, \quad (1.47)$$

tako da $\omega_j = \frac{1}{l'(x_j)}$, gdje je $l'(x_j)$ prvi izvod funkcije $l(x)$ u tački x_j . Sada l_j možemo zapisati u obliku

$$l_j(x) = l(x) \frac{\omega_j}{x - x_j}.$$

Primijetimo da svi sabirci u sumi (1.43) sadrže član $l(x)$, koji ne zavisi od j . Zbog toga taj član možemo izbaciti ispred sume tako da dobijamo

$$p(x) = l(x) \sum_{j=0}^n \frac{w_j}{x - x_j} f_j. \quad (1.48)$$

Sada i Langranžova interpolaciona formula zahtijeva $O(n^2)$ flopova za izračunavanje brojeva w_j i nakon što nam ti brojevi budu poznati potrebno je još $O(n)$ flopova za izračunavanje polinoma p . Rutishauser [39] je nazvao (1.48) prvom baricentričnom formulom.

Iz (1.47) zaključujemo da su za uključivanje novog čvora x_{n+1} potrebna dva proračuna:

1. podijeliti svaki ω_j , $j = 0, 1, \dots, n$ sa $x_j - x_{n+1}$ (jedan flop za svaku tačku), što će nas koštati $n + 1$ flopova;
2. izračunati ω_{n+1} pomoću formule (1.47), gdje nam treba još $n + 1$ flopova.

Langranžova interpolaciona formula se tako može poboljšati sa $O(n)$ flopova. Rekursivnom primjenom ovog poboljšanja uz minimiziranje broja dijeljenja, dobijamo sljedeći algoritam za izračunavanje $\omega_j = \omega_j^{(n)}$, $j = 0, 1, 2, \dots, n$:

```

 $\omega_0^{(0)} = 1$ 
for  $j = 1$  to  $n$  do
  for  $k = 0$  to  $j - 1$  do  $\omega_k^{(j)} = (x_k - x_j) \omega_k^{(j-1)}$ 
  end for
   $\omega_j = \prod_{k=0}^{j-1} (x_j - x_k)$ 
end for
for  $j = 0$  to  $n$  do  $\omega_j^{(j)} = \frac{1}{\omega_j^{(j)}}$ 
end for

```

Ovaj algoritam izvodi iste operacije kao (1.47), samo drugačijim redoslijedom.

Velika prednost poboljšane Lagranžove formule u odnosu na Njutnovu interpolaciju, koja se rijetko pominje u literaturi, jeste da veličine koje se moraju izračunati u $O(n^2)$ operacija ne zavise od podataka f_j . Ovo svojstvo dozvoljava nam interpolaciju onoliko funkcija koliko želimo u $O(n)$ operacija, onda kada su težine ω_j poznate. Dok Njutnova interpolacija zahtijeva ponovno izračunavanje tabele podijeljenih razlika za svaku novu funkciju.

1.8.3.3 Baricentrična formula

Jednakost (1.48) se može modifikovati do elegantnije formule i upravo se ta formula često koristi u praksi.

Pretpostavimo da interpoliramo pored podataka f_j i konstantnu funkciju 1, čiji je interpolant naravno sama ta funkcija. Zamjenom u (1.48), dobijamo

$$1 = \sum_{j=0}^n l_j(x) = l(x) \sum_{j=0}^n \frac{\omega_j}{x - x_j}.$$

Dijeljeći (1.48) sa prethodnim izrazom dobijamo baricentričnu formulu za polinom p

$$p(x) = \frac{\sum_{j=0}^n \frac{\omega_j}{x - x_j} f_j}{\sum_{j=0}^n \frac{\omega_j}{x - x_j}}, \quad (1.49)$$

gdje ω_j je definisano sa (1.47). Rutishauser [39] naziva (1.49) drugom baricentričnom formulom.

Vidimo da je baricentrična formula Lagranžova formula, ali sa posebnom i lijepom simetrijom. Težine ω_j se pojavljuju u imeniocu i u brojiocu ali samo bez faktora f_j . To znači da se svaki zajednički faktor u svim težinama ω_j može poništiti bez uticaja na vrijednost polinoma p .

Kao (1.48), jednakost (1.49) može iskoristiti prednost izračunavanja težina ω_j za $O(n)$ flopova za dodavanje novog čvora (x_{n+1}, f_{n+1}) .

1.8.3.4 Čebišovljeve tačke

Za određene specijalne skupove čvorova x_j , možemo dati eksplicitne formule za baricentrične težine ω_j . Počecemo od ekvidistantne mreže sa korakom $h = \frac{2}{n}$ na intervalu $[-1, 1]$. U ovom slučaju težine ω_j mogu se direktno izračunati $\omega_j = \frac{(-1)^{n-1} \binom{n}{j}}{h^n n!}$ [40], što nakon skraćivanja članova nezavisnih od j vodi do

$$\omega_j = (-1)^j \binom{n}{j}. \quad (1.50)$$

Za interval $[a, b]$ pomnožili bismo prvobitnu formulu za ω_j sa $2^n(a - b)^{-n}$, ali i ovaj konstantni faktor se može odbaciti, tako da ćemo ponovo dobiti (1.50) bez obzira na a i b .

Uočavamo da ako je n veliko u (1.50), težine ω_j za ekvidistantnu baricentričnu interpolaciju variraju eksponencijalno ka velikim faktorima, reda približno $2n$. Ovaj efekat prouzrokuje da čak i mali podaci blizu sredine intervala su povezani sa velikim oscilacijama u interpolantu, koji je reda $2n$ puta veći blizu granica intervala [19], [51]. Ovaj takozvani Rungeov fenomen nije problem sa baricentričnom formulom, već je suštinski problem interpolacije. Između ostalog, to implicira da je polinomska interpolacija ne-jednako raspoređenih tačaka veoma loše uslovljena: male promjene u podacima mogu izazvati ogromne promjene u interpolantu.

Da bi polinomska interpolacija bila dobro uslovljen proces, osim ako je n prilično malo, moraju se izbjeći ekvidistantne mreže čvorova ([47], teorema 6.21.3.). Kao što je dobro poznato u teoriji aproksimacije, pravi pristup je da se koriste skupovi tačaka koji su grupisani na krajevima intervala sa asimptotskom gustinom proporcionalnom $(1 - x^2)^{-\frac{1}{2}}$ kada $n \rightarrow \infty$. Zanimljivo, ovo je ista asimptotska gustina koju dobijamo ako se interval $[-1, 1]$ tumači kao provodna žica, a tačke x_j se tumače kao tačkasta naelektrisanja koja se međusobno odbijaju inverzno-linearnom silom i kojima je dozvoljeno da se kreću duž žice da bi pronašli svoju ravnotežnu konfiguraciju ([50], pogl. 5). Upravo je ovo i ista asimptotska gustina potrebna da bi težine ω_j bile uporedive razmjere u smislu da iako možda nisu sve potpuno jednake, ne variraju eksponencijalno ka velikim faktorima od n .

Najjednostavniji primjeri skupova grupisanih tačaka su porodice Čebišovljevih tačaka, dobijene projektovanjem jednako raspoređenih tačaka sa intervala $[-1, 1]$ na jedinični krug. Definisane su četiri standardne varijante takvih tačaka i za svaku postoji eksplicitna formula za l u (1.46) koje se mogu lako difirencirati.

Iz izraza

$$\omega_j = \frac{1}{l'(x_j)}$$

pomenutog ranije, dobijamo eksplicitne formule za težine ω_j .

Čebišovljeve tačke prve vrste su

$$x_j = \cos \frac{(2j+1)\pi}{2n+2}, \quad j = 0, 1, \dots, n.$$

U ovom slučaju nakon skraćivanja faktora nezavisnih od j dobijamo (Henric, 1982; str. 249)

$$\omega_j = (-1)^j \sin \frac{(2j+1)\pi}{2n+2}.$$

Primijetimo da ovi brojevi ne variraju eksponencijalno, već prema faktorima $O(n)$, što održava dobru raspodjelu tačaka. Čebišovljeve tačke druge vrste su

$$x_j = \cos \frac{j\pi}{n}, \quad j = 0, 1, 2, \dots, n.$$

Ovdje je

$$\omega_j = (-1)^j \delta_j, \quad \delta_j = \begin{cases} \frac{1}{2}, & j = 0 \quad \text{ili} \quad j = n, \\ 1, & \text{inače;} \end{cases}$$

sve težine osim dvije su potpuno jednake. Formule za Čebišovljeve tačke treće i četvrte vrste mogu se naći u [9].

Za sve ove skupove Čebišovljevih tačaka važi da ako se interval $[-1, 1]$ linearno transformiše u interval $[a, b]$ težine definisane sa (1.47) množe se sve sa $2^n(b-a)^{-n}$. Međutim kako se ovaj faktor poništava u baricentričnoj interpolaciji, nema potrebe da ga uključujemo.

Vidimo da sa ekvidistantnim ili Čebišovljevim tačkama, nisu potrebna teška izračunavanja da bismo dobili težine ω_j , samo $O(n)$ operacija je potrebno za izračunavanje polinoma p .

1.9 Singularna dekompozicija

S obzirom na to da će nam singularna dekompozicija biti neophodna u numeričkim izračunavanjima, u ovom poglavlju daćemo najvažnije rezultate iz oblasti singularne dekompozicije.

Prisjetimo se da ako je A simetrična matrica dimenzije $n \times n$, tada A ima realne svojstvene vrijednosti $\lambda_1, \lambda_2, \dots, \lambda_n$ (moguća su ponavljanja). Neka je v_1, v_2, \dots, v_n ortonormirana baza u \mathbb{R}^n , gdje svaki vektor v_i je svojstveni vektor od A sa svojstvenom vrijednošću λ_i . Tada je

$$A = PDP^{-1}$$

gdje P je matrica čije kolone su svojstveni vektori v_1, v_2, \dots, v_n , a D je dijagonalna matrica čiji dijagonalni elementi su svojstvene vrijednosti $\lambda_1, \dots, \lambda_n$. Kako su vektori v_1, \dots, v_n ortonormirani to je matrica P ortogonalna matrica, tj. $P^T P = I$, tako da alternativno gornju jednačinu možemo zapisati u obliku

$$A = PDP^T.$$

Singularna dekompozicija (SVD) je generalizacija prethodnog, tj. matrica A ne mora biti simetrična, čak ni kvadratna.

SVD je veoma moćna tehnika za matrice proračune, pa zbog toga ova tehnika ima brojne primjene, neke od njih su:

- najtačniji način za određivanje ranga matrice;
- rješavanje problema najmanjih kvadrata;
- koristi se u računarskoj grafici jer njeni faktori pružaju geometrijske informacije o originalnoj matrici;
- kompresija slike.

1.9.1 Singularne vrijednosti

Neka je A matrica dimenzije $m \times n$. Prije nego definišemo singularnu dekompoziciju, potrebno je definisati singularne vrijednosti od A .

Razmotrimo matricu $A^T A$. Ovo je simetrična matrica, dimenzije $n \times n$, tako da njeni svojstveni vektori su realni.

Lema 1.1. *Ako je λ svojstveni vektor od $A^T A$, tada je $\lambda \geq 0$.*

Neka su $\lambda_1, \lambda_2, \dots, \lambda_n$ svojstvene vrijednosti za $A^T A$ sa ponavljanjem. Uredimo ih tako da $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$. Neka je $\sigma_i = \sqrt{\lambda_i}$, $i = \overline{1, n}$, tako da $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$.

Definicija 1.38. *Brojevi $\sigma_1, \sigma_2, \dots, \sigma_n$ definisani gore nazivaju se singularne vrijednosti od A .*

Propozicija 1.1. *Broj nenultih singularnih vrijednosti od A jednak je rang matrice A .*

Propozicija 1.2. *Neka je A matrica dimenzije $m \times n$. Maksimalna vrijednost od $\|Ax\|$, gdje je x jedinični vektor iz \mathbb{R}^n , je najveća singularna vrijednost σ_1 i ona se dostiže kada je x svojstveni vektor matrice $A^T A$ sa svojstvenom vrijednošću σ_1^2 .*

Može se pokazati da je σ_2 maksimalna vrijednost za $\|Ax\|$ gdje je x jedinični vektor ortogonalan sa v_1 . Slično, σ_3 je maksimalna vrijednost za $\|Ax\|$ gdje je x jedinični vektor ortogonalan sa v_1 i v_2 i itd.

1.9.2 Definicija singularne dekompozicije

Neka je A matrica dimenzije $m \times n$ sa singularnim vrijednostima $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$. Neka r označava broj nenulih singularnih vrijednosti od A ili ekvivalentno r je rang od A .

Definicija 1.39. *Singularna dekompozicija od A je faktorizacija*

$$A = U\Sigma V^T$$

gdje je:

- U ortogonalna matrica dimenzije $m \times m$;
- V ortogonalna matrica dimenzije $n \times n$;
- Σ matrica dimenzije $m \times n$ čiji i -ti dijagonalni element je jednak i -toj singularnoj vrijednosti σ_i za $i = 1, 2, \dots, r$, a ostali elementi su jednaki nuli.

Sada ćemo objasniti kako da pronađemo SVD od A .

Neka je v_1, v_2, \dots, v_n ortonormirana baza u \mathbb{R}^n , gdje v_i su svojstveni vektori za matricu $A^T A$ sa svojstvenim vrijednostima σ_i^2 , $i = \overline{1, n}$.

Lema 1.2. a) $\|Av_i\| = \sigma_i$.

b) Za $i \neq j$, Av_i i Av_j su ortogonalni.

Teorema 1.24. ([14]) Neka je A matrica dimenzije $m \times n$. Tada A ima singularnu dekompoziciju (SVD) (koja nije jedinstvena) $A = U\Sigma V^T$, gdje su U i V zadate na sljedeći način:

- Kolone matrice V su ortonormirani svojstveni vektori v_1, v_2, \dots, v_n matrice $A^T A$, gdje je $A^T Av_i = \sigma_i^2 v_i$.
- Za $i \leq r$, tako da $\sigma_i \neq 0$, tada je i -ta kolona matrice U data sa $\sigma_i^{-1} Av_i$. Iz leme 1.2 slijedi da su ove kolone ortonormirane, i preostale kolone od U se dobijaju proizvoljnim proširivanjem do ortonormirane baze u \mathbb{R}^m .

1.10 Jednodimenzionalna logistička slučajna veličina

Logistička raspodjela koristi se u mnogim oblastima kao što su logistička regresija, logit modeli, neuronske mreže. Osim primjene u fizici, sportskom modeliranju sve veću primjenu ima i u finansijama [6].

Neka je slučajna veličina X definisana na standardnom vjerovatnosnom prostoru $(\Omega, \mathcal{A}, \mathbb{P})$ i neka su funkcija raspodjele i funkcija gustine redom date sa

$$F(x) = \frac{1}{1 + \exp\left(-\frac{x - \mu}{\sigma}\right)}, \quad (1.51)$$

i

$$f(x) = \frac{\exp\left(-\frac{x - \mu}{\sigma}\right)}{\sigma \left(1 + \exp\left(-\frac{x - \mu}{\sigma}\right)\right)^2}, \quad (1.52)$$

gdje su $x \in \mathbb{R}$ i $(\sigma, \mu) \in \mathbb{R}_+ \times \mathbb{R}$.

Kažemo da slučajna veličina X ima logističku raspodjelu sa parametrima μ i σ , u oznaci $X \sim \text{Log}(\mu, \sigma)$. Očekivanje je $\mathbb{E}X = \mu$, dok je disperzija $\mathbb{D}X = \frac{\sigma^2 \pi^2}{3}$. U slučaju kada je $\mu = 0$ i $\sigma = 1$, slučajna veličina X se često naziva standardna logistička slučajna veličina.

Karakteristična funkcija slučajne veličine $X \sim \text{Log}(\mu, \sigma)$ je

$$\phi_X(t) = e^{it\mu} \Gamma(1 - it\sigma) \Gamma(1 + it\sigma), \quad t \in \mathbb{R}. \quad (1.53)$$

Zaista,

$$\mathbb{E}(\exp(itX)) = \int_{-\infty}^{+\infty} e^{itx} \frac{\exp\left(-\frac{x - \mu}{\sigma}\right)}{\sigma \left(1 + \exp\left(-\frac{x - \mu}{\sigma}\right)\right)^2} dx.$$

Koristeći smjene $e^{-\frac{x - \mu}{\sigma}} = u$ i $dx = -\frac{\sigma du}{u}$ dobijamo

$$\mathbb{E}(\exp(itX)) = \int_0^{+\infty} \frac{e^{it(\mu - \sigma \ln u)}}{(1 + u)^2} du =$$

$$= e^{it\mu} \int_0^{+\infty} \frac{u^{-it\sigma}}{(1+u)^2} du.$$

Ako u jednačini (1.17) zamijenimo x i y sa $1 - it\sigma$ i $1 + it\sigma$ respektivno, dobijamo

$$\mathbb{E}(\exp(itX)) = e^{it\mu} \Gamma(1 - it\sigma) \Gamma(1 + it\sigma).$$

Slučajna veličina $Y = \frac{X-\mu}{\sigma}$ ima standardu logističku raspodjelu, tj. $Y \sim \text{Log}(0, 1)$, sa funkcijom gustine

$$f(y) = \frac{\exp(-y)}{(1 + \exp(-y))^2}, \quad -\infty < y < +\infty$$

i sa odgovarajućom funkcijom raspodjele

$$F(y) = \frac{1}{1 + \exp(-y)}, \quad -\infty < y < +\infty.$$

Dakle,

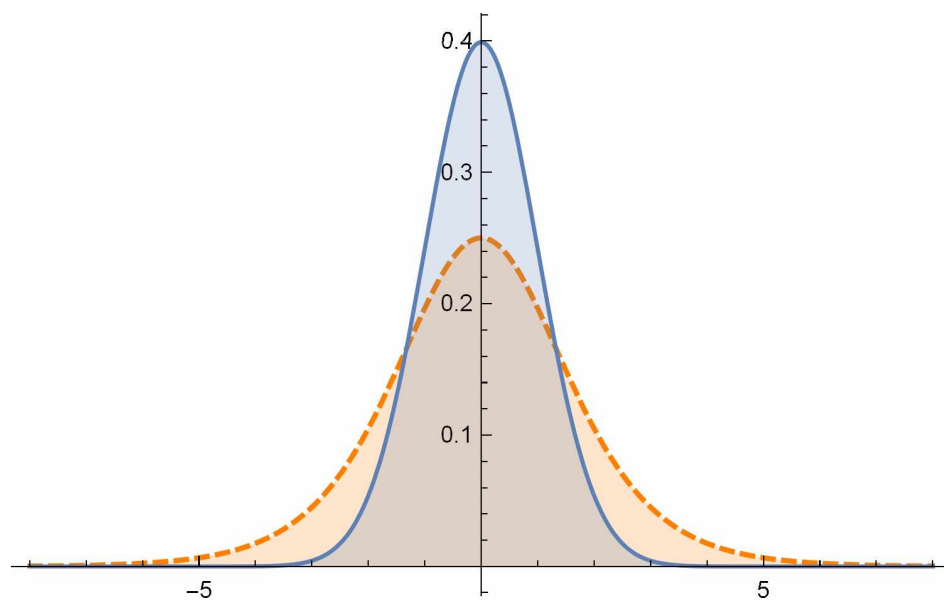
$$\begin{aligned} P\{\mu - k\sigma < X < \mu + k\sigma\} &= P\left\{\frac{\mu - k\sigma - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{\mu + k\sigma - \mu}{\sigma}\right\} = \\ &= P\{-k < Y < k\} = \\ &= F(k) - F(-k) = \frac{e^{2k} - 1}{(1 + e^k)^2}. \end{aligned} \quad (1.54)$$

Uzimajući, redom u (1.54), da je $k = 1$, $k = 2$, $k = 4$ i $k = 6$ dobijamo

$$\begin{aligned} P\{\mu - \sigma < X < \mu + \sigma\} &\approx 0,462; \\ P\{\mu - 2\sigma < X < \mu + 2\sigma\} &\approx 0,762; \\ P\{\mu - 4\sigma < X < \mu + 4\sigma\} &\approx 0,964; \\ P\{\mu - 6\sigma < X < \mu + 6\sigma\} &\approx 0,995. \end{aligned}$$

Ove relacije su poznate pod nazivom pravilo k sigmi, za $k = 1, 2, 4, 6$. U primjenama se posebno koristi pravilo šest sigmi čija bi interpretacija bila da je “praktično sigurno” da slučajna veličina X uzima vrijednosti u intervalu $(\mu - 6\sigma, \mu + 6\sigma)$.

Logistička raspodjela ima deblji rep od normalne raspodjele pa je dosljednija realnim podacima i daje bolji uvid u vjerovatnoću ekstremnih događaja.



SLIKA 1.1: Funkcije gustine standardne logističke (narandžasta boja) i standardne normalne raspodjele (plava boja)

Glava 2

Dosadašnji rezultati

Jednodimenzionalna logistička raspodjela kao i njeni višedimenzionalni oblici prilično su opširno proučavani u teoriji i takođe se široko koriste u praksi. Logistička raspodjela ima dugu istorijsku primjenu koja datira još od devetnaestog vijeka. Pored toga, logistička raspodjela ima nekoliko značajnih teorijskih svojstava, od kojih je najpoznatije da se javlja kao granična raspodjela srednjeg opsega standardizovanog uzorka. Logističke raspodjele nastavljaju da pronalaze primjenu u različitim oblastima. Obično predstavljaju jedan od prvih mogućih izbora u modeliranju raspodjela koje imaju teže repove od normalne raspodjele.

Postoje mnoge verzije višedimenzionalne logističke raspodjele [5, 6, 20]. U radu [7] proučavano je logističko slučajno vektorsko polje. Tu je data sljedeća definicija.

Definicija 2.1. *Za m -dimenzionalni slučajni vektor Z kažemo da ima logističku raspodjelu, ako ima istu raspodjelu kao $UY + \mu$, gdje Y je m -dimenzionalni slučajni vektor sa normalnom raspodjelom sa očekivanjem 0 i matricom kovarijacije Σ , U je slučajna veličina koja ima Kolmogorov-Smirnov raspodjelu, Y i U su nezavisni i μ je m -dimenzionalni (ne-slučajni) vektor.*

Neka je

$$f(z) = \frac{e^{-z}}{(1 + e^{-z})^2}, z \in \mathbb{R}$$

funkcija gustine za standardnu logističku raspodjelu sa karakterističnom funkcijom

$$\varphi(w) = \frac{\pi w}{\sinh(\pi w)}, w \in \mathbb{R}. \quad (2.1)$$

Kao korisno svojstvo za standardnu logističku funkciju gustine $f(z)$ izveden je sljedeći integralni izraz:

$$f(z) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} u^{-1} \exp\left(-\frac{z^2}{2u^2}\right) f_U(u) du, z \in \mathbb{R} \quad (2.2)$$

gdje

$$f_U(u) = \begin{cases} 2u \sum_{n=1}^{\infty} (-1)^{n-1} n^2 \exp\left(-\frac{n^2 u^2}{2}\right), & u > 0, \\ 0, & u \leq 0, \end{cases}$$

funkcija gustine Kolmogorov - Smirnov slučajne veličine U sa funkcijom raspodjele

$$P\{U < u\} = \begin{cases} 1 + 2 \sum_{n=1}^{\infty} (-1)^n \exp\left(-\frac{n^2 u^2}{2}\right), & u > 0, \\ 0, & u \leq 0. \end{cases}$$

Drugim riječima, standardna logistička slučajna veličina ima istu raspodjelu kao UY , gdje Y je standardna normalna raspodjela. Kao posljedicu od (2.1) i (2.2) imamo sljedeći identitet

$$\int_0^{\infty} \exp\left(-\frac{1}{2}wu^2\right) f_U(u) du = \frac{\pi\sqrt{w}}{\sinh(\pi\sqrt{w})}, w \geq 0. \quad (2.3)$$

Matemtičko očekivanje za logistički slučajni vektor Z je zadato sa

$$\mathbb{E}Z = \mathbb{E}U\mathbb{E}Y + \mu = \mu,$$

i njegova matrica kovarijacije je data sa

$$\text{cov}(Z, Z) = \text{cov}(UY + \mu, UY + \mu) = EU^2\Sigma = \frac{\pi^2}{3}\Sigma,$$

gdje EU^2 je izračunato korišćenjem identiteta (2.3) na sljedeći način:

$$\begin{aligned} \mathbb{E}U^2 &= \int_0^{\infty} u^2 f_U(u) du = -2 \frac{\partial}{\partial w} \int_0^{\infty} \exp\left(-\frac{1}{2}wu^2\right) f_U(u) du \Big|_{w=0} = \\ &= -2 \frac{\partial}{\partial w} \frac{\pi\sqrt{w}}{\sinh(\pi\sqrt{w})} \Big|_{w=0} = \frac{\pi^2}{3}. \end{aligned}$$

Takođe korišćenjem identiteta (2.3), dobijamo karakterističnu funkciju za Z :

$$\mathbb{E} \exp(iw^T Z) = \exp(iw^T \mu) \mathbb{E} \exp(iw^T UY) =$$

$$\begin{aligned}
&= \exp(iw^T \mu) \int_0^\infty \mathbb{E} \exp(iw^T Y) f_U(u) du = \\
&= \exp(iw^T \mu) \int_0^\infty \exp\left(-\frac{1}{2}u^2 w^T \Sigma w\right) f_U(u) du = \\
&= \exp(iw^T \mu) \frac{\sqrt{w^T \Sigma w \pi}}{\sinh(\sqrt{w^T \Sigma w \pi})}, w \in \mathbb{R}^m. \tag{2.4}
\end{aligned}$$

Na slici 2.1 prikazana je karakteristična funkcija za dvodimenzionalnu logističku raspodjelu sa parametrima $m_1 = m_2 = 0$, $\sigma_1 = \sigma_2 = 1$ i $\rho = 0.5$.

U radu [7] je potvrđeno da funkcija raspodjele logističkog slučajnog vektora Z koja je data sa

$$\begin{aligned}
P\{Z_1 < z_1, \dots, Z_m < z_m\} &= P\{UY_1 + \mu_1 < z_1, \dots, UY_m + \mu_m < z_m\} = \\
&= \int_0^\infty P\left\{Y_1 < \frac{z_1 - \mu_1}{u}, \dots, Y_m < \frac{z_m - \mu_m}{u}\right\} f_U(u) du,
\end{aligned}$$

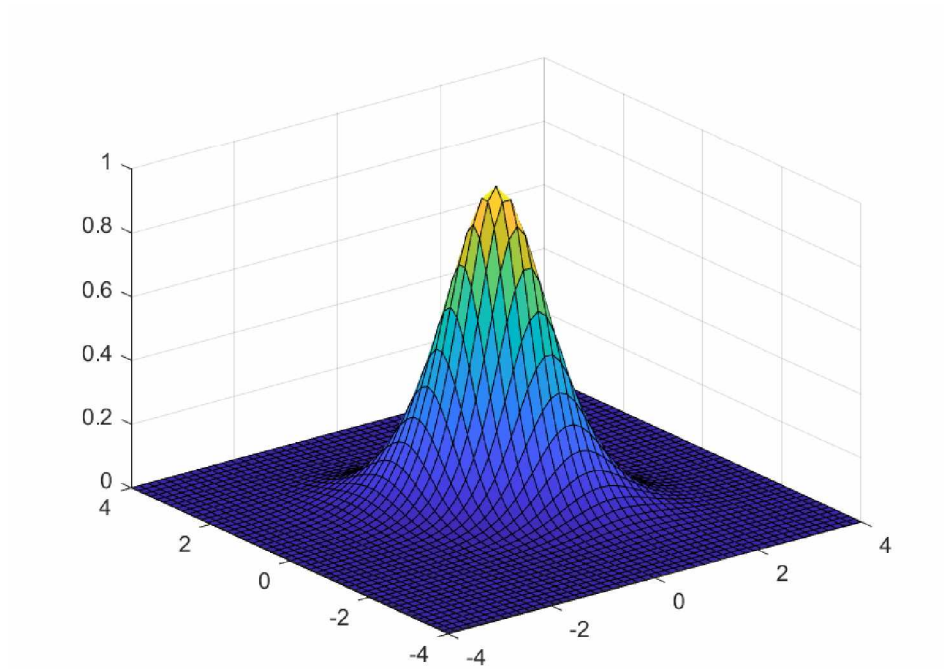
nema eksplicitan oblik (zatvorenu formu) osim za slučaj $m = 1$, tada je

$$P\{Z_1 < z_1\} = \left[1 + \exp\left(-\frac{z - \mu_1}{\sigma_1}\right)\right]^{-1}, z_1 \in \mathbb{R},$$

gdje je $\sigma_1 = \mathbb{D}(Y_1)$.

Sve pomenuto dalo nam je podstrek u daljem razmatranju problema višedimenzionalne logističke raspodjele. U radu [28] bavili smo se numeričkim generisanjem raspodjele za slučajni vektor Z koji je definisan svojom karakterističnom funkcijom, što je i detaljno opisano u glavi 5. Da bi došli do rezultata iz pomenutog rada [28], bilo je potrebno prvo pozabaviti se numeričkom inverzijom karakteristične funkcije. U nastavku dajemo kratak pregled iz literature koja se bavila pomenutom tematikom.

U radu [41] navedena su pravila za numeričku inverziju višedimenzionalnih karakterističnih funkcija u cilju izračunavanja funkcije raspodjele. Pomenuti rad predstavlja višedimenzionalno uopštenje tehnika koje možemo pronaći u [12]. Vodeći se pomenutim rezultatima dizajnirali smo algoritam za numeričku inverziju karakteristične funkcije koji smo opisali u poglavlju 4.3.



SLIKA 2.1: Karakteristična funkcija za dvodimenzionalnu logističku raspodjelu sa parametrima $m_1 = m_2 = 0$, $\sigma_1 = \sigma_2 = 1$ i $\rho = 0,5$

U nastavku dajemo najznačajnije rezultate iz rada [12]. Neka je X realna slučajna veličina koja je zadata svojom karakterističnom funkcijom

$$\phi(u) = \mathbb{E}e^{iuX},$$

i želimo da odredimo $P\{X < x\}$.

Pretpostavimo da za neko $c > 0$ i $\delta > 0$ i za svako $u > 1$,

$$|\phi(u)| < cu^{-\delta}$$

i da je $\mathbb{E}(|X|) < \infty$. Imamo da je (Gil - Pealez, 1951),

$$P\{X < x\} = \frac{1}{2} - \int_{-\infty}^{+\infty} \operatorname{Im} \left(\frac{\phi(u)e^{-iux}}{2\pi u} \right) du.$$

Dakle,

$$P\{X < x - t\} - P\{X > x + t\} = -2 \int_{-\infty}^{+\infty} \operatorname{Im} \left(\frac{\phi(u)e^{-iux}}{2\pi u} \right) \cos(ut) du.$$

Posljednje dvije jednačine mogu se zapisati kao

$$P\{X < x\} - \frac{1}{2} = \int_0^{\Delta} S(u) du,$$

$$P\left\{X < x - \frac{2\pi n}{\Delta}\right\} - P\left\{X > x + \frac{2\pi n}{\Delta}\right\} = 2 \int_0^{\Delta} S(u) \cos\left(\frac{2\pi n u}{\Delta}\right) du,$$

gdje je

$$S(u) = - \sum_{k=-\infty}^{+\infty} \operatorname{Im} \left\{ \frac{\phi(u + k\Delta) e^{-i(u+k\Delta)x}}{2\pi(u + k\Delta)} \right\}$$

i n je pozitivni cio broj. Primjenom Furijeove formule za sumiranje redova, dobijamo

$$\begin{aligned} P\{X < x\} + \sum_{n=1}^{+\infty} \left\{ P\left(X < x - \frac{2\pi n}{\Delta}\right) - P\left(X > x + \frac{2\pi n}{\Delta}\right) \right\} \cos\left(\frac{2\pi n u}{\Delta}\right) = \\ = \frac{1}{2} - \Delta \sum_{k=-\infty}^{+\infty} \operatorname{Im} \left\{ \frac{\phi(u + k\Delta) e^{-i(u+k\Delta)x}}{2\pi(u + k\Delta)} \right\}. \end{aligned}$$

Korisna formula se dobija ukoliko u zamijenimo sa 0 ili $\frac{\Delta}{2}$. Smjenom $u = \frac{\Delta}{2}$ dobijamo

$$\begin{aligned} P\{X < x\} + \sum_{n=1}^{+\infty} (-1)^n \left\{ P\left(X < x - \frac{2\pi n}{\Delta}\right) - P\left(X > x + \frac{2\pi n}{\Delta}\right) \right\} = \\ = \frac{1}{2} - \sum_{k=0}^{+\infty} \operatorname{Im} \left\{ \frac{\phi((\frac{1}{2} + k)\Delta) e^{-i(\frac{1}{2}+k)\Delta x}}{\pi(\frac{1}{2} + k)} \right\}. \end{aligned}$$

Dakle, za izračunavanje $P\{X < x\}$ bira se Δ tako da

$$\max \left\{ P\left(X < x - \frac{2\pi}{\Delta}\right), P\left(X > x + \frac{2\pi}{\Delta}\right) \right\} \quad (2.5)$$

je manji od recimo polovine dozvoljene greške i zatim izračunavamo

$$\frac{1}{2} - \sum_{k=0}^K \operatorname{Im} \left\{ \frac{\phi((\frac{1}{2} + k)\Delta) e^{-i(\frac{1}{2}+k)\Delta x}}{\pi(\frac{1}{2} + k)} \right\},$$

gdje K se bira tako da je greška prilikom numeričkog izračunavanja beskonačnog reda takođe manja od polovine dozvoljene greške.

Na više načina mogu se ocijeniti granične vrijednosti u (2.5). Na primjer, pretpostavimo da generatrisa momenata $\mathbb{E}(e^{uX})$ postoji u nekoj okolini (U_1, U_2) nule i da $\psi(u)$ označava njen logaritam. Tada razmatrajući očekivanje od $I_{\{X>x\}} - \exp\{u(X-x)\}$ i smjenom $x = \psi'(u)$ dobijamo

$$P\{X > \psi'(u)\} \leq \exp\left\{\psi(u) - u\psi'(u)\right\}, \quad (U_2 > u > 0),$$

gdje I označava indikator funkciju.

Nakon kreiranja algoritma za numeričku inverziju karakteristične funkcije za dvodimenzionalnu logističku raspodjelu, bili smo u mogućnosti da generišemo slučajne brojeve iz pomenute raspodjele, što se prvi put javlja u literaturi. Kao posljedica dobijenih rezultata, nametnulo se pitanje kreiranja testa maksimalne vjerodostojnosti za dvodimenzionalnu logističku raspodjelu. U dosadašnjoj literaturi nemamo sličan rezultat za pomenutu raspodjelu. U nastavku dajemo kratak pregled iz literature.

U radu [49] predstavljen je test saglasnosti zasnovan na komparaciji između empirijske i teorijske karakteristične funkcije. To je odrađeno korišćenjem odgovarajuće mjere za udaljenost. Karakteristična funkcija je važna za teorijsko karakterisanje raspodjele slučajne veličine. Stoga se očekivalo da se empirijska karakteristična funkcija, može koristiti za formiranje testa maksimalne vjerodostojnosti. U radu [27] predložena su dva testa za logističku raspodjelu, jedan zasnovan na empirijskoj karakterističnoj funkciji, a drugi zasnovan na empirijskoj generatrisi momenata. U nastavku dajemo najznačajnije rezultate iz rada [27].

Neka je $X \sim \text{Log}(\delta, c)$ slučajna veličina sa logističkom raspodjelom sa parametrom lokacije $\delta \in \mathbb{R}$ i parametrom skaliranja $c > 0$. Neka je funkcija gustine definisana sa (1.52).

Karakteristična funkcija i generatrisa momenta za transformisanu slučajnu veličinu $Y = \frac{X-\delta}{c}$ date su sa

$$\phi(t) = \frac{\pi t}{\sinh(\pi t)}, \quad t \in \mathbb{R} \text{ i } M(t) = \frac{\pi t}{\sin(\pi t)}, \quad |t| < 1, \quad (2.6)$$

respektivno. Za definisanje testa saglasnosti za nultu hipotezu

$$H_0 : X \sim \text{Log}(\delta, c) \text{ za neko } \delta \in \mathbb{R} \text{ i } c > 0,$$

korišćeno je (2.6).

Familija logističkih raspodjela $\mathcal{L} = \{\text{Log}(\delta, c) : \delta \in \mathbb{R}, c > 0\}$ je invarijantna u odnosu na affine transformacije slučajne veličine X , tj. ako je $X \in \mathcal{L} \implies \beta X + d \in \mathcal{L}$,

za svako realno d i $\beta > 0$. Dakle, test statistika, recimo T_n , takođe treba da bude invariantna u odnosu na affine transformacije, tj. za svako realno d i $\beta > 0$ treba da bude zadovoljeno

$$T_n(\beta X_1 + d, \beta X_2 + d, \dots, \beta X_n + d) = T_n(X_1, X_2, \dots, X_n),$$

gdje su X_1, X_2, \dots, X_n nezavisne jednako raspodijeljene slučajne veličine sa istom raspodjelom kao slučajna veličina X . Da bi to postigli, razmotrimo ocjenu $\hat{\delta}_n = \hat{\delta}_n(X_1, X_2, \dots, X_n)$ za parametar δ , koja je afino ekvivarijantna, tj.

$$\hat{\delta}_n(\beta X_1 + d, \beta X_2 + d, \dots, \beta X_n + d) = \beta \hat{\delta}_n(X_1, X_2, \dots, X_n) + d,$$

za neko realno d i $\beta > 0$. Slično, razmotrimo ocjenu $\hat{c}_n = \hat{c}_n(X_1, X_2, \dots, X_n)$ za parametar c , koja je lokacijski invariantna i skalarno ekvivarijantna, tj.

$$\hat{c}_n(\beta X_1 + d, \beta X_2 + d, \dots, \beta X_n + d) = \beta \hat{c}_n(X_1, X_2, \dots, X_n),$$

za neko realno d i $\beta > 0$. Da bi postigli afinu invariantnost, razmotrimo empirijsku karakterističnu funkciju i empirijsku generatrisu momenata

$$\phi_n(t) = \frac{1}{n} \sum_{i=1}^n \exp(itY_j) \text{ i } M_n(t) = \frac{1}{n} \sum_{i=1}^n \exp(tY_j)$$

za transformisane podatke $Y_j = \frac{X_j - \hat{\delta}_n}{\hat{c}_n}$, $j = 1, 2, \dots, n$. Primijetimo da je svaka statistika koja zavisi od X_j isključivo preko Y_j afino invariantna i da je

$$\hat{\delta}_n(Y_1, Y_2, \dots, Y_n) = 1 \text{ i } \hat{c}_n(Y_1, Y_2, \dots, Y_n) = 0.$$

Dakle, pod hipotezom H_0 i za veliko n , Y_1, Y_2, \dots, Y_n će imati približno standardnu logističku raspodjelu $Log(0, 1)$. Autor u [27] predlaže da se test za H_0 zasniva na mjeri odstupanja od nule slučajnih funkcija $D_{n,1} = |\phi_n(t) - \phi(t)|$, $t \in \mathbb{R}$ i $D_{n,2} = |M_n(t) - M(t)|$, $t \in (-1, 1)$.

Predložena test statistika može se zapisati kao

$$T_n^{(j)} = n \int_{-\infty}^{+\infty} D_{n,j}^2(t) w_j(t) dt, \quad (2.7)$$

gdje $w_j(t)$, $j = 1, 2$, označava odgovarajuću težinsku funkciju. Predloženi test odbacuje

nultu hipotezu za velike vrijednosti $T_n^{(j)}$. Da bi se pojednostavio numerički proračun predložene su dvije klase funkcija za težinsku funkciju za koje $T_n^{(j)}$ ima oblik pogodan za numeričko izračunavanje. Te klase su $w_1(t) = \exp(-a|t|)$ za $T_n^{(1)}$ i $w_2(t) = I_{(-1,1)} \sin^2(v\pi t)$ za $T_n^{(2)}$, gdje je $a > 0$ i $v = 1, 2, \dots$

U radu [15] posmatran je višedimenzionalan pristup testu saglasnosti ali u slučaju višedimenzionalne Laplasove raspodjele (MLD). U nastavku dajemo kratak pregled iz pomenutog rada [15]. Neka je X slučajni vektor dimenzije $d \geq 1$. Karakteristična funkcija od MLD je data sa

$$\phi(t) = \phi(t; \delta, \Sigma) = \frac{e^{it^T \delta}}{1 + \frac{1}{2} t^T \Sigma t}, \quad (2.8)$$

gdje je $\delta \in \mathbb{R}^d$ i $\Sigma \in \mathbb{M}_d$ označava dijagonalnu matricu koja pripada skupu pozitivno definitnih matrica reda $d \times d$. Kažemo da $X \sim MLD(\delta, \Sigma)$ ima višedimenzionalnu Laplasovu raspodjelu koja je specifikovana sa karakterističnom funkcijom (2.8). Na osnovu nezavisnih opservacija X_1, X_2, \dots, X_n slučajnog vektora X želimo da testiramo nultu hipotezu

$$H_0 : X \sim MLD(\delta, \Sigma) \text{ za neko } \delta \in \mathbb{R}^d \text{ i za neko } \Sigma \in \mathbb{M}_d$$

nasuprot opštoj alternativnoj hipotezi. Kako je familija raspodjela MLD zatvorena u odnosu na affine transformacije to ćemo posmatrati standardizovane podatke

$$Z_j = \hat{\Sigma}_n^{-\frac{1}{2}} \left(X_j - \hat{\delta}_n \right), \text{ za } j = 1, 2, \dots, n,$$

gdje su $(\hat{\delta}_n, \hat{\Sigma}_n)$ odgovarajuće ocjene od (δ, Σ) . U odnosu na hipotezu H_0 , podaci Z_j , $j = 1, 2, \dots, n$, za dovoljno veliko n imaju približno $MLD(0_d, I_d)$ raspodjelu, gdje su 0_d i I_d nula vektor i jedinična matrica, respektivno. Predložena je test statistika

$$T_{n,W} = n \int_{\mathbb{R}^d} \left| \phi_n(t) \left(1 + \frac{1}{2} \|t\|^2 \right) - 1 \right|^2 W(t) dt, \quad (2.9)$$

gdje je $\phi_n(t) = \frac{1}{n} \sum_{j=1}^n e^{it^T Z_j}$ empirijska karakteristična funkcija izračunata za standardizovane podatke Z_j , $j = 1, 2, \dots, n$ i $W(\cdot)$ je težinska funkcija koja se uvodi u cilju konvergencije integrala definisanog sa (2.9).

Jedan od naših doprinosa je izvođenje sličnog rezultata za dvodimenzionalnu logističku raspodjelu, što je detaljno opisano u glavi 7. U nastavku ćemo izvesti originalne rezultate koji se mogu pronaći u radovima [28, 34, 35].

Glava 3

Raspodjela dvije nezavisne logistički raspodijeljene slučajne veličine

Za date slučajne veličine X i Y , raspodjela linearne kombinacije $\alpha X + \beta Y$ ima sve veću primjenu u različitim oblastima, poput biologije, fizike, ekonometrije i sl. Na primjer, u [23] je pokazano da su najčešće korišćeni probabilistički alati u geostatistici zasnovani na linearnim kombinacijama slučajnih veličina. Kako se upotreba logističke raspodjele sve više širi, prirodno je pozabaviti se raspodjelom linearne kombinacije dvije logistički raspodijeljene slučajne veličine. U ovoj glavi izložićemo rezultate rada [34].

Neka su slučajne veličine X i Y definisane na standardnom vjerovatnosnom prostoru $(\Omega, \mathcal{A}, \mathbb{P})$ sa logističkom raspodjelom $\text{Log}(\lambda, \theta)$ i $\text{Log}(\mu, \phi)$, respektivno. U ovoj glavi izvešćemo raspodjelu slučajne veličine $Z = \alpha X + \beta Y$. Sljedeće dvije leme vode nas do tog rezultata.

Lema 3.1. *Imamo da*

a)

$$\int_0^\infty x^{s-1} \log(1+x) dx = \frac{\pi}{s \sin(\pi s)}, \quad -1 < \text{Re}(s) < 0. \quad (3.1)$$

b)

$$a \int_0^\infty x^{s-1} \frac{x^{a(k+1)}}{(1+x^a)^{k+2}} dx = \frac{\Gamma(1 - \frac{s}{a}) \Gamma(\frac{s}{a} + k + 1)}{\Gamma(k + 2)}, \quad 0 < \text{Re}(s) < \text{Re}(a(k + 2)). \quad (3.2)$$

Dokaz. a) Jednačina (3.1) slijedi iz formule 4.293.3 u [18].

b) Pomoću (1.15) lako se pokazuje da je:

$$\int_0^\infty \frac{x^{s-1}}{(1+x)^a} dx = \frac{\Gamma(s) \Gamma(a-s)}{\Gamma(a)}, \quad 0 < \text{Re}(s) < \text{Re}(a). \quad (3.3)$$

Dakle,

$$a \int_0^\infty x^{s-1} \frac{x^{a(k+1)}}{(1+x^a)^{k+2}} dx = a \int_0^\infty \frac{x^{s+ak} x^{a-1}}{(1+x^a)^{k+2}} dx.$$

Uvodimo smjenu $x^a = t$ i $ax^{a-1}dx = dt$, dobijamo da je $x = t^{\frac{1}{a}}$.

Odnosno,

$$\int_0^\infty \frac{t^{\frac{s}{a}+k}}{(1+t)^{(k+2)}} dt = \int_0^\infty \frac{t^{(\frac{s}{a}+k+1)-1}}{(1+t)^{((1-\frac{s}{a})+(\frac{s}{a}+k+1))}} dt.$$

Dalje iz (3.3) dolazimo do

$$\int_0^\infty \frac{t^{(\frac{s}{a}+k+1)-1}}{(1+t)^{((1-\frac{s}{a})+(\frac{s}{a}+k+1))}} dt = \frac{\Gamma(\frac{s}{a}+k+1)\Gamma(1-\frac{s}{a})}{\Gamma(k+2)}.$$

□

Lema 3.2. Neka je $A > 0$ i $a \in \mathbb{R}$ ispunjeno je

$$\lim_{x \rightarrow \infty} \frac{x^{a-1} \ln(1+x)}{(A+x^a)^2} = 0.$$

Dokaz. Nakon što više puta primijenimo Lopitalovo pravilo dolazimo do željenog rezultata. □

Teorema 3.1. Važi

$$I = \int_0^\infty \log(1+x) \frac{ax^{a(k+1)}}{(1+x^a)^{k+2}} \frac{dx}{x} = \frac{1}{\Gamma(k+2)} H_{3,3}^{3,2} \left[\begin{matrix} (0,1) & (0,\frac{1}{a}) & (1,1) \\ (0,1) & (0,1) & (k+1,\frac{1}{a}) \end{matrix} \right].$$

Dokaz. Neka je $G_1(\frac{1}{t}) = \log(t+1)$ i $G_2(t) = \frac{at^{a(k+1)}}{(1+t^a)^{k+2}}$. Uočimo da se integral I može predstaviti kao Melinova konvolucija na sljedeći način:

$$I = (G_1 *_M G_2)(1) = \int_0^\infty G_1\left(\frac{1}{t}\right) G_2(t) \frac{dt}{t}.$$

Na osnovu (1.34), integral I može biti napisan kao inverzna Melinova transformacija proizvoda Melinovih slika $\hat{G}_1(-s)$ i $\hat{G}_2(s)$ funkcija $G_1\left(\frac{1}{t}\right)$ i $G_2(t)$, respektivno. Dakle,

$$I = \frac{1}{2\pi i} \int_C \hat{G}_1(-s) \hat{G}_2(s) ds, \quad (3.4)$$

gdje je C kontura u kompleksnoj ravni od $\gamma - i_\infty$ do $\gamma + i_\infty$, $Re(s) = \gamma$, tako da $\frac{b_j+v}{B_j}$ i $\frac{a_j-1-v}{A_j}$ leže sa različite strane konture C , gdje su $a_j \in \{0,1\}$, $b_j \in \{0,k+1\}$,

$A_j, B_j \in \{1, \frac{1}{a}\}$ i $v = 0, 1, 2, \dots$

Na osnovu leme 3.1 imamo

$$\hat{G}_1(s) = \int_0^\infty x^{s-1} \log(x+1) dx = \frac{\pi}{s \sin(\pi s)}$$

i

$$\hat{G}_2(s) = \int_0^\infty x^{s-1} \frac{ax^{a(k+1)}}{(1+x^a)^{k+2}} dx = \frac{\Gamma(1 - \frac{1}{a})\Gamma(\frac{s}{a} + k + 1)}{\Gamma(k + 2)}.$$

Primjenom (3.4) dobijamo:

$$I = \frac{1}{2\pi i} \int_C \frac{\pi}{s \sin(\pi s)} \frac{\Gamma(1 - \frac{1}{a})\Gamma(\frac{s}{a} + k + 1)}{\Gamma(k + 2)} ds. \quad (3.5)$$

Dalje, primjenom osobina za gama funkciju i Ojlerove formule za refleksiju (1.18) imamo da je

$$s \sin(\pi s) = \frac{\pi \Gamma(1 + s)}{\Gamma(s)\Gamma(s)\Gamma(1 - s)}. \quad (3.6)$$

Zamjenom (3.6) u (3.5) dobijamo:

$$I = \frac{1}{2\pi i} \frac{1}{\Gamma(k + 2)} \int_C \frac{\Gamma(s)\Gamma(s)\Gamma(1 - s)\Gamma(1 - \frac{1}{a})\Gamma(\frac{s}{a} + k + 1)}{\Gamma(s + 1)} ds.$$

Korišćenjem definicije 1.33 za Foksovu H funkciju dobijamo traženi rezultat. \square

Teorema 3.2. *Neka su $X \sim \text{Log}(\lambda, \theta), Y \sim \text{Log}(\mu, \phi)$ nezavisne slučajne veličine. Neka je*

$$a = \frac{\phi\beta}{\theta\alpha}; \quad A = \exp \left\{ -\frac{1}{\theta} \left(\frac{z - \mu\beta}{\alpha} - \lambda \right) \right\}.$$

Tada funkcija raspodjele $F_Z(z)$ slučajne veličine $Z = \alpha X + \beta Y$ je data sa:

a) ako je $\alpha\beta > 0$

$$F_Z(z) = aA^{-\frac{1}{a}} H_{3,3}^{3,2} \left[A^{-\frac{1}{a}} \mid \begin{matrix} (0, 1) & (\frac{1}{a}, -\frac{1}{a}) & (1, 1) \\ (0, 1) & (0, 1) & (2 - \frac{1}{a}, \frac{1}{a}) \end{matrix} \right] - (a - 1)A^{-\frac{1}{a}} H_{3,3}^{3,2} \left[A^{-\frac{1}{a}} \mid \begin{matrix} (0, 1) & (\frac{1}{a}, -\frac{1}{a}) & (1, 1) \\ (0, 1) & (0, 1) & (1 - \frac{1}{a}, \frac{1}{a}) \end{matrix} \right]; \quad (3.7)$$

b) ako je $\alpha\beta < 0$

$$F_Z(z) = 1 - aA^{\frac{2}{a}} H_{3,3}^{3,2} \left[A^{\frac{1}{a}} \mid \begin{matrix} (0, 1) & (-\frac{2}{a}, \frac{1}{a}) & (1, 1) \\ (0, 1) & (0, 1) & (2 - \frac{2}{a}, \frac{1}{a}) \end{matrix} \right]$$

$$+ (a-1)A^{\frac{2}{a}}H_{3,3}^{3,2}\left[A^{\frac{1}{a}}\left|\begin{array}{ccc}(0,1) & (-\frac{2}{a},\frac{1}{a}) & (1,1) \\ (0,1) & (0,1) & (1-\frac{2}{a},\frac{1}{a})\end{array}\right.\right]. \quad (3.8)$$

Dokaz. Daćemo dokaz za $\alpha\beta > 0$.

Funkcija raspodjele $F_Z(z) = P\{\alpha X + \beta Y \leq z\}$ može se izraziti kao

$$F_Z(z) = \int_{\mathbb{R}} F_X\left(\frac{z - \beta y}{\alpha}\right) f_Y(y) dy. \quad (3.9)$$

Zamjenom (1.51) i (1.52) u (3.9) dolazimo do

$$F_Z(z) = \frac{1}{\phi} \int_{\mathbb{R}} \frac{e^{-\frac{(y-\mu)}{\phi}} dy}{\left(1 + e^{-\frac{(y-\mu)}{\phi}}\right)^2 \left(1 + e^{-\frac{1}{\theta}[(z-\beta y)/\alpha - \lambda]}\right)}.$$

Uvodeći smjenu $x = e^{-\frac{(y-\mu)}{\phi}}$ dobijamo

$$F_Z(z) = \int_0^\infty \frac{x^a dx}{(1+x)^2 (A+x^a)} \quad (3.10)$$

Dakle,

$$\begin{aligned} F_Z(z) &= \int_0^\infty \frac{dx}{(1+x)^2} dx - A \int_0^\infty \frac{dx}{(1+x)^2 (A+x^a)} = \\ &= 1 - A \int_0^\infty \frac{dx}{(1+x)^2 (A+x^a)}. \end{aligned}$$

Označimo sa I_1 integral na desnoj strani u posljednjoj jednakosti. Korišćenjem parcijalne integracije dobijamo

$$I_1 = \frac{1}{A} - \int_0^\infty \frac{ax^{a-1}}{(1+x)(A+x^a)^2} dx.$$

Odnosno,

$$I_1 = \frac{1}{A} - I_2,$$

gdje je $I_2 = \int_0^\infty \frac{ax^{a-1}}{(1+x)(A+x^a)^2} dx$.

Da bi izračunali I_2 koristimo parcijalnu integraciju sa smjenom $u = \frac{ax^{a-1}}{(A+x^a)^2}$ i $v = \log(1+x)$. Tako da,

$$I_2 = \left. \frac{ax^{a-1} \log(1+x)}{(A+x^a)^2} \right|_0^\infty - \int_0^\infty \log(x+1) \frac{a(a-1)x^{a-2}(A+x^a) - 2a^2x^{2a-2}}{(A+x^a)^3} dx.$$

Primjenom leme 3.2. dobijamo

$$I_2 = I_3 - I_4,$$

gdje su $I_3 = \int_0^\infty \log(1+x) \frac{2a^2 x^{2a-2}}{(A+x^a)^3} dx$ i $I_4 = \int_0^\infty \log(1+x) \frac{a(a-1)x^{a-2}(A+x^a)}{(A+x^a)^3} dx$.

Za određivanje I_3 i I_4 koristimo metod iz teoreme 3.1. Tako da,

$$I_3 = \frac{1}{2\pi i} \int_C \hat{G}_1(-s) \hat{G}_2(s) ds, \quad (3.11)$$

gdje su

$$\hat{G}_1(s) = \int_0^\infty x^{s-1} \log(x+1) dx = \frac{\pi}{s \sin(\pi s)} \quad (3.12)$$

i

$$\begin{aligned} \hat{G}_2(s) &= \int_0^\infty x^{s-1} \frac{2a^2 x^{2a-1}}{(A+x^a)^3} dx = \\ &= \frac{1}{A^3} \int_0^\infty x^{s-1} \frac{2a^2 x^{2a-1}}{(1+(\frac{x}{A^{\frac{1}{a}}}))^3} dx. \end{aligned}$$

Primjenom smjena $x = A^{\frac{1}{a}} t$ i $dx = A^{\frac{1}{a}} dt$ dobijamo

$$\hat{G}_2(s) = 2aA^{\frac{s-1-a}{a}} \int_0^\infty t^{(s-1)-1} \frac{at^{a(1+1)}}{(1+t^a)^{1+2}} dt.$$

Ukoliko s i k u lemi 3.1 (b) zamijenimo sa $s-1$ i 1 respektivno, imamo da je

$$\hat{G}_2(s) = aA^{\frac{s-1-a}{a}} \Gamma(2 - \frac{1}{a} + \frac{s}{a}) \Gamma(1 + \frac{1}{a} - \frac{s}{a}). \quad (3.13)$$

Kombinujući (3.11), (3.12) i (3.13) dolazimo do sljedećeg rezultata:

$$I_3 = aA^{-\frac{a+1}{a}} H \left[A^{-\frac{1}{a}} \left| \begin{array}{ccc} (0, 1) & (-\frac{1}{a}, \frac{1}{a}) & (1, 1) \\ (0, 1) & (0, 1) & (2 - \frac{1}{a}, \frac{1}{a}) \end{array} \right. \right]. \quad (3.14)$$

Slično,

$$I_4 = (a-1)A^{-\frac{a+1}{a}} H \left[A^{-\frac{1}{a}} \left| \begin{array}{ccc} (0, 1) & (-\frac{1}{a}, \frac{1}{a}) & (1, 1) \\ (0, 1) & (0, 1) & (1 - \frac{1}{a}, \frac{1}{a}) \end{array} \right. \right]. \quad (3.15)$$

I_2 dobijamo kao razliku dva prethodna rezultata. Znamo da je $I_1 = \frac{1}{A} - I_2$. Dakle,

$$\int_0^\infty \frac{x^a}{(1+x)^2(A+x^a)} dx = AI_2.$$

Iz posljednje jednakosti slijedi traženi rezultat u (3.7).

S druge strane, u slučaju kada je $\alpha\beta < 0$, početni integral možemo lako dobiti iz (3.9) smjenom $a = -k$, gdje je $k > 0$. Nakon dvostruke parcijalne integracije kao i nakon primjene istih tehnika kao u prethodnom slučaju dolazimo do rezultata u (3.8). \square

Daćemo neke specijalne slučajeve prethodne teoreme.

Napomena 2. Ako su $\alpha = 1$ i $\beta = 1$ u (3.7) dobijamo funkciju raspodjele za $Z = X + Y$ za $a = \phi\theta^{-1}$ i $A = \exp\{-\frac{1}{\theta}(z - \mu - \lambda)\}$.

Posljedica 3.1. Pretpostavimo da su X i Y nezavisne slučajne veličine sa standardnom logističkom raspodjelom. Tada je funkcija raspodjele od $Z = X + Y$ data sa

$$F_Z(z) = \frac{e^z(e^z - 1 - z)}{(e^z - 1)^2}. \quad (3.16)$$

Dokaz. Za standardnu logističku raspodjelu $a = 1$, $A = e^{-z}$ zbog toga se (3.10) svodi na

$$F_Z(z) = \int_0^\infty \frac{x \, dx}{(1+x)^2(e^{-z} + x)} = \frac{1 - e^{-z} - ze^{-z}}{(1 - e^{-z})^2},$$

što je ekvivalentno izrazu (3.16). Izvođenje posljednjeg izraza je rutinsko. \square

Napomena 3. Kako je funkcija gustine standardne logističke raspodjele simetrična, onda funkcija raspodjele od $Z = X - Y$ ima isti oblik kao (3.16).

Teorema 3.3. Neka su $X \sim \text{Log}(\lambda, \theta)$, $Y \sim \text{Log}(\mu, \phi)$ nezavisne slučajne veličine. Tada je funkcija gustine slučajne veličine $Z = \alpha X + \beta Y$ data sa

$$f_Z(z) = H_{2,2}^{2,2} \left[e^{z - \alpha\lambda - \beta\mu} \mid \begin{matrix} (0, |\alpha|\theta) & (0, |\beta|\phi) \\ (1, |\alpha|\theta) & (1, |\beta|\phi) \end{matrix} \right]. \quad (3.17)$$

Dokaz. Primjenom osobina za karakterističnu funkciju kao i izraza (1.53) dobijamo karakterističnu funkciju za Z

$$\begin{aligned} \phi_Z(t) &= \mathbb{E}(\exp(iZ)) = \mathbb{E}(\exp(it(\alpha X + \beta Y))) = \mathbb{E}(\exp(it\alpha X))\mathbb{E}(\exp(it\beta Y)) = \\ &= e^{it(\alpha\lambda + \beta\mu)} \Gamma(1 - it|\alpha|\theta) \Gamma(1 + it|\alpha|\theta) \Gamma(1 - it|\beta|\phi) \Gamma(1 + it|\beta|\phi). \end{aligned}$$

Korišćenjem formule za inverziju (1.3), gustina od Z može se izraziti na sljedeći način

$$f_Z(z) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{it(\alpha\lambda + \beta\mu - z)} \Gamma(1 - it|\alpha|\theta) \Gamma(1 + it|\alpha|\theta) \Gamma(1 - it|\beta|\phi) \Gamma(1 + it|\beta|\phi) dt. \quad (3.18)$$

Smjenama $it = j$ i $dt = \frac{dt}{i}$ prethodni izraz svodimo na

$$f_Z(z) = \frac{1}{2i\pi} \int_{-\infty}^{+\infty} e^{j(\alpha\lambda + \beta\mu - z)} \Gamma(1 - j|\alpha|\theta) \Gamma(1 + j|\alpha|\theta) \Gamma(1 - j|\beta|\phi) \Gamma(1 + j|\beta|\phi) dj. \quad (3.19)$$

Traženi rezultat (3.17) slijedi nakon direktne primjene definicije 1.33. \square

3.1 Algoritam za računanje vrijednosti Foksove H funkcije

U prethodnom poglavlju izveli smo raspodjelu za linearnu kombinaciju dvije nezavisne logistički raspodijeljene slučajne veličine. Pošto dobijena raspodjela zavisi od Foksove H funkcije bilo je neophodno razviti odgovarajući alogoritam za računanje njenih vrijednosti.

U ovom poglavlju opisaćemo algoritam koji smo implementirali na osnovu rezultata iz [26] i [44]. Cio algoritam je pripremljen u softveru Wolfram Mathematica i može se pronaći u dodatku A.

U radovima [56] i [37] možemo pronaći alternativan Mathematica kod za Foksovu H funkciju.

Algoritam implementiran u [56] koristi ekvivalenciju između Foksove H i Meijerove G funkcije date relacijom (1.25) za sve slučajeve kada su parametri Foksove H funkcije racionalni brojevi i autor koristi tri različite konture kao što je opisano u [26].

Za razliku od algoritma u [56], mi predlažemo algoritam koji automatski razdvaja polove gama funkcija. Nakon što se polovi uoče oni se razdvajaju pravom. Koliko nam je poznato, algoritam za izračunavanje Foksove H funkcije koji ovde predlažemo pojavljuje se prvi put u literaturi.

Ulazni parametri za Foksovu H funkciju dati su strukturom liste, tj.

$$a = \{ \{ \{ a_1, A_1 \}, \dots, \{ a_n, A_n \} \}, \{ \{ a_{n+1}, A_{n+1} \}, \dots, \{ a_p, A_p \} \} \}$$

i

$$b = \{ \{ \{ b_1, B_1 \}, \dots, \{ b_m, B_m \} \}, \{ \{ b_{m+1}, B_{m+1} \}, \dots, \{ b_q, B_q \} \} \}.$$

Parametri su specifikovani definicijom 1.33.

Za izračunavanje proizvoda gama funkcija, koji se pojavljuje u (1.21) koristili smo ugrađenu funkciju `Product`. Na primjer, imenilac računamo na sljedeći način:

```
Product[Gamma[1-a[[1,j,1]]-a[[1,j,2]]*s],
{j,1,Length[a[[1]]]}]*
*Product[Gamma[b[[1,j,1]]+b[[1,j,2]]*s],
{j,1,Length[b[[1]]]}],
```

gdje `a[[]]` vraća djelove ili podnizove liste `a`.

Važan zadatak je da se odredi prava koja razdvaja polove od $\Gamma(b_i + B_i s)$ i $\Gamma(1 - a_j - A_j s)$ za sve $i = 1, \dots, m$ i $j = 1, \dots, n$.

U suštini, to je tačno ako je $A_j(b_i + k) \neq B_i(a_j - 1 - l)$, za sve $i = 1, \dots, m$, $j = 1, \dots, n$, i $k, l = 0, 1, 2, \dots$. Koristeći jednakosti (1.22) i (1.23) to možemo odraditi na sljedeći način:

```
PolesLeft[b,p]=
DeleteDuplicates[Table[-(b[[1,i,1]]+k)/b[[1,i,2]],
{i,1,Length[b[[1]]]}, {k,0,p}]]
PolesRight[a,p]=
DeleteDuplicates[Table[(1-a[[1,j,1]]+k)/a[[1,j,2]],
{j,1,Length[a[[1]]]}, {k,0,p}]].
```

Sada, da bismo odredili poziciju konture L definišemo realni parametar γ . To je odradeno kreiranjem funkcije `ParameterGamma[a_, b_]`. Odnosno, parametar γ predstavlja središnju vrijednost između dva najbliža pola za dvije grupe gama funkcija $\Gamma(b_j + sB_j)$ i $\Gamma(1 - a_\lambda - sA_\lambda)$. Nakon što smo odredili parametar γ , ostaje da provjerimo, sa funkcijom `NumPolesRight`, da li je bilo koji od desnih polova pogrešno dodat pravoju L .

Tada kontura počinje u $\gamma - i\infty$ i završava u $\gamma + i\infty$, gdje $\gamma \in \mathbb{R}$ razdvaja polove od $\Gamma(b_i + B_i s)$ i $\Gamma(1 - a_j - A_j s)$ za svako $i = 1, \dots, m$ i $j = 1, \dots, n$. Ovdje se, integral iz (1.21) definisan nad konačnim intervalima `ParameterGamma[a_, b_] - i D` i `ParameterGamma[a_, b_] + i D` računa korišćenjem ugrađene funkcije `NIntegrate`, gdje je D neki veliki broj, mi smo uzeli $D = 10000$.

U slučaju kada je neki od polova pogrešno dodat pravoju L moramo oduzeti reziduale u polovima od vrijednosti rezultata funkcije `NIntegrate`. To postizemo korišćenjem funkcije `NResidue`.

Konačno, algoritam za numeričko izračunavanje Foksove H funkcije pozivamo na sljedeći način

```
Hfunction[{{a_1, A_1}, ..., {a_n, A_n}},
{{a_{n+1}, A_{n+1}}, ..., {a_p, A_p}}},
{{b_1, B_1}, ..., {b_m, B_m}},
{{b_{m+1}, B_{m+1}}, ..., {b_q, B_q}}}, z, pg]
```

gdje sa pg zadajemo željenu preciznost.

Uporedili smo naš algoritam za računanje vrijednosti Foksove H funkcije sa algoritmom predloženim u [37]. Zaključili smo da taj algoritam pravi određena odstupanja. Na primjer, razmatrali smo vrijednost za funkciju raspodjele u tački $y = 12$ (tabela B.2, dodatak B) algoritam iz [37] odredio je vrijednost $7,49999999998e-01$, što je prilično daleko od stvarne vrijednosti kao i od vrijednosti date u tabeli B.1 iz dodatka B. Još jedan primjer odstupanja je u slučaju kad nedostaje član $\prod_{j=1}^n \Gamma(1 - a_j - A_j s)$ u izrazu 1.21. Iz nekih razloga algoritam iz [37] vraća vrijednost nula. Na primjer, koristeći naš algoritam za računanje sljedeće Foksove H funkcije

$$H_{1,2}^{2,0} \left[0, 3 \mid \begin{array}{cc} (0, 8; 1) & --- \\ (1, 1; 1) & (0, 1; 1) \end{array} \right],$$

dobijamo $6,21654039663e-01$, dok algoritam implementiran u [37] vraća 0.

Očigledno da algoritam koji smo prethodno opisali je mnogo pouzdaniji u odnosu na algoritam predložen u [37].

U tabeli 3.1 uporedili smo rezultate dobijene numeričkim računanjem (3.10) i rezultate dobijene primjenom algoritma za Foksovu H funkciju za izračunavanje (3.7) i (3.8).

TABELA 3.1: Upoređivanje dva metoda za izračunavanje funkcije raspodjele u određenim tačkama

Parametri	Tačka	Numerički	H-funkcija
$\alpha = 1; \beta = 1; \lambda = 0, 2; \theta = 0, 5; \mu = 0, 5; \phi = 0, 3$	0,2	0,305565	0,305565
$\alpha = 0, 5; \beta = 1; \lambda = 0, 1; \theta = 0, 4; \mu = 0, 3; \phi = 0, 1$	1	0,947609	0,947609
$\alpha = 2; \beta = 5; \lambda = -0, 5; \theta = 1; \mu = -0, 1; \phi = 0, 5$	-3	0,363831	0,363831
$\alpha = 3; \beta = 15; \lambda = 3, 6; \theta = 3; \mu = 1, 5; \phi = 2$	8	0,311664	0,311664

Dalje, kako smo teoremom 3.1 dobili funkciju raspodjele, možemo numerički dobiti percentilne vrijednosti rješavajući jednačinu $F_Z(z_p) = p$, $0 < p < 1$, tj. p -ti percentil,

za raspodjelu Z dobili smo numerički rješavajući jednačinu

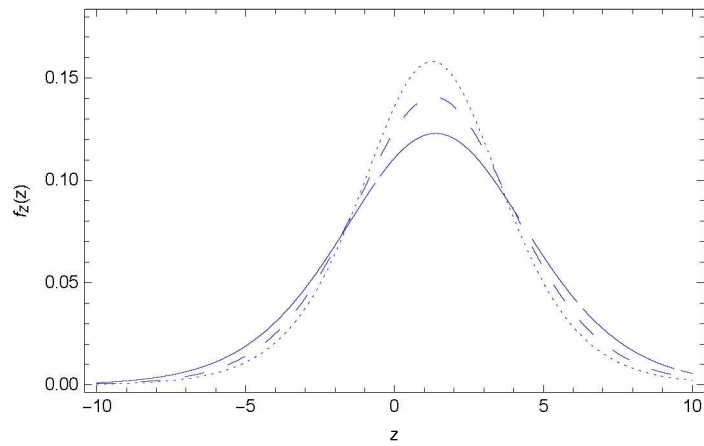
$$\int_0^{z_p} f_Z(z) dz = p, \quad 0 < p < 1.$$

Izabrali smo određene vrijednosti za parametre raspodjele slučajne veličine Z i korišćenjem algoritma za numeričko izračunavanje Foksove H funkcije odredili smo odgovarajuće percentile. Koristili smo ugrađenu funkciju `FindRoot` u softveru `Mathematica` i rezultati za $p \in \{0,9; 0,95; 0,975; 0,99\}$ dati su u tabeli 3.2. Rezultati takode odgovaraju raspodjelama prikazanim na slici 3.1.

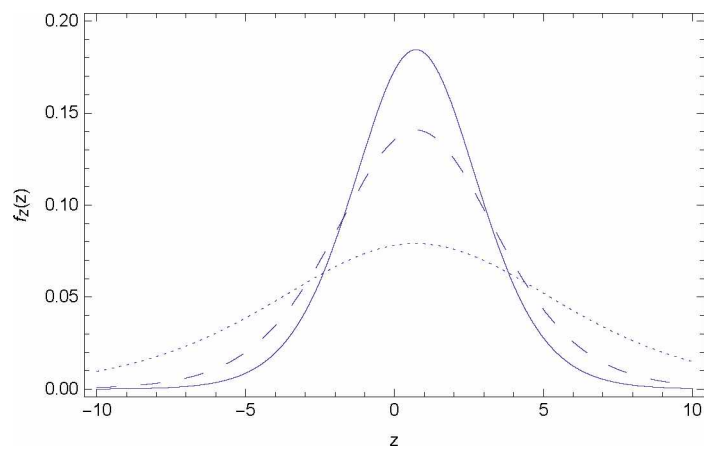
U sljedećoj glavi bavili smo uopštenjem dobijenih rezultata u slučaju kada posmatramo linearnu kombinaciju više od dvije slučajne veličine.

TABELA 3.2: Percentilne vrijednosti za raspodjelu slučajne veličine Z sa parametrima $\lambda = 2$; $\theta = 2, 5$; $\mu = 0, 3$ i $\phi = 2$

α	β	p			
		0,90	0,95	0,975	0,99
0,6	0,2	4,71681	5,65286	6,22347	6,59990
0,6	0,4	5,12195	6,26769	7,08954	7,69108
0,6	0,6	5,70531	7,02901	8,04942	8,84698
0,8	0,2	6,03490	7,03297	7,60399	7,97022
0,8	0,4	6,48875	7,83215	8,71709	9,33097
0,8	0,6	7,00356	8,56149	9,68476	10,5150
1	1	9,50706	11,6929	13,3500	14,6297
1	1,5	11,2473	13,9701	16,1605	17,9547
1	2	13,1836	16,5130	19,2708	21,5994
0,6	-0,4	4,56612	5,12255	5,67798	6,4109
0,5	-0,5	3,54889	4,2035	4,85642	5,7215
1	-1	7,09777	8,40701	9,71289	11,443



(A) Funkcije gustine za raspodjelu Z sa parametrima $\alpha = 0,6$; $\lambda = 2$; $\phi = 2$; $\mu = 0,3$; $\theta = 2,5$; $\beta = 0,2$ (tačkasta linija); $\beta = 0,4$ (isprekidana linija) i $\beta = 0,6$ (puna linija)



(B) Funkcije gustine za raspodjelu Z sa parametrima $\alpha = 0,6$; $\beta = 0,4$; $\phi = 2$; $\mu = 0,3$; $\lambda = 1$, $\theta = 5$ (tačkasta linija); $\theta = 2,5$ (isprekidana linija); $\theta = 1,67$ (puna linija)

SLIKA 3.1: Funkcije gustine za raspodjelu Z

Glava 4

Raspodjela linearne kombinacije ($n > 2$) nezavisnih logistički raspodijeljenih slučajnih veličina

U ovoj glavi izložićemo rezultate iz rada [35].

Razmotrimo linearnu kombinaciju n nezavisnih logistički raspodijeljenih slučajnih veličina, recimo

$$Y = \sum_{k=1}^n c_k X_k,$$

gdje c_k su poznati realni koeficijenti i $X_k \sim \text{Log}(\mu_k, \beta_k)$ su nezavisne slučajne veličine sa logističkim raspodjelama sa parametrima lokacije μ_k ($\mu_k \in \mathbb{R}$) i parametrima razmjere β_k ($\beta_k > 0$) za $k = 1, \dots, n$, sa funkcijama gustine definisanim sa (1.52) i karakterističnim funkcijama

$$\phi_{X_k}(t) = e^{i\mu_k t} (\pi\beta_k t) \operatorname{csch}(\pi\beta_k t).$$

Ovdje je, i kompleksna jedinica, $i = \sqrt{-1}$ i $\operatorname{csch}(\cdot)$ je hiperbolički kosekans. Dalje, koristeći nezavisnost slučajnih veličina X_k , $k = 1, 2, \dots, n$ i osobine karakteristične funkcije dobijamo da je karakteristična funkcija od Y definisana sa

$$\phi_Y(t) = \prod_{k=1}^n \phi_{X_k}(c_k t) = e^{it \sum_{k=1}^n c_k \mu_k} \prod_{k=1}^n (\pi c_k \beta_k t) \operatorname{csch}(\pi c_k \beta_k t). \quad (4.1)$$

U [25] obrađen je problem izračunavanja tačne raspodjele linearne kombinacije nezavisnih logističkih slučajnih varijabli i razvijene su dvije skoro tačne aproksimacije za

ovu raspodjelu. Ove dvije aproksimacije su korisne za različite svrhe, prva za slučajeve gdje je brzina računanja važna, a druga za slučajeve gdje je potrebna visoka preciznost.

U sljedećem poglavlju dali smo tačnu raspodjelu slučajne veličine Y primjenom Foksove H funkcije. Osim algoritma za numeričko računanje Foksove H funkcije, koristili smo i metod za numeričku inverziju karakteristične funkcije. U daljem radu, upoređićemo ova dva pristupa.

4.1 Rapodjela linearne kombinacije n nezavisnih logistički raspodijeljenih slučajnih veličina primjenom Foksove H funkcije

Teorema 4.1. *Funkcija gustine slučajne veličine Y je data sa*

$$f_Y(y) = H_{n,n}^{n,n} \left[u(y) \left| \begin{array}{ccc} (0, |c_1|\beta_1) & \cdots & (0, |c_n|\beta_n) \\ (1, |c_1|\beta_1) & \cdots & (1, |c_n|\beta_n) \end{array} \right. \right], \quad (4.2)$$

dok je funkcija raspodjele data sa

$$F_Y(y) = 1 - H_{n+1,n+1}^{n+1,n} \left[u(y) \left| \begin{array}{cccc} (0, |c_1|\beta_1) & \cdots & (0, |c_n|\beta_n) & (1, 1) \\ (1, |c_1|\beta_1) & \cdots & (1, |c_n|\beta_n) & (0, 1) \end{array} \right. \right], \quad (4.3)$$

gdje je $u(y) = \exp \left(y - \sum_{k=1}^n c_k \mu_k \right)$.

Dokaz. Izraz (4.2) možemo lako dobiti primjenom metoda karakterističnih funkcija. Karakterističnu funkciju od Y možemo izraziti na sljedeći način

$$\phi_Y(t) = \mathbb{E}(e^{itY}) = \mathbb{E} \left(e^{it \sum_{k=1}^n c_k X_k} \right) = \prod_{k=1}^n \mathbb{E}(e^{it c_k X_k}).$$

Da bismo pojednostavili zapis, izvešćemo karakterističnu funkciju za $c_1 X_1$, tj. kada je $k = 1$. Sada korisćenjem (1.52) imamo

$$\mathbb{E}(e^{it c_1 X_1}) = \int_{-\infty}^{+\infty} e^{it c_1 x_1} \frac{\exp \left(-\frac{x_1 - \mu_1}{\beta_1} \right)}{\beta_1 \left(1 + \exp \left(-\frac{x_1 - \mu_1}{\beta_1} \right) \right)^2} dx_1.$$

Smjenom $\exp\left(-\frac{x_1-\mu_1}{\beta_1}\right) = u$ posljednji izraz svodimo na

$$\mathbb{E}\left(e^{itc_1X_1}\right) = \int_0^{+\infty} \frac{\exp(itc_1(\mu_1 - \beta_1 \ln u))}{(1+u)^2} du = e^{itc_1\mu_1} \int_0^{+\infty} \frac{u^{itc_1\beta_1}}{(1+u)^2} du. \quad (4.4)$$

S druge strane, očigledno da (4.4) predstavlja oblik beta integrala, tako da

$$\mathbb{E}\left(e^{itc_1X_1}\right) = e^{itc_1\mu_1} \Gamma(1+itc_1\beta_1) \Gamma(1-itc_1\beta_1) = e^{itc_1\mu_1} \Gamma(1+it|c_1|\beta_1) \Gamma(1-it|c_1|\beta_1). \quad (4.5)$$

Dakle, zaključujemo da je karakteristična funkcija slučajne veličine Y data sa

$$\phi_Y(t) = e^{it \sum_{k=1}^n c_k \mu_k} \prod_{k=1}^n \Gamma(1+it|c_k|\beta_k) \Gamma(1-it|c_k|\beta_k). \quad (4.6)$$

Primjenom teoreme za inverziju karakteristične funkcije dobijamo funkciju gustine za Y . Zaista,

$$f_Y(y) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{it(\sum_{k=1}^n c_k \mu_k - y)} \prod_{k=1}^n \Gamma(1+it|c_k|\beta_k) \Gamma(1-it|c_k|\beta_k) dt.$$

Smjenom $it = s$ dobijamo

$$f_Y(y) = \frac{1}{2\pi i} \int_{-\infty}^{+\infty} e^{-s(y - \sum_{k=1}^n c_k \mu_k)} \prod_{k=1}^n \Gamma(1+|c_k|\beta_k s) \Gamma(1-|c_k|\beta_k s) ds. \quad (4.7)$$

Sada, (4.2) slijedi iz definicije 1.33.

Da bismo izveli (4.3) posmatrajmo funkciju preživljavanja slučajne veličine Y , definisane sa (1.1).

Dalje, ako uvedemo da je $t(x) = x - \sum_{k=1}^n c_k \mu_k$ tada iz (4.7) dobijamo

$$\begin{aligned} \overline{F}_Y(y) &= \frac{1}{2\pi i} \int_y^{+\infty} \int_L \prod_{k=1}^n e^{-st(x)} \Gamma(1+|c_k|\beta_k s) \Gamma(1-|c_k|\beta_k s) ds dx = \\ &= \frac{1}{2\pi i} \int_L \prod_{k=1}^n \Gamma(1+|c_k|\beta_k s) \Gamma(1-|c_k|\beta_k s) \int_y^{+\infty} e^{-st(x)} dx ds. \end{aligned}$$

Korišćenjem identiteta $s = \frac{\Gamma(1+s)}{\Gamma(s)}$ dolazimo do

$$\int_y^{+\infty} e^{-s(x+a)} dx = \frac{e^{-s(y+a)}}{s} = \frac{\Gamma(s) e^{-s(x+a)}}{\Gamma(1+s)}.$$

Sada, lako izvodimo da je

$$\overline{F}_Y(y) = \frac{1}{2\pi i} \int_L e^{-st(x)} \frac{\Gamma(s)}{\Gamma(1+s)} \prod_{k=1}^n \Gamma(1 + |c_k| \beta_k s) \Gamma(1 - |c_k| \beta_k s) ds. \quad (4.8)$$

Iz (4.8), korišćenjem definicije 1.33 i relacije $F_Y(y) = 1 - \overline{F}_Y(y)$, dobijamo izraz (4.3) za funkciju raspodjele. \square

Napomena 4. Korišćenjem identiteta $\Gamma(1+z)\Gamma(1-z) = \pi z \csc(\pi z) = i\pi z \operatorname{csch}(i\pi z)$, gdje je \csc kosekant funkcija, (4.6) svodimo na (4.1).

Rezultat iz sljedeće teoreme nam omogućava da računamo matematičko očekivanje za nenegativnu slučajnu veličinu Y na dva različita načina, što ćemo kasnije iskoristiti za numeričko upoređivanje algoritama.

Teorema 4.2. Neka je $Y = \sum_{k=1}^n c_k X_k$ nenegativna slučajna veličina. Tada očekivanje od Y možemo izraziti kao

$$\mathbb{E}(Y) = H_{n+2, n+2}^{n+2, n} \left[z \left| \begin{array}{cccccc} (0, |c_1| \beta_1) & \dots & (0, |c_n| \beta_n) & (1, 1) & (1, 1) \\ (1, |c_1| \beta_1) & \dots & (1, |c_n| \beta_n) & (0, 1) & (0, 1) \end{array} \right. \right], \quad (4.9)$$

gdje je $z = \exp \left(- \sum_{k=1}^n c_k \mu_k \right)$.

Dokaz. Korišćenjem (4.2) očekivanje slučajne veličine Y se može izraziti kao

$$\mathbb{E}(Y) = \int_0^{+\infty} y H_{n, n}^{n, n} \left[z \left| \begin{array}{cccc} (0, |c_1| \beta_1) & \dots & (0, |c_n| \beta_n) \\ (1, |c_1| \beta_1) & \dots & (1, |c_n| \beta_n) \end{array} \right. \right] dy.$$

Primjenom definicije za Foksovu H funkciju (1.33) kao i imajući u vidu sljedeći rezultat

$$\int_0^{+\infty} y e^{-s(y+a)} dy = \frac{e^{-sa}}{s^2} = \frac{\Gamma(s)\Gamma(s)e^{-sa}}{\Gamma(1+s)\Gamma(1+s)} \quad (4.10)$$

jednostavan račun vodi nas do (4.9). \square

Napomena 5.

- a) Ako slučajna veličina Y uzima negativne vrijednosti njeno očekivanje se ne može izraziti preko Foksove H funkcije, jer nije ispunjena konvergencija integrala (4.10).
- b) Matematičko očekivanje od slučajne veličine Y koja uzima vrijednosti iz \mathbb{R} može se izračunati primjenom linearnosti matematičkog očekivanja i korišćenjem činjenice

da je $\mathbb{E}(X_k) = \mu_k$, $k = 1, \dots, n$, tj.

$$\mathbb{E}(Y) = \sum_{k=1}^n c_k \mu_k. \quad (4.11)$$

4.2 Raspodjela linearne kombinacije n nezavisnih slučajnih veličina. Primjena metoda za numeričku inverziju karakteristične funkcije

U ovom poglavlju predstavimo jedan od osnovnih numeričkih metoda za računanje raspodjele za linearnu kombinaciju nezavisnih slučajnih veličina, tj. opisaćemo dobro poznati i često korišćeni metod za numeričku inverziju karakteristične funkcije.

Prisjetimo se da je funkcija gustine slučajne veličine Y apsolutno neprekidnog tipa sa karakterističnom funkcijom $\phi_Y(t)$ data inverznom Furijeovom transformacijom, tj. imamo da je

$$f_Y(y) = \frac{1}{\pi} \int_0^\infty \operatorname{Re} (e^{-ity} \phi_Y(t)) dt. \quad (4.12)$$

Čak, ako je y tačka neprekidnosti funkcije raspodjele od Y , recimo $F(y) = P\{Y < y\}$, Gil-Pelaez je u [17] izveo metod za inverziju apsolutno integrabilne karakteristične funkcije $\phi(t)$ na $(-\infty, +\infty)$, koji je pogodan za numeričko generisanje funkcije raspodjele $F_Y(y)$,

$$F_Y(y) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \operatorname{Im} \left(\frac{e^{-ity} \phi_Y(t)}{t} \right) dt. \quad (4.13)$$

Uopšteno govoreći, numeričko izračunavanje Furijeove transformacije je važan i dobro proučen problem, često povezan sa izračunavanjem integrala visoko oscilatornih (kompleksnih) funkcija. Konkretno, korišćene su metode za invertovanje karakteristične funkcije za dobijanje funkcije raspodjele koje su predložene u [1], [41], [46], [53].

Algoritmi za numeričku inverziju karakteristične funkcije zasnovani na (4.12) i (4.13) uspješno su implementirani u MATLAB repozitorijumu CharFunTool, pogledati [55], gdje se nalaze potrebne funkcije i alati za numeričku inverziju karakteristične funkcije, koji se koristi za izračunavanje funkcije raspodjele, funkcije gustine i kvantilne funkcije.

U opštem slučaju, integrali u (4.12) i (4.13) mogu se izračunati primjenom bilo kojeg poznatog metoda numeričke kvadrature. Pokazano je da je za ovaj tip problema

najefikasnija aproksimacija integrala trapeznim pravilom (kao što je dato u [41]), tj.

$$f_Y(y) \approx \frac{\delta}{\pi} \sum_{j=0}^{\infty} w_j \operatorname{Re} \left(e^{-i\delta y(j+m)} \phi_Y(\delta(j+m)) \right), \quad (4.14)$$

$$F_Y(y) \approx \frac{1}{2} - \frac{\delta}{\pi} \sum_{j=0}^{\infty} w_j \operatorname{Im} \left(\frac{e^{-i\delta y(j+m)} \phi_Y(\delta(j+m))}{\delta(j+m)} \right), \quad (4.15)$$

gdje je $\delta > 0$ korak integracije i $0 \leq m \leq 1$ je srednja tačka korišćena u integraciji. U suštini, ako je $m = 0$, dobijamo metod lijevih pravougaonika (MLP), ako je $m = 1$ dobijamo metod desnih pravougaonika (MDP). Srednja vrijednost od MLP i MDP vodi do dobro poznatog trapeznog pravila. Tačnije, zamjenom beskonačnih suma u (4.14) i (4.15) konačnim sumama dobijamo aproksimaciju zasnovanu na trapeznom pravilu

$$f_Y(y) \approx \frac{\delta}{\pi} \sum_{j=0}^N w_j \operatorname{Re} \left(e^{-it_j y} \phi_Y(t_j) \right), \quad (4.16)$$

$$F_Y(y) \approx \frac{1}{2} - \frac{\delta}{\pi} \sum_{j=0}^N w_j \operatorname{Im} \left(\frac{e^{-it_j y} \phi_Y(t_j)}{t_j} \right), \quad (4.17)$$

gdje je N dovoljno veliki cio broj, recimo $N = 2^{10}$, w_j su odgovarajuće težine trapezoidne kvadrature (tj. $w_0 = w_N = \frac{1}{2}$ i $w_j = 1$ za $j = 1, \dots, N-1$), i $t_j = j\delta$ za $j = 0, \dots, N$ su ekvidistantni čvorovi (tako da $t_1 - t_0 = t_2 - t_1 = \dots = t_N - t_{N-1} =: \delta$) iz intervala $[0, T]$, za dovoljno veliko T , gdje je $T = \delta N$ tako da je $|\phi_Y(T)| \leq \epsilon$ za datu grešku tolerancije ϵ .

Alternativno, ako je $m = 0,5$ dobijamo metod srednjih pravougaonika, koji vodi do Rimanove sume. U radovima [12] i [41] je pokazano da je ovaj metod kvadrature pogodniji za analizu greške. Odnosno, greška indukovana numeričkom integracijom, tj. apsolutna razlika između (4.13) i (4.15) je mnogo manja ovdje. U radu [41] u posljedici 1, apsolutna vrijednost greške je ograničena sa

$$\max \left(F_Y \left(y - \frac{2\pi}{\delta} \right), 1 - F_Y \left(y + \frac{2\pi}{\delta} \right) \right). \quad (4.18)$$

Kao što je istaknuto u [12] i [41], moguće je predložiti kriterijum za izbor δ tako da ova greška bude što manja. U suštini, ako postoji generatrisa momenata $M_Y(t)$ tada je $1 - F_Y(y) \leq M_Y(t) \exp(-ty)$ za malo $t > 0$. Neka je $K(t) = \log(M_Y(t))$, birajući da je $y = K'(t)$ imamo da je

$$1 - F_Y(K'(t)) \leq \exp(K(t) - tK'(t)) \text{ i } F_Y(K'(t)) \leq \exp(K(-t) + tK'(t)). \quad (4.19)$$

Dakle, možemo ograničiti grešku integracije (4.18) pravilnim izborom za δ .

Optimalni izbor za δ je povezan sa rangom r domena raspodjele Y . Ako je domen apsolutno neprekidne slučajne veličine Y beskonačan, recimo $(-\infty, +\infty)$, definisali smo domen raspodjele Y u pokretnom zarezu kao interval $D = (D_{min}, D_{max})$ tako da je $F_Y(D_{min}) = \epsilon_0$ i $F_Y(D_{max}) = 1 - \epsilon_0$, gdje ϵ_0 je relativna preciznost u pokretnom zarezu (za 64-bit IEEE pokretni zarez, možemo koristiti vrijednost $\epsilon_0 = 2.2 \times 10^{-16}$, kao što se koristi u Matlabu). Tako da, donja granica od D je data kao $D_{min} = q_{\epsilon_0}$ i gornja granica od D je data sa $D_{max} = q_{1-\epsilon_0}$, gdje q_{ϵ_0} i $q_{1-\epsilon_0}$ su kvantilne funkcije raspodjele Y . Dakle, optimalan izbor za δ je $\delta = \frac{2\pi}{r}$, gdje je $r = D_{max} - D_{min}$. Stoga, na osnovu (4.18) apsolutna vrijednost greške indukovane numeričkom integracijom je ograničena sa ϵ_0 za svako $y \in D$.

Međutim, u praktičnim situacijama, nemoguće je odrediti tražene vrijednosti za kvantile jer ne znamo tačnu funkciju raspodjele. S druge strane, poznata je karakteristična funkcija i pomoću nje je moguće ocijeniti potrebne momente i korišćenjem odgovarajućih metoda za aproksimaciju (kao na primjer aproksimaciju zasnovanu na (4.19) ili pravilo šest sigmi sa korišćenjem Čebišovljeve nejednakosti), moguće je pronaći dobru ocjenu za interval domena tako da obuhvatimo veliki dio date raspodjele.

Pojedinačni izbor za δ i N utiče na ukupnu grešku aproksimacije, tj. to je zbir grešaka koje dobijamo prilikom integracije i prilikom zamjene beskonačke sume sa konačnom sumom. Odnos između njih veoma zavisi od karakteristične funkcije ϕ_Y . U opštem slučaju, ako je karakteristična funkcija analitička, formalno je pokazano da primjena trapeznog pravila vodi do veoma preciznog proračuna sa eksponencijalnim padom greške integracije, pogledati u [13].

Problem invertovanja karakteristične funkcije linearne kombinacije nezavisnih slučajnih veličina sa logističkom raspodjelom je pogodan za primjenu navedenog metoda. Karakteristična funkcija obično brzo opada ka nuli za velike vrijednosti svog argumenta t i velika preciznost se često dobija (za velike domene) od $N \approx 100$ (obično uzimamo $N = 2^7 = 128$).

4.3 Uprođivanje numeričke efikasnosti algoritama

U ovom poglavlju, upoređićemo efikasnost algoritma za numeričko računanje vrijednosti za Foksovu H funkciju.

Prvo, razmotrimo $Y = \sum_{i=1}^n c_i X_i$, gdje su parametri za $n = 4$ nezavisne logistički raspodijeljene slučajne veličine $X_i : \text{Log}(\mu_i, \beta_i)$ definisani sa $\mu = [-4, -1, 2, 3]$, $\beta = [1, 1, 1, 1]$ i koeficijenti su $c = [1, 1, 1, 1]$. Ako je $\beta = [1, 1, 1, 1]$ primjenom jednakosti (1.25) Foksova H funkcija se svodi na Meijerovu G funkciju. To nam omogućava da

uporedimo predloženi algoritam za numeričko računanje Foksove H funkcije sa već implementiranim algoritmom **MeijerG** u **Wolfram Mathematica**. Dobijeni rezultati su prikazani u tabeli 4.1. Uočavamo da oba algoritma prikazuju skoro jednake vrijednosti što nas navodi da zaključimo da algoritam za računanje Foksove H funkcije radi dobro.

TABELA 4.1: Gustina (PDF) i funkcija raspodjele (CDF) linearne kombinacije nezavisnih logistički raspodijeljenih slučajnih veličina primjenom algoritma za Foksovu H funkciju u Meijerovu G funkciju.

y	Meijerova G funkcija		Foksova H funkcija		Apsolutna razlika	
	PDF	CDF	PDF	CDF	PDF	CDF
-5	0,0401932948686445	0,081442443796992	0,0401933	0,0814424	$3,46945e-17$	$6,93889e-17$
-1	0,109124286474134	0,387635075995027	0,109124	0,387635	$1,29341e-14$	$1,19904e-14$
1	0,109124286474134	0,612364924004973	0,109124	0,612365	$1,29341e-14$	$1,19904e-14$
5	0,0401932948686445	0,9185575562030083	0,0401933	0,918558	$3,46945e-17$	$1,11022e-16$

U opštem slučaju nije moguće odrediti tačne vrijednosti za Foksovu H funkciju. Međutim, za testiranje numeričke efikasnosti našeg algoritma iskoristićemo dobro poznate identitete između Foksove H funkcije i nekih specijalnih funkcija, koje su implementirane u **Wolfram Mathematica**. Na primjer, za upoređivanje iskoristili smo identitete (1.26) - (1.29). Označimo sa \mathcal{L} lijevu stranu a sa \mathcal{R} desnu stranu u prethodnim identitetima.

Sada, ćemo prikazati kako su specijane funkcije iz (1.26)–(1.29) date u **Wolfram Mathematica**:

- a) Gama funkcija $\Gamma(z)$ - **Gamma[z]**,
- b) Gausova hipergeometrijska funkcija ${}_2F_1(b, a; c; -)$ - **Hypergeometric2F1[a,b,c,x]**,
- c) Mittag - Lefflerova funkcija $E_{a,b}(z)$ - **MittagLefflerE[a,b,z]**,
- d) Modifikovana Beselova funkcija druge vrste $Y_u(z)$ - **BesselK[u,z]**.

U nastavku su dati rezultati dobijenim primjenom našeg algoritma i prethodno navedenih specijanih funkcija, razlike u decimalnim ciframa su označene crvenom bojom.

- a) Jednakost (1.26): Ako je $u = [0, 8; 2, 5; -1, 5]$ i $z = [0, 2; 0, 6; 0, 8]$ dobijamo

$$\mathcal{L} = [1, 00622170788e+00; 4, 10521817607e-01; 5, 70719130855e+00]$$

i

$$\mathcal{R} = [1, 00622170788e+00; 4, 10521817607e-01; 5, 70719130855e+00].$$

b) Jednakost (1.27): Za $a = [2, 3; 0, 2]$, $b = [1, 4; 0, 3]$, $c = [2, 1; 0, 4]$ i $z = [5, 2; 0, 8]$, dobijamo vrijednosti

$$\mathcal{L} = [1, 66666666667e-01; 1, 64609053498e-02; 5, 66378208567e+00]$$

i

$$\mathcal{R} = [1, 66666666667e-01; 1, 64609053498e-02; 5, 66378208567e+00].$$

c) Jednakost (1.28): Ako je $\beta = [0, 1; 0, 5; 0, 8]$, $\alpha = [5, 1; 0, 5]$ i $z = [-3, 3; 0, 1]$, dobijamo

$$\mathcal{L} = [-2, 27115061653e-03; 3, 48556703885e+01; 9, 05699128711e-01]$$

i

$$\mathcal{R} = [-2, 27115061656e-03; 3, 48556703885e+01; 9, 8201869232e-01].$$

d) Jednakost (1.29): Za $a = [0, 2; 0, 3; 0, 5]$, $k = [0, 3; 0, 8; 1, 5]$ i $z = [0, 4; 2; 3, 5]$, imamo da je

$$\mathcal{L} = [-7, 04362495373e-01; 2, 73656603993e-02; 3, 48862274121e-01]$$

i

$$\mathcal{R} = [-7, 04362495373e-01; 2, 73656603993e-02; 3, 48862274121e-01].$$

Još jedan od načina da testiramo performanse našeg algoritma za računanje vrijednosti Foksove H funkcije je upoređivanje rezultata dobijenih u (4.9) i (4.11). Razmotrili smo slučajnu veličinu Y koja je linearna kombinacija $n = 4$ nezavisne logistički raspodjeljene slučajne veličine sa paramterima $\mu = [1; 2; 3; 0]$, $\beta = [0, 1; 0, 2; 0, 3; 0, 4]$ i koeficijentima $c = [1; 2; 3; 4]$. Sada, ako koristimo (4.9), vrijednost za matematičko očekivanje od Y je

$$\mathbb{E}(Y) = 1,40005089835e+01,$$

dok primjenom (4.11) dobijamo da je $\mathbb{E}(Y) = 14$. Za prethodne proračune koristili smo da je parametar preciznosti $\text{pg}=10$.

```

% Set the parameters, coefficients, y-values
% and probabilities of the calculated quantiles
mu    = [-4 -1 2 3];
beta  = [0.1 0.2 0.3 0.4];
c      = [1 2 3 4];
y      = [-5:0.5:30];
prob  = [0.80, 0.85, 0.90, 0.925, 0.95, 0.975, 0.99, 0.995, 0.999];

% Set the characteristic function
% of the linear combination of independent Logistic RVs
cf     = @(t) cf_Logistic(t,mu,beta,c);

% Set the optional control parameters
% for the numerical inversion algorithm
clear options
options.N = 2^12;

% Numerical inversion of the characteristic function
result = cf2DistGP(cf,y,prob,options);

```

SLIKA 4.1: Kod u MATLAB-u za izračunavanje PDF/CDF/QF od linearne kombinacije nezavisnih logistički raspodijeljenih slučajnih veličina korišćenjem CharFunTool.

4.4 Funkcija raspodjele, funkcija gustine i kvantilna funkcija linearne kombinacije nezavisnih logistički raspodijeljenih slučajnih veličina

Ovdje smo razmatrali $Y = \sum_{i=1}^n c_i X_i$, gdje su parametri za $n = 4$ nezavisne logistički raspodijeljene slučajne veličine $X_i \sim \text{Log}(\mu_i, \beta_i)$ dati sa $\mu = [-4; -1; 2; 3]$, $\beta = [0, 1; 0, 2; 0, 3; 0, 4]$ i koeficijenti su $c = [1; 2; 3; 4]$ (tj. $\mu_1 = -4$, $\mu_2 = -1$, $\mu_3 = 2$, $\mu_4 = 3$, $\beta_1 = 0, 1$, $\beta_2 = 0, 2$, $\beta_3 = 0, 3$, $\beta_4 = 0, 4$ i $c_1 = 1$, $c_2 = 2$, $c_3 = 3$, $c_4 = 4$).

Korišćenjem CharFunTool, funkcija gustine (PDF) i funkcija raspodjele (CDF) mogu se izračunati u proizvoljnim (unaprijed zadatim) tačkama y i kvantilna funkcija (QF) može se izračunati za proizvoljne (unaprijed zadate) vjerovatnoće $prob$, primjenom koda sa slike 4.1. Vrijednosti za funkcije raspodjele i gustine slučajne veličine Y izračunate u zadatim tačkama y prikazane su u tabeli B.1 iz dodatka B.

S druge strane, u istim tačkama y izračunali smo vrijednosti za funkcije raspodjele i gustine slučajne veličine Y primjenom algoritma za Foksovu H funkciju. Dobijeni rezultati su prikazani u tabeli B.2 iz dodatka B. Vrijeme računanja za sve funkcije raspodjele je 21,8281 s, dok za računanje funkcije gustine je utrošeno 19,3906 s. Za kalkulacije koristili smo parametar preciznosti $pg=10$. Upoređujući rezultate u tabelama

B.1 i B.2 iz dodatka B, označili smo niz različitih decimalnih cifara crvenom bojom. Možemo ponovo zaključiti da algoritam za numeričko računanje vrijednosti Foksove H funkcije ima visoku tačnost (u odnosu na zadati parametar prezicnosti). Za zadate vjerovatnoće

$$prob = [0, 80; 0, 85; 0, 90; 0, 925; 0, 95; 0, 975; 0, 99; 0, 995; 0, 999],$$

algoritam za invertovanje karakteristične funkcije izračunao je sljedeće kvantilne vrijednosti: $q_{0,80} = 14,713388739545$, $q_{0,85} = 15,371371147975$, $q_{0,90} = 16,226441137362$, $q_{0,925} = 16,796753149398$, $q_{0,95} = 17,564281554317$, $q_{0,975} = 18,809343228029$, $q_{0,99} = 20,377166657952$, $q_{0,995} = 21,529552051583$, $q_{0,999} = 24,151327981610$.

S druge strane, ako u istim vjerovatnoćama primijenimo algoritam za Foksovu H funkciju dobijamo sljedeće kvantilne vrijednosti: $q_{0,80} = 14,713388740514$, $q_{0,85} = 15,371371149811$, $q_{0,90} = 16,226441142073$, $q_{0,925} = 16,796753157383$, $q_{0,95} = 17,564281572793$, $q_{0,975} = 18,809343304232$, $q_{0,99} = 20,377167144685$, $q_{0,995} = 21,529554015765$, $q_{0,999} = 24,151377527702$. Dakle, oba algoritma daju iste kvantile vrijednosti.

Provjerimo još vrijeme koje je potrebno za izračunavanje kvantilnih vrijednosti pomoću algoritma za Foksovu H funkciju. Ovi rezultati su prikazani u tabeli 4.2.

Dakle, možemo zaključiti da algoritam za numeričko računanje Foksove H funkcije je veoma pouzdan u poređenju sa algoritmom za numeričku inverziju karakteristične funkcije. Sve navedeno upućuje da se algoritam za numeričko računanje Foksove H funkcije može efikasno koristiti u praksi kao što ćemo pokazati u glavi 6.

TABELA 4.2: Vrijeme izračunavanja, u sekundama, za približne vrijednosti za $P(Y \leq q)$ dobijene korišćenjem algoritma za Foksovu H funkciju, gdje q su kvantili.

p	q	vrijeme (s)
0,80	14,713388740514	3,26563
0,85	15,371371149811	3,17188
0,90	16,226441142073	2,95313
0,925	16,796753157383	3,46875
0,95	17,564281572793	3,48438
0,975	18,809343304232	3,67188
0,99	20,377167144685	3,89063
0,995	21,529554015765	3,78125
0,999	24,151377527702	3,60938

Glava 5

Numerička inverzija dvodimenzionalne logističke raspodjele

U ovoj glavi izložićemo rezultate rada [28], gdje je na osnovu dvodimenzionalne karakteristične funkcije izvedena raspodjela za dvodimenzionalnu logističku slučajnu veličinu.

5.1 Karakteristična funkcija dvodimenzionalne logističke raspodjele

Razmotrimo funkciju

$$\phi(t_1, t_2) = 2e^{i(t_1 m_1 + t_2 m_2)} B(1 - i\sigma_1 t_1, 1 - i\sigma_2 t_2, 1 + i\sigma_1 t_1 + i\sigma_2 t_2), \quad (5.1)$$

gdje je i imaginarna jedinica, $m_1, m_2, t_1, t_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$ i B je beta funkcija definisana sa (1.19) za $k = 3$.

Teorema 5.1. *Funkcija $\phi(t_1, t_2)$ definisana sa (5.1) je karakteristična funkcija.*

Dokaz. Primjenom Bochnerove teoreme dovoljeno je pokazati da je $\phi(t_1, t_2)$ pozitivno definitna i da je $\phi(0, 0) = 1$. Posljednji uslov lako provjeravamo.

Pokažimo pozitivnu definitnost. Dovoljno je pokazati da je

$$\sum_{j,k=1}^n c_j c_k e^{i b(j,k)} B\left(1 - i\sigma_1 a_1^{(j,k)}, 1 - i\sigma_2 a_2^{(j,k)}, 1 + i\left(\sigma_1 a_1^{(j,k)} + \sigma_2 a_2^{(j,k)}\right)\right) > 0, \quad (5.2)$$

gdje su $a_1^{(j,k)} = t_1^{(j)} - t_1^{(k)}$, $a_2^{(j,k)} = t_2^{(j)} - t_2^{(k)}$, $b(j,k) = a_1^{(j,k)}m_1 + a_2^{(j,k)}m_2$ i c_j, c_k realni brojevi za svako $j, k = 1, \dots, n$.

Primjenom (1.16) i (1.19) izraz (5.2) svodimo na

$$\begin{aligned} \sum_{j,k=1}^n c_j c_k e^{i b(j,k)} G(j, k) &= \sum_{j,k=1}^n c_j c_k e^{i b(j,k)} I(j, k) = \\ &= \iiint_{\mathbb{R}_+} \sum_{j,k=1}^n c_j c_k e^{i b(j,k)} x^{-i\sigma_1 a_1^{(j,k)}} s^{-i\sigma_2 a_2^{(j,k)}} u^{i\sigma_1 a_1^{(j,k)} + i\sigma_2 a_2^{(j,k)}} e^{-x-s-u} dx ds du = \\ &= \iiint_{\mathbb{R}_+} \left| \sum_{j=1}^n c_j e^{i(m_1 t_1^{(j)} + m_2 t_2^{(j)})} x^{-i\sigma_1 t_1^{(j)}} s^{-i\sigma_2 t_2^{(j)}} u^{i(\sigma_1 t_1^{(j)} + \sigma_2 t_2^{(j)})} \right|^2 e^{-x-s-u} dx ds du \geq 0, \end{aligned}$$

gdje su

$$\begin{aligned} G(j, k) &= \Gamma(1 - i\sigma_1 a_1^{(j,k)}) \Gamma(1 - i\sigma_2 a_2^{(j,k)}) \Gamma(1 + i\sigma_1 a_1^{(j,k)} + i\sigma_2 a_2^{(j,k)}), \\ I(j, k) &= \int_0^{+\infty} x^{-i\sigma_1 a_1^{(j,k)}} e^{-x} dx \int_0^{+\infty} s^{-i\sigma_2 a_2^{(j,k)}} e^{-s} ds \int_0^{+\infty} u^{i\sigma_1 a_1^{(j,k)} + i\sigma_2 a_2^{(j,k)}} e^{-u} du \end{aligned}$$

i $\mathbb{R}_+ = [0, +\infty)$. □

Dakle, sada je očigledano da funkcija definisana sa (5.1) je karakteristična funkcija.

Kako je $\phi(t_1, 0) = e^{it_1 m_1} B(1 - i\sigma_1 t_1, 1 + i\sigma_1 t_1)$ i $\phi(0, t_2) = e^{it_2 m_2} B(1 - i\sigma_2 t_2, 1 + i\sigma_2 t_2)$, odnosno primjenom teoreme (1.15) $\phi(t_1, 0) = e^{it_1 m_1} \Gamma(1 - i\sigma_1 t_1) \Gamma(1 + i\sigma_1 t_1)$ i $\phi(0, t_2) = e^{it_2 m_2} \Gamma(1 - i\sigma_2 t_2) \Gamma(1 + i\sigma_2 t_2)$. Uočavamo iz (1.53) da se marginalne karakteristične funkcije poklapaju sa karakterističnom funkcijom logističke raspodjele sa parametrima m_1, σ_1 i m_2, σ_2 , respektivno.

Definicija 5.1. *Raspodjela slučajnog vektora (X, Y) čija je funkcija raspodjele data sa (5.1) je dvodimenzionalna logistička raspodjela sa parametrima $m_1, m_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$ ili kraće*

$$(X, Y) \sim \text{BivLogistic}(m_1, m_2, \sigma_1, \sigma_2).$$

Štaviše, marginalne raspodjele vektora (X, Y) imaju logističku raspodjelu, tj. $X \sim \text{Logistic}(m_1, \sigma_1)$ i $Y \sim \text{Logistic}(m_2, \sigma_2)$.

U sljedećoj lemi izvešćemo mješoviti moment slučajnog vektora (X, Y) .

Lema 5.1. *Neka je (X, Y) slučajni vektor sa $BivLogistic(m_1, m_2, \sigma_1, \sigma_2)$ raspodjelom. Tada važi da je*

$$\mathbb{E}(XY) = m_1 m_2 + \sigma_1 \sigma_2 \frac{\pi^2}{6}. \quad (5.3)$$

Dokaz. Neka je $\phi(t_1, t_2)$ dvodimenzionalna karakteristična funkcija slučajnog vektora (X, Y) , lako pokazujemo da je

$$\mathbb{E}(XY) = \frac{1}{i^2} \frac{\partial^2 \phi(t_1, t_2)}{\partial t_1 \partial t_2} \Big|_{t_1=0, t_2=0}. \quad (5.4)$$

Nakon zamjene izraza (5.1) u (5.4), dobijamo da je

$$\begin{aligned} \frac{\partial^2 \phi(t_1, t_2)}{\partial t_1 \partial t_2} = & -m_1 m_2 e^{ia(t_1, t_2)} \{ \Gamma(1 - i\sigma_1 t_1) \Gamma(1 - i\sigma_2 t_2) \Gamma(b(t_1, t_2)) + \\ & -m_1 \sigma_2 \Gamma(1 - i\sigma_1 t_1) \Gamma(1 - i\sigma_2 t_2) \Gamma(b(t_1, t_2)) [\psi(b(t_1, t_2)) - \psi(1 - i\sigma_2 t_2)] + \\ & -m_2 \sigma_1 \Gamma(1 - i\sigma_1 t_1) \Gamma(1 - i\sigma_2 t_2) \Gamma(b(t_1, t_2)) [\psi(b(t_1, t_2)) - \psi(1 - i\sigma_1 t_1)] + \\ & -\sigma_1 \sigma_2 \Gamma(1 - i\sigma_1 t_1) \Gamma(1 - i\sigma_2 t_2) \Gamma(b(t_1, t_2)) \psi(1 - i\sigma_1 t_1) \times \\ & [\psi(1 - i\sigma_2 t_2) - \psi(b(t_1, t_2))] + i\sigma_1 \Gamma(1 - i\sigma_1 t_1) \times \\ & [-i\sigma_2 \Gamma(1 - i\sigma_2 t_2) \Gamma(b(t_1, t_2)) \psi(1 - i\sigma_2 t_2) \psi(b(t_1, t_2)) + \\ & i\sigma_2 \Gamma(1 - i\sigma_2 t_2) \Gamma(1 + i\sigma_1 t_1 [\psi^2(b(t_1, t_2)) + \psi'(b(t_1, t_2))])] \} , \end{aligned}$$

gdje su $a(t_1, t_2) = m_1 t_1 + m_2 t_2$, $b(t_1, t_2) = 1 + i\sigma_1 t_1 + i\sigma_2 t_2$, ψ je digama funkcija data u definiciji 1.29 i ψ' prvi izvod digama funkcije.

Dakle, ako u posljednju jednačinu uvrstimo $t_1 = t_2 = 0$ zamjenom dobijenog izraza u (5.4) dobijamo (5.3). \square

Sada možemo da izračunamo Pirsonov koeficijent korelacije za slučajni vektor (X, Y) iz definicije 5.1.

Teorema 5.2. *Neka je $(X, Y) \sim BivLogistic(m_1, m_2, \sigma_1, \sigma_2)$. Tada je*

$$\rho_{X,Y} = \frac{1}{2},$$

gdje je $\rho_{X,Y}$ Pirsonov koeficijent korelacije između slučajnih veličina X i Y .

Dokaz. Iz definicije 5.1 slijedi da slučajne veličine X i Y imaju logističku raspodjelu, pa je $\mathbb{E}X = m_1$, $\mathbb{E}Y = m_2$, $\mathbb{D}(X) = \frac{\pi^2 \sigma_1^2}{3}$ i $\mathbb{D}(Y) = \frac{\pi^2 \sigma_2^2}{3}$. Primjenom leme 5.1 i prethodnih rezultata i zamjenom u (1.2) dolazimo do tvrđenja teoreme. \square

Iz svega navedenog lako dolazimo do sljedećeg rezultata.

Posljedica 5.1. *Neka je $(X, Y) \sim \text{BivLogistic}(m_1, m_2, \sigma_1, \sigma_2)$. Tada je*

$$C = \begin{pmatrix} \frac{\pi^2 \sigma_1}{3} & \sigma_1 \sigma_2 \frac{\pi^2}{6} \\ \sigma_1 \sigma_2 \frac{\pi^2}{6} & \frac{\pi^2 \sigma_2}{3} \end{pmatrix}, \quad (5.5)$$

gdje je C matrica kovarijacije slučajnog vektora (X, Y) .

5.2 Uslovna karakteristična funkcija

Razmotrimo slučajni vektor (X, Y) iz definicije 5.1. Glavni cilj u ovom poglavlju je da izvedemo uslovnu karakterističnu funkciju za slučajnu veličinu $X|Y$.

Teorema 5.3. *Neka je slučajni vektor $(X, Y) \sim \text{BivLogistic}(m_1, m_2, \sigma_1, \sigma_2)$. Tada karakteristična funkcija slučajne veličine $X|Y$ je data sa*

$$\phi(u|y) = \frac{e^{im_1 u} \Gamma(1 - i\sigma_1 u) \Gamma(2 + i\sigma_1 u)}{\left(1 + e^{-\frac{(y-m_2)}{\sigma_2}}\right)^{i\sigma_1 u}}. \quad (5.6)$$

Dokaz. Primjenom definicije za dvodimenzionalnu karakterističnu funkciju imamo da je

$$\phi(u, v) = \mathbb{E}e^{i(uX+vY)} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{i(ux+vy)} f(x, y) dx dy, \quad (5.7)$$

gdje je $f(x, y)$ dvodimenzionalna funkcija gustine vektora (X, Y) .

Razmotrimo relaciju za uslovnu funkciju gustine, tj.

$$f(x, y) = g(x|y)h(y),$$

gdje je $g(x|y)$ funkcija gustine slučajne veličine $X|Y$, a $h(y)$ je funkcije gustine slučajne veličine Y . Dalje je

$$\phi(u|y) = \int_{-\infty}^{+\infty} e^{iux} g(x|y) dx,$$

gdje je $\phi(u|y)$ uslovna karakteristična funkcija od slučajne veličine $X|Y$.

Sada jednakost (5.7) možemo zapisati u obliku

$$\phi(u, v) = \int_{-\infty}^{+\infty} e^{ivy} h(y) \int_{-\infty}^{+\infty} e^{iux} g(x|y) dx dy = \int_{-\infty}^{+\infty} e^{ivy} \phi(u|y) h(y) dy.$$

Primjenom inverzne Furijeove transformacije (teorema 1.20) slijedi da je

$$\phi(u|y) h(y) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ivy} \phi(u, v) dv. \quad (5.8)$$

Neka je $\phi(u, v)$ karakteristična funkcija definisana u (5.1) i $h(y)$ je funkcija gustine slučajne veličine $Y \sim \text{Log}(m_2, \sigma_2)$.

Jednakost (5.8) glasi

$$\phi(u|y) = \frac{e^{ium_1} \Gamma(1 - i\sigma_1 u)}{2\pi g(y)} \int_{-\infty}^{+\infty} e^{iv(m_2 - y)} \Gamma(1 - i\sigma_2 v) \Gamma(1 + i\sigma_1 u + i\sigma_2 v) dv.$$

Gama funkciju možemo predstaviti kao Furijeovu transformaciju na sljedeći način

$$\mathcal{F} \left(e^{(1+i\sigma_1 u) \frac{t}{\sigma_2}} e^{-e^{\frac{t}{\sigma_2}}} ; w \right) = \frac{\sigma_2}{\sqrt{2\pi}} \Gamma(1 + i\sigma_1 u + i\sigma_2 w) \quad (5.9)$$

i

$$\mathcal{F} \left(e^{\frac{t-y+m_2}{\sigma_2}} e^{-e^{\frac{t-y+m_2}{\sigma_2}}} ; w \right) = \frac{\sigma_2}{\sqrt{2\pi}} e^{iw(y-m_2)} \Gamma(1 + i\sigma_2 w). \quad (5.10)$$

Koristeći Parsevalovu jednakost (teorema 1.21) za Furijeove transformacije date sa (5.9) i (5.10) dobijamo

$$\begin{aligned} & \frac{\sigma_2^2}{2\pi} \int_{-\infty}^{+\infty} e^{iv(m_2 - y)} \Gamma(1 - i\sigma_2 v) \Gamma(1 + i\sigma_1 u + i\sigma_2 v) dv = \\ & = \int_{-\infty}^{+\infty} e^{\frac{t-y+m_2}{\sigma_2}} e^{-e^{\frac{t-y+m_2}{\sigma_2}}} e^{(1+i\sigma_1 u) \frac{t}{\sigma_2}} e^{-e^{\frac{t}{\sigma_2}}} dt \end{aligned} \quad (5.11)$$

Sada, riješićemo integral na desnoj strani jednakosti (5.11). Smjenom $e^{\frac{t}{\sigma_2}} = s$ po-
menuti integral svodimo na

$$\int_{-\infty}^{+\infty} e^{\frac{t-y+m_2}{\sigma_2}} e^{-e^{\frac{t-y+m_2}{\sigma_2}}} e^{(1+i\sigma_1 u)\frac{t}{\sigma_2}} e^{-e^{\frac{t}{\sigma_2}}} dt = \sigma_2 \int_0^{+\infty} e^{-s(1+e^{\frac{m_2-y}{\sigma_2}})} s^{1+i\sigma_1 u} e^{\frac{m_2-y}{\sigma_2}} ds.$$

Integral u posljednoj jednakosti rješavamo koristeći smjenu $s(1 + e^{\frac{m_2-y}{\sigma_2}}) = h$

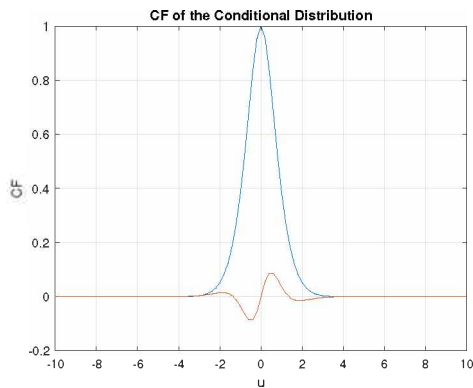
$$\sigma_2 e^{\frac{m_2-y}{\sigma_2}} \int_0^{+\infty} e^{-h} \frac{h^{1+i\sigma_1 u}}{\left(1 + e^{\frac{m_2-y}{\sigma_2}}\right)^{2+i\sigma_1 u}} dh = \frac{\sigma_2 \Gamma(2 + i\sigma_1 u) e^{\frac{m_2-y}{\sigma_2}}}{\left(1 + e^{\frac{m_2-y}{\sigma_2}}\right)^{2+i\sigma_1 u}}.$$

Odnosno, nakon sređivanja dobijamo

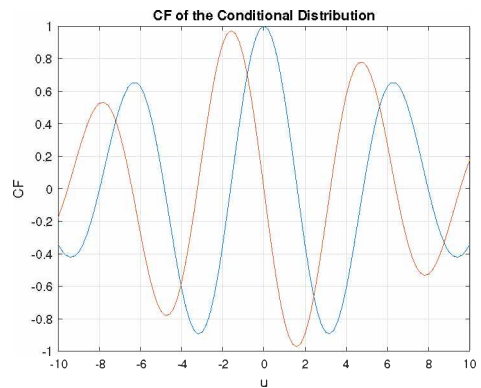
$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{iv(m_2-y)} \Gamma(1 - i\sigma_2 v) \Gamma(1 + i\sigma_1 u + i\sigma_2 v) dv = \frac{\Gamma(2 + i\sigma_1 u) e^{\frac{m_2-y}{\sigma_2}}}{\sigma_2 \left(1 + e^{\frac{m_2-y}{\sigma_2}}\right)^{2+i\sigma_1 u}}.$$

Konačno, stižemo do izraza za uslovnu karakterističnu funkciju (5.6). □

Na slici 5.1 (A) prikazana je uslovna karakteristična funkcija za standardnu logističku
raspodjelu za $y = 0$, dok je slici 5.1 (B) karakteristična funkcija za raspodjelu $X|Y$,
gdje je $(X, Y) \sim BivLogistic(-1, 2, 0.1, 0.2)$ za $y = 2$.



(A) standardna logistička raspodjela
za $y = 0$



(B) logistička raspodjela sa
parametrima $m_1 = -1$, $m_2 = 2$,
 $\sigma_1 = 0.1$, $\sigma_2 = 0.2$ za $y = 2$

SLIKA 5.1: Karakterististična funkcija za uslovnu raspodjelu

5.3 Algoritam za numeričku inverziju dvodimenzionalne karakteristične funkcije

Prvo, navešćemo rezultate koji su dobijeni u radu [42]. Pretpostavimo da ϕ je poznata karakteristična funkcija i neka je F njoj odgovarajuća funkcija raspodjele. Pokazano je da u jednodimenzionalnom slučaju za funkcije F i ϕ važi sljedeća relacija

$$F(x) = \frac{1}{2} - \frac{1}{2\pi} \int_0^{+\infty} \left(\Delta_t \left[\frac{\phi(t)e^{-itx}}{it} \right] \right) dt, \quad (5.12)$$

gdje je $\Delta_t[y(t)] = y(t) + y(-t)$.

U radu [17] izvedena je alternativna formula za inverziju zasnovana samo na integraciji realne funkcije. Funkcija gustine u tom slučaju data je sa

$$f(x) = \frac{1}{\pi} \int_0^{+\infty} \operatorname{Re}(e^{-itx} \phi(t)) dt \approx \frac{1}{\pi} h_1 \sum_{v=0}^{+\infty} \operatorname{Re}(e^{-ih_1(v+0.5)x} \phi(h_1(v+0.5))), \quad (5.13)$$

za svako $h_1 > 0$.

Ako pretpostavimo da je x tačka neprekidnosti funkcije raspodjele F , tada imamo da je

$$F(x) = \frac{1}{2} - \frac{1}{2\pi} \int_0^{+\infty} \operatorname{Im} \left(\frac{e^{-itx} \phi(t)}{t} \right) dt \approx \frac{1}{2} - \frac{1}{2\pi} h_1 \sum_{v=0}^{+\infty} \operatorname{Im} \left(\frac{e^{-ih_1(v+0.5)x} \phi(h_1(v+0.5))}{h_1(v+0.5)} \right), \quad (5.14)$$

za svako $h_1 > 0$. Razmotrimo sada raspodjelu slučajnog vektora $X = (X_1, X_2)$, koja je specifikovana svojom karakterističnom funkcijom $\phi(t_1, t_2)$. U radu [42] data je sljedeća relacija

$$F(x_1, x_2) = \frac{F(x_1) + F(x_2)}{2} + \frac{1}{4\pi^2} \int_0^{+\infty} \int_0^{+\infty} \left(\Delta_{t_1} \cdot \Delta_{t_2} \left[\frac{\phi(t_1, t_2) e^{-i(t_1 x_1 + t_2 x_2)}}{it_1 it_2} \right] \right) dt_1 dt_2 - \frac{1}{4}, \quad (5.15)$$

gdje $F(x_1)$ i $F(x_2)$ su marginalne funkcije raspodjele date sa (5.12).

Korišćenjem teoreme 2 i posljedice 2 iz rada [42], posljednja relacija se može predstaviti kao

$$\begin{aligned}
F(x_1, x_2) &= \frac{F(x_1) + F(x_2)}{2} - \frac{1}{2\pi^2} \int_0^{+\infty} \int_{-\infty}^{+\infty} \operatorname{Re} \left(\frac{\phi(t_1, t_2) e^{-i(t_1 x_1 + t_2 x_2)}}{t_1 t_2} \right) dt_1 dt_2 - \frac{1}{4} \approx \\
&\approx \frac{F(x_1) + F(x_2)}{2} - \frac{1}{2\pi^2} h_1 h_2 \sum_{v_1=0}^{+\infty} \sum_{v_2=-\infty}^{+\infty} \operatorname{Re} \left(\frac{\phi(b_1, b_2) e^{-i(b_1 x_1 + b_2 x_2)}}{b_1 b_2} \right) - \frac{1}{4},
\end{aligned} \tag{5.16}$$

gdje $F(x_1)$ i $F(x_2)$ su marginalne funkcije raspodjele date sa (5.14), i $b_1 = h_1(v_1 + 0.5)$, $b_2 = h_2(v_2 + 0.5)$ za svako $h_1, h_2 > 0$. Slično,

$$\begin{aligned}
f(x_1, x_2) &= \frac{1}{2\pi^2} \int_0^{+\infty} \int_{-\infty}^{+\infty} \operatorname{Re}(e^{-i(t_1 x_1 + t_2 x_2)} \phi(t_1, t_2)) dt_1 dt_2 \approx \\
&\approx \frac{1}{2\pi^2} h_1 h_2 \sum_{v_1=0}^{+\infty} \sum_{v_2=-\infty}^{+\infty} \operatorname{Re}(e^{-i(b_1 x_1 + b_2 x_2)} \phi(b_1, b_2)).
\end{aligned} \tag{5.17}$$

Algoritam je uspješno implementiran u MATLAB repozitorijumu CharFunTool (pogledati [55]). Kada smo dizajnirali algoritam za numeričku inverziju dvodimenzionalne karakteristične funkcije takođe smo izveli numerički algoritam za kopula funkciju. Ovo nam je omogućilo da računamo, na primjer, Pirsonov i Kendalov tau koeficijent korelacije.

Integrale u (5.16) i (5.17) smo računali aproksimativno, korišćenjem numeričke kvadrature. Ova procedura zahtjeva krišćenje Rimanovih suma na intervalima $(0, T) \times (-T, T)$, gdje $T = N \cdot dt$ je dovoljno velika gornja granica integracije u oblasti definisanosti raspodjele.

Ako optimalne vrijednosti za N i T nijesu poznate možemo prvo primijeniti dobro poznati metod za aproksimaciju pravilo šest sigmi. Primjenom tog metoda vrijednost za dt možemo izračunati na sljedeći način: $dt = 2\pi/(x_{max} - x_{min})$, gdje su $x_{max} = \text{mean}(X) + 6 \text{std}(X)$ i $x_{min} = \text{mean}(X) - 6 \text{std}(X)$, gdje $\text{mean}()$ očekivana vrijednost i $\text{std}()$ je standardna devijacija slučajnog vektora X (pretpostavljamo da postoje).

Na osnovu prethodnog, predlažemo sljedeću implementaciju algoritma za numeričku inverziju dvodimenzionalne karakteristične funkcije ϕ :

korak 1: Predstaviti karakterističnu funkciju ϕ slučajnog vektora (X_1, X_2) kao anonimnu funkciju ¹ sa dvodimenzionalnim ulaznim parametrom t , tj. kao funkciju koja računa vrijednosti karakteristične funkcije za proizvoljno $t = (t_1, t_2)$. Nakon toga, pomoću karakteristične funkcije ϕ odrediti marginalne karakteristične funkcije ϕ_1 i ϕ_2 kao anonimne jednodimenzionalne funkcije sa ulaznim parametrom t_1 i t_2 , respektivno $(\phi_1(t_1) = \phi([t_1, 0])$ i $\phi_2(t_2) = \phi([0, t_2])$);

korak 2: Specifikovati matricu dimenzije $(N \times 2)$ za vrijednosti x . Funkcija raspodjele i funkcija gustine date su na lijevoj strani u jednakostima (5.16) i (5.17). Izračunati njihove vrijednosti u svim elementima specifikovane matrice. Alternativno, možemo specifikovati niz dimenzije (1×2) , gdje prva koordinata $x\{1\}$ odgovara vektoru sa x_1 vrijednostima i druga koordinata $x\{2\}$ je vektor koji sadrži vrijednosti od x_2 . Potom ćemo konstruisati mrežu koja sadrži vrijednosti za x pomoću ugrađene funkcije u Matlabu `meshgrid` na sljedeći način: $x = \text{meshgrid}(x\{1\}, x\{2\})$;

korak 3: Izračunati matematičko očekivanje i standardnu devijaciju za slučajni vektor $X = (X_1, X_2)$ pomoću marginalnih karakterističnih funkcija. Korišćenjem metoda konačnih razlika za numeričko diferenciranje dobili smo sljedeće ocjene

$$\begin{aligned}\mathbb{E}(X_1) &= \left(\frac{8\text{Im}(\phi_1(\delta))}{5\delta} - \frac{2\text{Im}(\phi_1(2\delta))}{5\delta} + \frac{8\text{Im}(\phi_1(3\delta))}{105\delta} - \frac{2\text{Im}(\phi_1(4\delta))}{280\delta} \right), \\ \mathbb{E}(X_1^2) &= \left(\frac{205}{72\delta^2} - \frac{16\text{Re}(\phi_1(\delta))}{5\delta^2} + \frac{2\text{Re}(\phi_1(2\delta))}{5\delta^2} - \frac{16\text{Re}(\phi_1(3\delta))}{315\delta^2} + \frac{2\text{Re}(\phi_1(4\delta))}{560\delta^2} \right), \\ \sigma(X_1) &= \sqrt{\mathbb{E}(X_1^2) - \mathbb{E}^2(X_1)}\end{aligned}$$

gdje je δ odabrani nivo tolerancije za numeričko diferenciranje (kao zadanu vrijednost koristili smo $\delta = 10^{-4}$). Slično, računamo $\mathbb{E}(X_2)$ i $\sigma(X_2)$ tako da je $\mathbb{E}(X) = [\mathbb{E}(X_1), \mathbb{E}(X_2)]$ i $\sigma(X) = [\sigma(X_1), \sigma(X_2)]$;

korak 4: Primijeniti pravilo šest sigmi i izračunati dt kao što je opisano iznad. Zatim odrediti vrijednosti za t_1, t_2, t u kojima će se računati dvodimenzionalna karakteristična funkcija kao i njene marginalne funkcije. Dobijene rezultate čuvamo kao `cft1, cft2` i `cft` kao što je i prikazano Matlabovim kodom:

¹Anonimna funkcija je jednostavna, jednolinijska funkcija koja ne mora da se čuva u posebnoj datoteci. Anonimne funkcije se mogu kreirati u komandnom prozoru ili u bilo kojim skriptama ili funkciji koju definiše korisnik. Sintaksa za anonimnu funkciju je sljedeća: `fnhandlevar = @(arguments) functionbody`.

```

t1 = (0.5+(0:N))'*dt(1);
t2 = (0.5+(0:N))'*dt(2);
t3 = (0.5+(-N:N))'*dt(2);
cft1 = cf([t1 0*t1]);
cft2 = cf([0*t2 t2]);
[tt1,tt2] = meshgrid(t1,t3);
t = [tt1(:) tt2(:)];
cft = cf(t);

```

korak 5: U specifikovanim vrijednostima x_1 i x_2 odrediti marginalne funkcije gustine i marginalne funkcije raspodjele slučajnog vektora (X_1, X_2) . Koristiti jednakosti (5.13) i (5.14) kao i objašnjenje iz algoritma 1;

Algoritam 1 Određivanje marginalnih funkcija gustine (pdf1) i raspodjele (cdf1)

```

1: E1= exp(-i * x1 * t1')
2: pdf1 = real(E1 * cft1) * dt(1) /pi
3: cdf1 = 0.5 - (imag(E1 * cft1 ./ t1) * dt(1)/ pi)

```

korak 6: Odrediti dvodimenzionalne funkcije raspodjele (cdf) i gustine (pdf) kao što je opisano u algoritmu 2, tj. korišćenjem jednakosti (5.16) i (5.17);

Algoritam 2 Određivanje dvodimenzionalnih funkcija raspodjele (cdf) i gustine (pdf)

```

1: c = dt(1) * dt(2) / (2* pi*pi)
2: E = exp(-i* x * t')
3: pdf = c * real(E * cft)
4: if ~ isMeshed then
5:   [f1,f2] = meshgrid(cdf1,cdf2)
6:   cdf = (f1 + f2)/2 - 0.25
7: else
8:   cdf = (cdf1 + cdf2)/2 - 0.25
9: end if
10: cdf = cdf(:)
11: cftt = cft ./ t(:,1) ./ t(:,2)
12: cdf = cdf - c * real(E*cftt)

```

korak 7: Odrediti interpolante od dvodimenzionalnih funkcija gustine i raspodjele korišćenjem dvodimenzionalne baricentrične interpolacije.

Algoritam `InterpBarycentric2D` vrši interpolaciju proizvoljne dvodimenzionalne funkcije, na osnovu vrijednosti funkcije koje su date $n \times m$ dimenzionalnom matricom, recimo `funOld`. Matricu (`funOld`) smo računali za sve kombinacije početnog

niza tačaka x (**xOld**) i y (**yOld**). Algoritam koristi dekompoziciju matrice **funOld** na singularne vrijednosti (SVD). Rezultat je faktORIZACIJA ove matrice na tri matrice tako da $\mathbf{funOld}_{n \times m} = U_{n \times n} S_{n \times m} V_{m \times m}^T$, gdje su U i V ortogonalne matrice, i T je oznaka za transponovanje matrica. Algoritam izvršava baricentričnu interpolaciju kolona matrica U i V . Rezultat je $N \times M$ dimenzionalna matrica **funNew** koja sadrži interpolirane vrijednosti ulazne funkcije, koju smo izračunali za sve kombinacije novih koordinata **xNew** i **yNew**. Alternativno, rezultat može biti vektor koji sadrži vrijednosti od **funNew** čija dužina zavisi od broja parova u navedenim kombinacijama $\mathbf{xyNew} = \{\mathbf{xNew}, \mathbf{yNew}\}$;

korak 8: Odrediti interpolante za marginalne kvantilne funkcije (**QF1** **QF2**) na osnovu kojih dobijamo odgovarajuće kvantile za bilo koju vjerovatnoću iz intervala $(0, 1)$. **QF** je izračunata iz preračunatih (poznatih) vrijednosti **xGiven** i **cdfGiven**, gdje je **xGiven** vektor koji sadrži vrijednosti od x u kojima smo preračunali funkciju raspodjele i **cdfGiven** je vektor koji sadrži vrijednosti od funkcije raspodjele u tačkama **xGiven**. Evaluacija je zasnovana na baricentričnoj interpolaciji;

korak 9: Generisati kopula funkciju raspodjele (**CopulaCdf**) i kopula funkciju gustine (**CopulaPdf**) za date $u_1, u_2 \in (0, 1)$. Prvo, u koraku 8, odrediti kvantilne funkcije **QF1** i **QF2** za marginalne raspodjele u tačkama u_1 i u_2 . Koristeći posljedicu 1.1, kopulu funkciju raspodjele možemo definisati kao

$$\text{CopulaCdf}(u_1, u_2) = F(\text{QF1}(u_1), \text{QF2}(u_2)),$$

gdje je F dvodimenzionalna funkcija raspodjele slučajnog vektora X . Zatim, u koraku 7, pomoću baricentrične interpolacije generisati ineterpolant kopula funkcije raspodjele. U algoritmu 3 prikazani su najznačajni koraci. **InterpBarycentric2D** je funkcija kreirana u koraku 7 koja interpolira kopula funkciju raspodjele, na osnovu $n \times m$ dimenzionalne matrice koja sadrži vrijednosti od dvodimenzionalne funkcije raspodjele **Zcdf** (predstavljene kao **funOld** u koraku 7). Dvodimenzionalna funkcija raspodjele se izračunava u svim kombinacijama od x_1 i x_2 . Algoritam 3 izračunava dvodimenzionalnu funkciju raspodjele u novim tačkama **QF1**(u_1) i **QF2**(u_2). Slično, ako u **InterpBarycentric2D** koristimo dvodimenzionalnu funkciju gustine od slučajnog vektora X , možemo generisati kopula funkciju gustine.

5.4 Algoritam za generisanje slučajnih brojeva

U ovom poglavlju opisaćemo postupak za generisanje slučajnih brojeva iz dvodimenzionalne slučajne veličine koja je specifikovana svojom karakterističnom funkcijom.

Algoritam 3 Kopula funkcija raspodjele

```
1: if iscell(u12) then
2:   u1 = u12{1}
3:   u2 = u12{2}
4:   x1New = QF1(u1)
5:   x2New = QF2(u2)
6:   x12New = {x1New, x2New}
7: else
8:   x12New = [QF1(u12(:, 1)), QF2(u12(:, 2))]
9: end if
10: CopulaCdf = InterpBarycentric2D(x1, x2, Zcdf, x12New)
```

Koristili smo metod koji se zasniva na korišćenju uslovne karakteristične funkcije, što se prvi put javlja u literaturi, što je detaljno opisano u radu [28].

Opis algoritma: Algoritam se zasniva na korišćenju marginalne i uslovne karakteristične funkcije (koje su poznate ili izvedene iz dvodimenzionalne karakteristične funkcije). Potom se aproksimiraju tražene kvantilne vrijednosti (vrijednosti uzorka) iz raspodjele korišćenjem odgovarajućih alogoritama za interpolaciju. Tačnije, `interp2` se korisiti za brzu dvodimenzionalnu interpolaciju i `interp1` ili više preciznija barocentrična inverpolacija (`InterpBarycentric`) se koriste za jednodimenzionalnu interpolaciju.

Prvo deklariramo ulazne vrijdnosti:

- `nPts` je broj tačaka korišćen za uspostavljanje inicijalnih vrijednosti za vjerovatnoću koje su izvedene korišćenjem Čebišovljevih tačaka sa intervala $[0, 1]$. Ako nije navedeno, tada je `nPts` = 101.
- `nTrapRule` je broj tačaka korišćen za trapezoidnu integraciju. Ako nije navedeno, onda je `nTrapRule` = 2^9 .
- `chebyPts` je broj Čebišoljevih tačaka korišćen u kalkulacijama za kreiranje vektora x , $x \in [x_{min}, x_{max}]$, u čijim vrijednostima računamo funkciju raspodjele koja je specifikovana svojom karakterističnom funkcijom. Ako nije navedeno, onda je `chebyPts` = 2^8 .
- `SixSigmaRule` je množitelj standardne devijacije kojim definišemo pravilo za ocjenjivanje domena definisanosti raspodjele. Ako nije navedeno, uzimamo da je `SixSigmaRule` = 6.

Dakle, imamo sljedeće korake.

1. Generisati `nPts` Čebišovljeve tačke (druge vrste) u datom intervalu (domen = $[0, 1]$), zajedno sa vektorom vrsta čiji elementi su težine za Clenshaw-Curtis kvadraturu (preko datog intervala $[0, 1]$).

2. Za određivanje kvantilnih vrijednosti marginalne slučajne veličine X_2 (neprekidna, jednodimenzionalna raspodjela) koja je specifikovana preko svoje karakteristične funkcije za integraciju koristiti Gil - Pelaez formulu za inverziju (5.14) sa Rimanovom sumom. Kvantile slučajne veličine X_2 računamo u vektoru čiji elementi su Čebišovljeve tačke od $[xMin, xMax]$ obima `chebyPts`.
3. Izvesti uslovnu karakterističnu funkciju iz dvodimenzionalne karakteristične funkcije, tj. karakterističnu funkciju slučajne veličine X_1 pod uslovom $X_2 = x_2$. U specijalnom slučaju, eksplicitni oblik uslovne karakteristične funkcije je poznat, u suprotnom uslovna karakteristična funkcija može se izračunati numerički u određenim vrijednostima njenog argumenta t , kao što je predloženo u [8]. Koristiti brzu Gil-Pelaez inverziju (5.14) sa Rimanovom sumom za izračunavanje kvantila od uslovne raspodjele $X_1|X_2 = x_2$.
4. Koristiti ugrađene funkcije u Matlabu `interp1` i `interp2` za brzu jednodimenzionalnu i dvodimenzionalnu linearnu interpolaciju i generisati slučajni vektor.

```
X2 = interp1(prOld,x_2,pNew(:,2));
X1 = interp2(p1Old,p2Old,x_1,pNew(:,1),pNew(:,2));
```

gdje je `pNew = rand(N,2)` i `prOld` predstavlja Čebišovljeve tačke generisane u prvom koraku, `[p1Old,p2Old] = meshgrid(prOld)`. Ovdje x_1 je matrica kvantilnih vrijednosti za uslovnu raspodjelu slučajne veličine X_1 , koja je izračunata u svim vrijednostima `prOld`, pod uslovom za sve vrijednosti iz $x_2 = QF2(prOld)$.

5. Na kraju, možemo izračunati Pirsonov koeficijent (ρ) kao $\rho = \text{corr}(X1, X2)$;

5.5 Kopula funkcija, Spirmanov ro i Kendalov tau za dvodimenzionalnu logističku raspodjelu

U ovom poglavlju prikazaćemo kopula funkciju koja odgovara dvodimenzionalnoj logističkoj raspodjeli koja je data preko svoje karakteristične funkcije (5.1).

U tabelama 5.1 i 5.2 prikazane su srednje vrijednosti za Spirmanov ro, Kendalov tau i Pirsonov koeficijent korelacije. Generisali smo 1000 uzoraka and izračunali smo srednje vrijednosti korelacionih koeficijenata zajedno sa njihovom standardnim greškama. Možemo uočiti da vrijednosti za standardne greške u tabelama 5.1 i 5.2 su veoma male. Ovo sugerise da je naš algoritam za generisanje slučajnih brojeva veoma precizan.

TABELA 5.1: Srednje vrijednosti za Kendalov tau, Spirmanov ρ i Pirsonov koeficijent korelacije u slučaju dvodimenzionalne logističke raspodjele sa parametrima $m = [1; 3]$ i $\sigma = [0, 1; 0, 3]$

N	Pirsonov koeficijent korelacije		Kendalov τ koeficijent		Spirmanov ρ koeficijent	
	\bar{x}	se	\bar{x}	se	\bar{x}	se
50	0,4896	$5,5015e - 04$	0,3302	$4,0342e - 04$	0,4660	$5,3125e - 04$
100	0,4970	$2,7323e - 04$	0,3338	$1,9624e - 04$	0,4745	$2,6168e - 04$
500	0,4983	$5,7171e - 05$	0,3318	$3,9233e - 05$	0,4755	$5,2646e - 05$
1000	0,4993	$2,9782e - 05$	0,3329	$2,0279e - 05$	0,4774	$2,7115e - 05$
2000	0,5000	$1,4399e - 05$	0,3337	$9,5740e - 06$	0,4787	$1,2818e - 05$
5000	0,4997	$5,8147e - 06$	0,3331	$3,9854e - 06$	0,4780	$5,3636e - 06$
10000	0,5001	$2,9187e - 06$	0,3334	$1,9570e - 06$	0,4784	$2,6235e - 06$

TABELA 5.2: Srednje vrijednosti za Kendalov tau, Spirmanov ρ i Pirsonov koeficijent korelacije u slučaju dvodimenzionalne logističke raspodjele sa parametrima $m = [-1; 2]$ i $\sigma = [0, 5; 0, 8]$

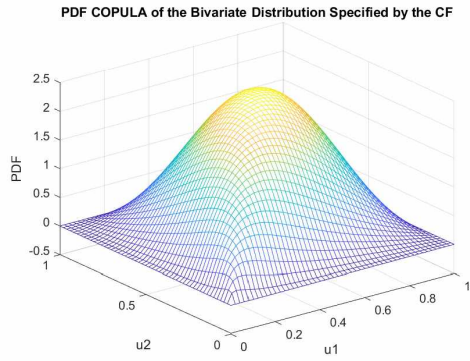
N	Pirsonov koeficijent korelacije		Kendalov τ koeficijent		Spirmanov ρ koeficijent	
	\bar{x}	se	\bar{x}	se	\bar{x}	se
50	0,4887	$5,5030e - 04$	0,3294	$4,0163e - 04$	0,4649	$5,2974e - 04$
100	0,4973	$2,7295e - 04$	0,3338	$1,9565e - 04$	0,4747	$2,6118e - 04$
500	0,4983	$5,7142e - 05$	0,3318	$3,9228e - 05$	0,4755	$5,2652e - 05$
1000	0,4990	$2,8935e - 05$	0,3330	$1,9725e - 05$	0,4776	$2,6229e - 05$
2000	0,5013	$1,4262e - 05$	0,3339	$9,9633e - 06$	0,4790	$1,3370e - 05$
5000	0,4999	$5,5453e - 06$	0,3334	$3,8690e - 06$	0,4784	$5,2173e - 06$
10000	0,5000	$2,9076e - 06$	0,3334	$1,9331e - 06$	0,4784	$2,5876e - 06$

Dokazali smo da je Pirsonov koeficijent korelacije konstantan za bilo koju vrijednost parametara (teorema 5.2). Iz pomenutih tabela dobijenih na osnovu numeričkih simulacija, možemo zaključiti da Spirmanov ρ i Kendalov τ imaju takođe konstantne vrijednosti, tj. ne zavise od vrijednosti parametara dvodimenzionalne logističke raspodjele.

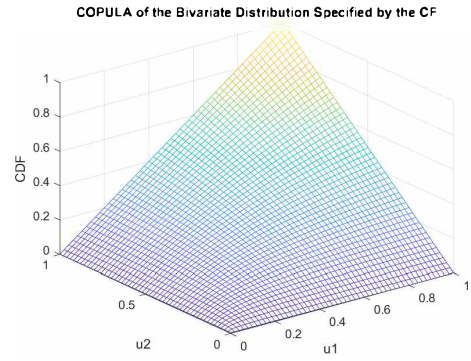
Na slikama 5.2–5.5 prikazane su kopula funkcija raspodjele i kopula funkcija gustine dobijene pomoću algoritma 3.

Korišćenjem algoritma za numeričku inverziju karakteristične funkcije koji smo opisali u poglavlju 4.2 možemo generisati funkciju gustine, funkciju raspodjele za dvodimenzionalnu logističku raspodjelu koja je specifikovana karakterističnom funkcijom (2.4).

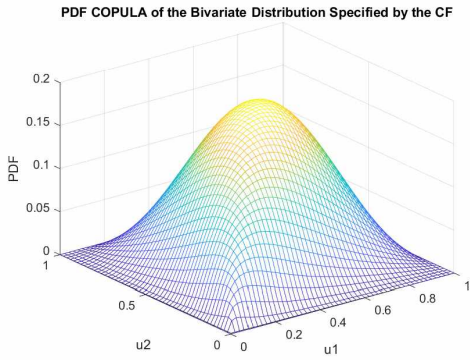
Na slikama (5.6) - (5.7) prikazane su kopula funkcija gustine i kopula funkcija raspodjele za dvodimenzionalnu logističku raspodjelu sa parametrima $m_1 = m_2 = 0$,



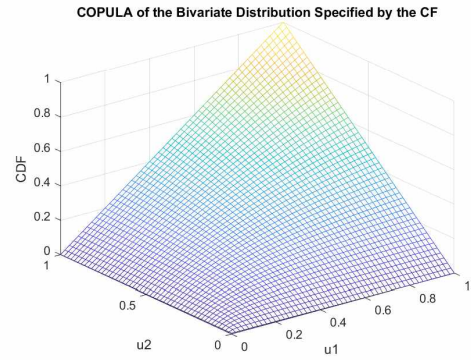
SLIKA 5.2: Kopula funkcija gustine dvodimenzionalne logističke raspodjele sa parametrima $m = [1; 3]$ i $\sigma = [0, 1; 0, 3]$



SLIKA 5.3: Kopula funkcija dvodimenzionalne logističke raspodjele sa parametrima $m = [1; 3]$ i $\sigma = [0, 1; 0, 3]$

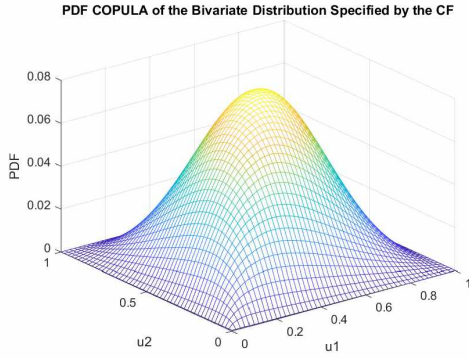


SLIKA 5.4: Kopula funkcija gustine dvodimenzionalne logističke raspodjele sa parametrima $m = [-1; 2]$ i $\sigma = [0, 5; 0, 8]$

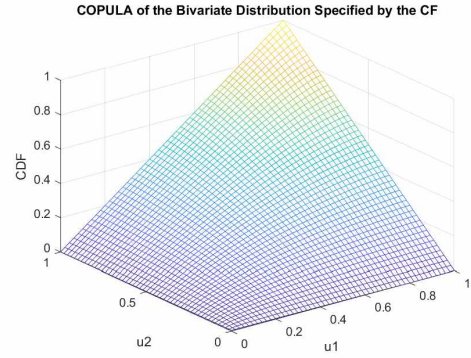


SLIKA 5.5: Kopula funkcija dvodimenzionalne logističke raspodjele sa parametrima $m = [-1; 2]$ i $\sigma = [0, 5; 0, 8]$

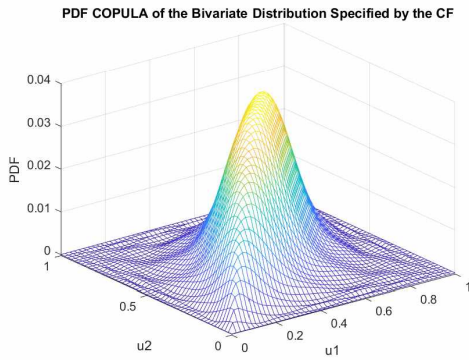
$\sigma_1 = \sigma_2 = 1$ i $\rho = 0,5$ a njene funkcije gustine i raspodjele prikazane su na slikama (5.10) - (5.11), respektivno.



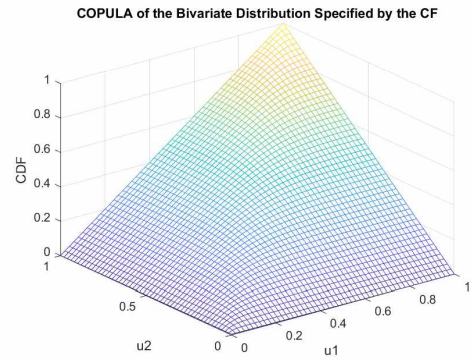
SLIKA 5.6: Kopula funkcija gustine dvodimenzionalne logističke raspodjele sa parametrima $m = [0, 0]$, $\sigma = [1, 1]$ i $\rho = 0,5$



SLIKA 5.7: Kopula funkcija dvodimenzionalne logističke raspodjele sa parametrima $m = [0, 0]$, $\sigma = [1, 1]$ i $\rho = 0,5$



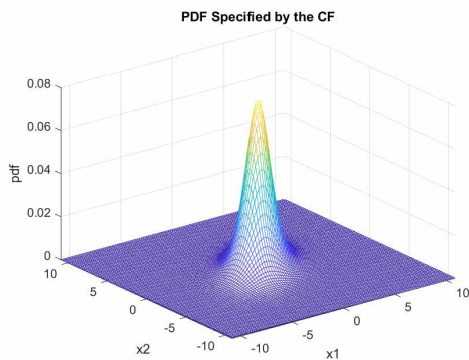
SLIKA 5.8: Kopula funkcija gustine dvodimenzionalne logističke raspodjele sa parametrima $m = [0, 1]$, $\sigma = [2, 2]$ i $\rho = 0,9$



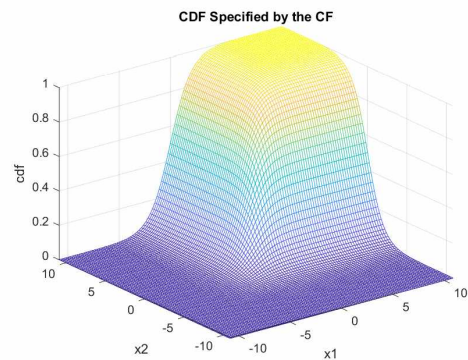
SLIKA 5.9: Kopula funkcija dvodimenzionalne logističke raspodjele sa parametrima $m = [0, 1]$, $\sigma = [2, 2]$ i $\rho = 0,9$

Pogodnost karakteristične funkcije (2.4) je ta što možemo generisati logističku raspodjelu za različite vrijednosti Pirsonovog koeficijenta $\rho \in (-1, 1)$. Na slikama (5.8) - (5.9) prikazane su kopula funkcija gustine i kopula funkcija raspodjele za dvodimenzionalnu logističku raspodjelu sa parametrima $m_1 = 0, m_2 = 1$, $\sigma_1 = \sigma_2 = 2$ i $\rho = 0,9$ a njene funkcije gustine i raspodjele prikazane su na slikama (5.12) - (5.13), respektivno.

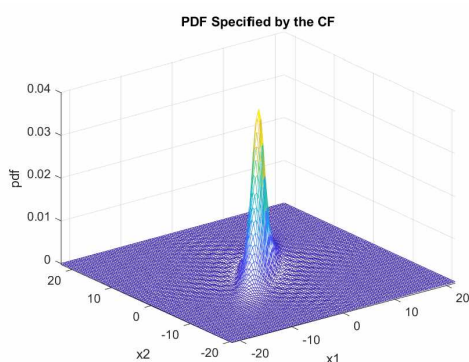
Upotrebom algoritama za numeričku inverziju možemo generisati raspodjelu linearne kombinacije nezavisnih dvodimenzionalnih logističkih slučajnih vektora čija je raspodjela specifikovana karakterističnom funkcijom (2.4). Neka je $Y = aX + bZ$, gdje je $a = (0, 5; 0, 25)$ i $b = (0, 25; 0, 5)$, $X = (X_1, X_2) \sim BivLogistic(2, 2, 1, 1)$, $Z = (Z_1, Z_2) \sim BivLogistic(3, 1, 2, 2)$ i $\rho = (0, 5; 0, 9)$. Na slikama (5.14) - (5.15) prikazane su funkcija raspodjele i funkcija gustine slučajne veličine Y , dok su na slikama (5.16) - (5.17) prikazane njene kopula funkcija raspodjele i gustine.



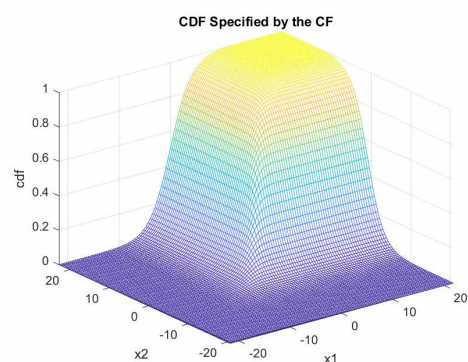
SLIKA 5.10: Funkcija gustine dvodimenzionalne logističke raspodjele sa parametrima $m = [0, 0]$, $\sigma = [1, 1]$ i $\rho = 0, 5$



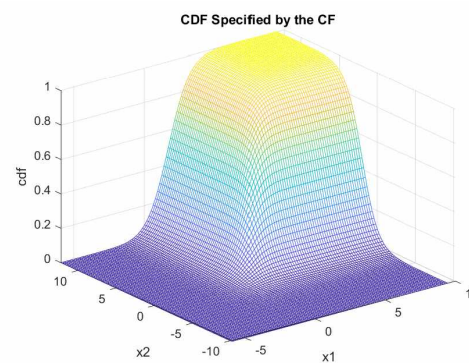
SLIKA 5.11: Funkcija raspodjele dvodimenzionalne logističke raspodjele sa parametrima $m = [0, 0]$, $\sigma = [1, 1]$ i $\rho = 0, 5$



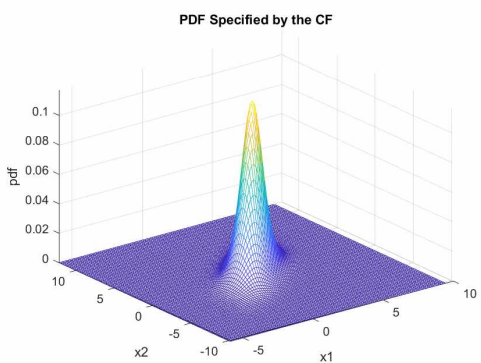
SLIKA 5.12: Funkcija gustine dvodimenzionalne logističke raspodjele sa parametrima $m = [0, 1]$, $\sigma = [2, 2]$ i $\rho = 0, 9$



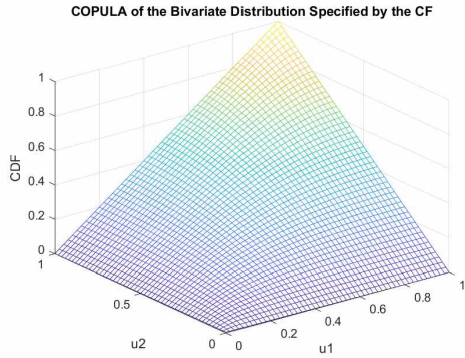
SLIKA 5.13: Funkcija raspodjele dvodimenzionalne logističke raspodjele sa parametrima $m = [0, 1]$, $\sigma = [2, 2]$ i $\rho = 0, 9$



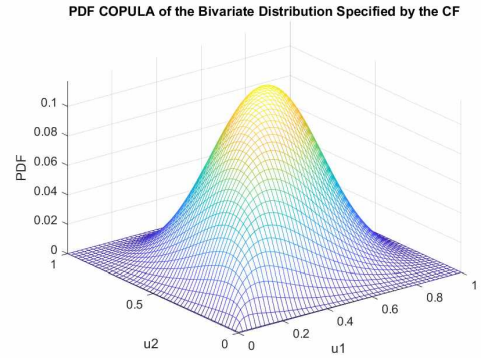
SLIKA 5.14: Funkcija raspodjele linearne kombinacije nezavisnih dvodimenzionalnih logistički raspodijeljenih slučajnih vektora



SLIKA 5.15: Funkcija gustine linearne kombinacije nezavisnih dvodimenzionalnih logistički raspodijeljenih slučajnih vektora



SLIKA 5.16: Kopula funkcija raspodjele linearne kombinacije nezavisnih dvodimenzionalnih logistički raspodijeljenih slučajnih vektora



SLIKA 5.17: Kopula funkcija gustine linearne kombinacije nezavisnih dvodimenzionalnih logistički raspodijeljenih slučajnih vektora

5.6 Dvije zavisne logistički raspodijeljene slučajne veličine

U ovom poglavlju razmotrićemo raspodjelu linearne kombinacije dvije zavisne logistički raspodijeljene slučajne veličine sa zajedničkom dvodimenzionalnom logističkom raspodjelom, koja je specifikovana karakterističnom funkcijom (5.1). Neka je $Z = \alpha X + \beta Y$, pri čemu $(X, Y) \sim BivLog(m_1, m_2, \sigma_1, \sigma_2)$, $\alpha, \beta \in \mathbb{R}$ i $m_1, m_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$.

Primijetimo, da je raspodjela slučajne veličine Z nastala transformacijom slučajnog vektora (X, Y) . Uvedimo pomoćnu slučajnu veličinu $V = X$. Sistem jednačina

$$z = \alpha x + \beta y, \quad v = x$$

ima samo jednu inverziju

$$y = \frac{z - \alpha v}{\beta}, \quad x = v.$$

Jakobijan ove transformacije jednak je

$$J = \begin{vmatrix} \frac{\partial x}{\partial v} & \frac{\partial x}{\partial z} \\ \frac{\partial y}{\partial v} & \frac{\partial y}{\partial z} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -\frac{\alpha}{\beta} & \frac{1}{\beta} \end{vmatrix} = \frac{1}{\beta}.$$

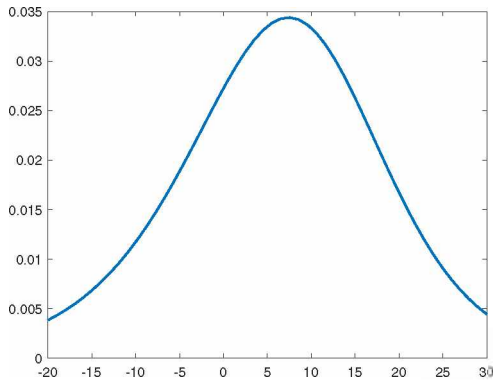
Primjenom teoreme 1.1 sada je

$$f_{V,Z}(v, z) = f_{X,Y}\left(v, \frac{z - \alpha v}{\beta}\right) |J|, \quad (5.18)$$

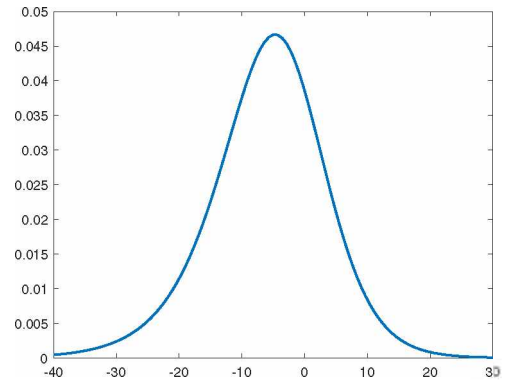
pa funkciju gustine slučajne veličine $Z = \alpha X + \beta Y$ nalazimo kao marginalnu funkciju iz (5.18):

$$f_Z(z) = \frac{1}{|\beta|} \int_{-\infty}^{+\infty} f_{X,Y} \left(v, \frac{z - \alpha v}{\beta} \right) dv = \frac{1}{|\beta|} \int_{-\infty}^{+\infty} f_{X,Y} \left(x, \frac{z - \alpha x}{\beta} \right) dx. \quad (5.19)$$

Integral iz (5.19) rješavamo numerički, gdje funkciju gustine definisanu sa (5.18) izvodimo korišćenjem algoritma 2. Na slikama 5.18 i 5.19 prikazana je funkcija gustine slučajne veličine $Z = \alpha X + \beta Y$, za $\alpha = 2$ i $\beta \in \{3, -3\}$, respektivno i $(X, Y) \sim BivLog(0, 2, 1, 2)$.



SLIKA 5.18: Funkcija gustine slučajne veličine $Z = 2X + 3Y$, gdje je $(X, Y) \sim BivLog(0, 2, 1, 2)$



SLIKA 5.19: Funkcija gustine slučajne veličine $Z = 2X - 3Y$, gdje je $(X, Y) \sim BivLog(0, 2, 1, 2)$

Glava 6

Primjena na realnim podacima

U ovom poglavlju korist ćemo raspodjelu linearne kombinacije logističkih slučajnih veličina za modelovanje finansijskih podataka i pokazati njenu upotrebu za izračunavanje nekih finansijskih instrumenata.

Portfolio predstavlja skup finansijske imovine koja je sastavljena od različitih finansijskih instrumenata (zalihe, obveznice, keš, stopa povrata). Odnosno to je skup ili kombinacija različitih hartija od vrijednosti, koje su u vlasništvu finansijskih institucija ili individualnih lica.

Neka se portfolio sastoji od n hartija od vrijednosti, tj. od n aktiva (roba, zaliha i sl.) $\{X_1, X_2, \dots, X_n\}$. Zaključnu cijenu svake aktive X_i , $i = 1, 2, \dots, n$ u trenutku t označimo sa p_t . Prinos hartije od vrijednosti od vremena $t - 1$ do vremena t definišemo kao

$$R_t = \frac{p_t - p_{t-1}}{p_{t-1}}. \quad (6.1)$$

Iz finansijske literature [2] je poznato da stopa prinosa portfolija je ponderisani prosjek prinosa na hartije od vrijednosti u portfoliju, a ponderi su određeni udjelom hartija od vrijednosti u portfoliju. To znači da se stopa prinos portfolija može izraziti kao konveksna linearna kombinacija stopa prinosa hartija od vrijednosti. Odnosno, portfolio P se izračunava kao linearna kombinacija aktiva, tj.

$$P = a_1 X_1 + a_2 X_2 + \dots + a_n X_n,$$

gdje su a_i , $i = 1, 2, \dots, n$ težinski koeficijenti (ponderi) za koje važi $\sum_{i=1}^n a_i = 1$.

Stopa prinosa portfolija P je

$$R_P = a_1 R_1 + a_2 R_2 + \dots a_n R_n,$$

gdje su R_i , $i = 1, 2, \dots, n$ stope prinosa pojedinačnih hartija od vrijednosti.

6.1 Primjer 1

Portfolio koje se sastoji od dvije različite hartije od vrijednosti može se izraziti kao

$$Z = \alpha X + \beta Y,$$

gdje su $0 \leq \alpha \leq 1$ i $0 \leq \beta \leq 1$ težinski koeficijenti portfolija koji predstavljaju udio aktiva X i Y u portfolio respektivno i $\alpha + \beta = 1$.

U ovom primjeru razmatrali smo da su težinski koeficijenti jednaki, tj. da svaki konstituent u portfoliju investira u istoj proporciji, a to je 0,5. Sa sajta <https://orfe.princeton.edu/~jqfan/fan/classes/504.html> preuzeli smo zaključne mjesečne cijene akcija preduzeća Jonson&Jonson i General Electric u periodu od 1. januara 1990. do 1. januara 1999. godine. Pretpostavićemo da su zaključne cijene akcija za dva preduzeća nezavisne slučajne veličine, jer dolaze iz dvije različite industrije.

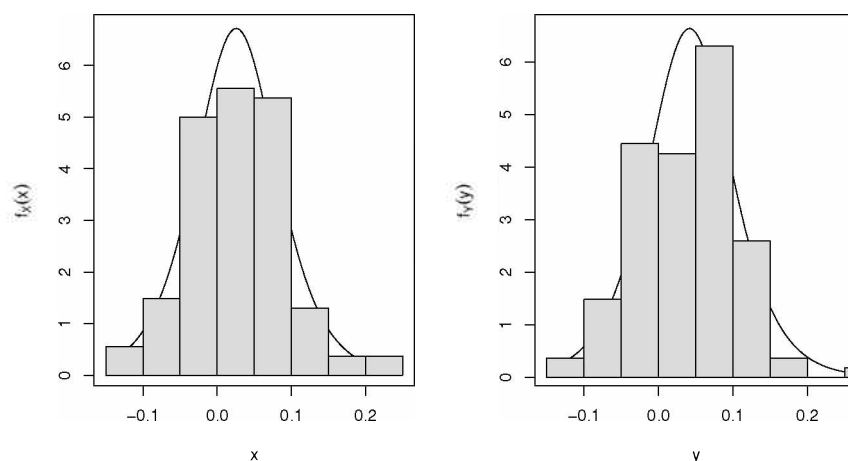
Označimo sa X i Y prinos akcija od Jonson&Jonson i General Electric respektivno. Najprije smo izračunali stope povrata R_X i R_Y koristeći formulu (6.1). Sljedeći korak je da provjerimo da li dobijeni uzorci dolaze iz logističke raspodjele. Ocijenili smo nepoznate parametre koristeći metod maksimalne vjerodostojnosti. To je odrađeno u statističkom softveru R korišćenjem funkcije `optim`, a početne vrijednosti odredili smo pomoću metoda momenata. Iz metodoloških razloga u nastavku primjenjujemo bootstrap metodu u testu Kolmogorov - Smirnova, kako bi opisali postupak u slučaju da nepoznati parametri nijesu parametri lokacije μ i skaliranja σ . U našem slučaju dovoljno je bilo primjeniti metod Monte Karla (vidjeti [4, 10]).

U testu Kolmogorov-Smirnova primjenili smo bootstrap metodu sa 10000 ponavljanja. Za više detalja o izračunavanju Kolmogorov - Smirnov statistike može se pronaći u [54]. Ocjene parametara sa standardnim greškama (*se*) kao i rezultati testa dati su tabeli 6.1.

Na osnovu bootstrap statistike (Kolmogorov - Smirnov test statistike i p vrijednosti) možemo da zaključimo da stope povrata za aktive X i Y dolaze iz logističke raspodjele, kao što je prikazano na slici 6.1.

TABELA 6.1: Ocjene nepoznatih parametara sa se za logističku raspodjelu i butstrap vrijednosti za Kolmogorov - Smirnov test statistiku D_{boot} i p_{boot} vrijednosti

Stope povrata	$\hat{\mu}$ (se)	$\hat{\beta}$ (se)	D_{boot}	p_{boot}
R_X	0,0256 ($2,045e - 07$)	0,0428 ($1,847e - 05$)	0,0849	0,6729
R_Y	0,0418 ($6,098e - 07$)	0,0428 ($1,894e - 05$)	0,1736	0,0710



SLIKA 6.1: Histogram i funkcija gustine za R_X i R_Y .

Podaci za portfolio Z su dati u tabeli 6.2. Koristeći teoreme 3.2 i 3.3 znamo da se raspodjela za Z može prikazati u obliku Foksove H funkcije. U statističkom softveru **Mathematica**, koristeći metod maksimalne vjerodostojnosti ocijenili smo parametre Foksove H funkcije. Koristili smo relaciju (1.25). Dobili smo sljedeće ocjene $\mu_1 = 0,0254$, $\mu_2 = 0,0410$ i $\beta_1 = \beta_2 = 0,0428$ i vrijednost za logaritamsku funkciju vjerodostojnosti je 160,4. Histogram stope prinosa portfolija Z i raspodjela od Z prikazani su na slici 6.2.

Jedna od osnovnih finansijskih kalkulacija je očekivana stopa prinosa portfolija i standardna devijacija portfolija. Ocjena za očekivanu stopu prinosa portfolija je

$$\frac{\mu_1 + \mu_2}{2} = 0,0332.$$

Dakle očekivana stopa prinosa portfolija je 3,32%.

Standardna devijacija portfolija pomaže u određivanju rizičnosti investicije u odnosu na očekivanu stopu prinosa. Portfolio koji ima nisku (visoku) standardnu devijaciju,

TABELA 6.2: Podaci portfolija Z

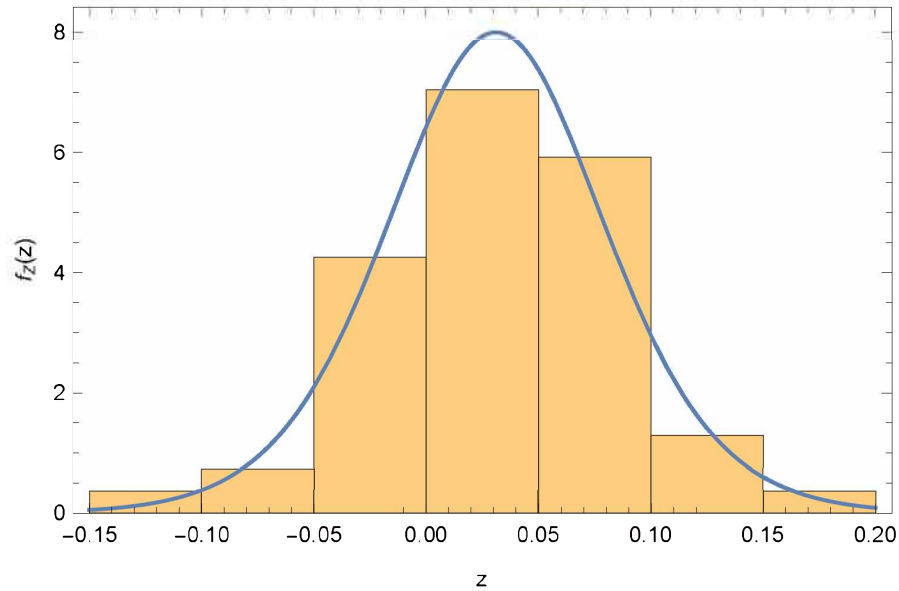
0,00795	0,11825	0,00250	0,10210
0,10225	0,03835	-0,10260	-0,05840
0,03915	0,06820	0,06030	0,16100
0,06800	0,16580	0,00095	0,02545
-0,00220	0,04360	0,01790	-0,04390
0,08350	-0,02425	0,20450	0,01610,
-0,00665	-0,02190	0,03810	0,00890
-0,01045	0,08875	-0,01510	0,01500
0,06870	0,04610	0,02430	-0,01050
-0,02865	0,04165	0,06940	0,02315
-0,00785	-0,00100	0,05695	-0,01480
0,08260	0,02490	0,06010	0,03130
-0,03755	-0,04335	0,06660	0,05705
-0,03305	0,11410	0,02705	0,01270
0,05910	-0,04085	0,07870	0,06155
0,01970	0,02810	0,08910	0,02650
0,00675	0,07880	-0,02020	0,08880
0,06555	0,06210	0,03795	0,11560
-0,02115	0,01585	0,01580	0,06195
0,04085	-0,04365	0,03765	0,07140
0,02865	0,07805	-0,05400	0,12010
-0,00515	-0,05640	0,15460	0,03470
0,07860	0,02190	-0,09240	0,05660
-0,02235	0,12000	0,02380	0,04295
0,06465	0,04450	-0,01310	-0,02770
0,08450	0,01405	-0,10170	0,06805
0,07665	0,01490	0,08355	0,02690

podrazumjeva da je rizik protfolija nizak (visok), a prinosi portfolija su (manje) stabilniji. Ocjena za standardnu devijaciju portfolija je

$$\sqrt{\mathbb{D}(Z)} = \frac{\pi\sqrt{3}}{6} \sqrt{\beta_1^2 + \beta_2^2} = 0,0549.$$

6.2 Primjer 2

Koristili smo `EuStockMarkets` skup podataka koji je dio standardne R instalacije [48]. Skup podataka `EuStockMarkets` sastoji se od dnevnih zaključnih cijena glavih evropskih akcijskih indeksa od 1991 do 1998, uključujući German DAX (Ibis), Swiss SMI, French CAC i UK FTSE. Korišćenjem (6.1) konvertovali smo zaključne cijene od



SLIKA 6.2: Histogram i funkcija gustine za stopu povrata portfolija Z

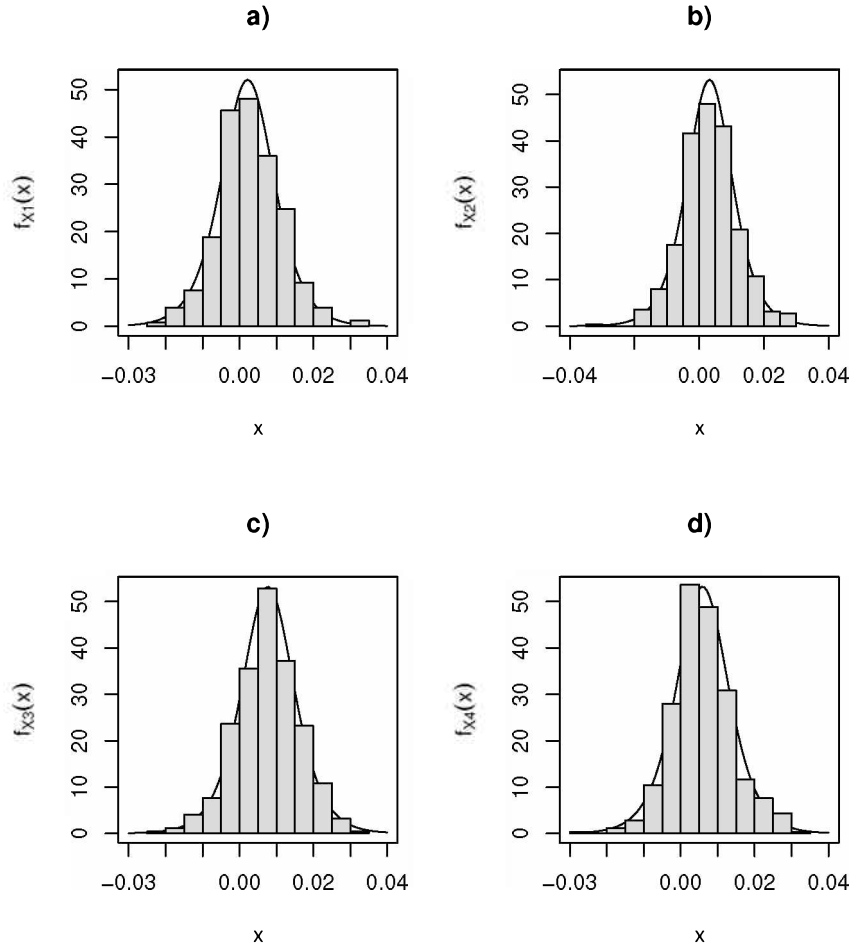
DAX, SMI, CAC i FTSE u stope povrata R_{DAX} , R_{SMI} , R_{CAC} i R_{FTSE} respektivno. Pretpostavićemo da su ove slučajne veličine nezavisne, jer predstavljaju stope povrata različitih industrija.

Sljedeći korak je da provjerimo da li dobijeni uzorci dolaze iz logističke raspodjele. Prvo smo ocijenili nepoznate parametre korišćenjem metoda maksimalne vjerodostojnosti. To je odrađeno upotrebom R funkcije `optim`, a početne vrijednosti su dobijene pomoću metoda momenata. U tabeli 6.3 date su ocjene i odgovarajuće standarde greške (*se*). Imajući u vidu da imamo ocijenjene parametre logističke raspodjele, direktna primje-

TABELA 6.3: Ocjene nepoznatih parametara sa *se* za logističku raspodjelu i butstrap vrijednosti za Kolmogorov - Smirnov test statistiku D_{boot} i p_{boot} vrijednosti

Stope povrata	$\hat{\mu}$ (<i>se</i>)	$\hat{\beta}$ (<i>se</i>)	D_{boot}	p_{boot}
R_{DAX}	0,0022 ($2e - 04$)	0,0048 ($e - 04$)	0,0164	0,9002
R_{SMI}	0,0032 ($2,3647e - 04$)	0,0047 ($9,8522e - 05$)	0,0102	0,9996
R_{CAC}	0,0077 ($2,3299e - 04$)	0,0047 ($9,9632e - 05$)	0,0162	0,9095
R_{FTSE}	0,0059 ($2,3439e - 04$)	0,0047 ($9,7920e - 05$)	0,0149	0,9526

na Kolmogorov - Smirnov testa ne bi bila prikladna. Da bi izbjegli ovu nedosljednost, korišćemo butstrap metod sa 10000 replikacija. Kao i u primjeru 1 iz metodoloških razloga biramo ovaj pristup. Za više detalja o izračunavanju Kolmogorov - Smirnov statistike D i p vrijednosti pomoću butstrap metoda pogledati [54]. Rezultati su prikazani u tabeli 6.3. Na osnovu butstrap statistike (Kolmogorov Smirnov test statistike D_{boot} i p_{boot} vrijednosti) zaključujemo da sve stope povrata dolaze iz logističke raspodjele, što je i prikazano na slici 6.3.



SLIKA 6.3: Histogram i funkcija gustine sljedećih podataka: a) R_{DAX} , b) R_{SMI} , c) R_{CAC} i d) R_{FTSE} .

Uzimajući u obzir sve navedeno, možemo uvesti slučajnu veličinu definisanu kao konveksnu linearnu kombinaciju stopa povrata kao

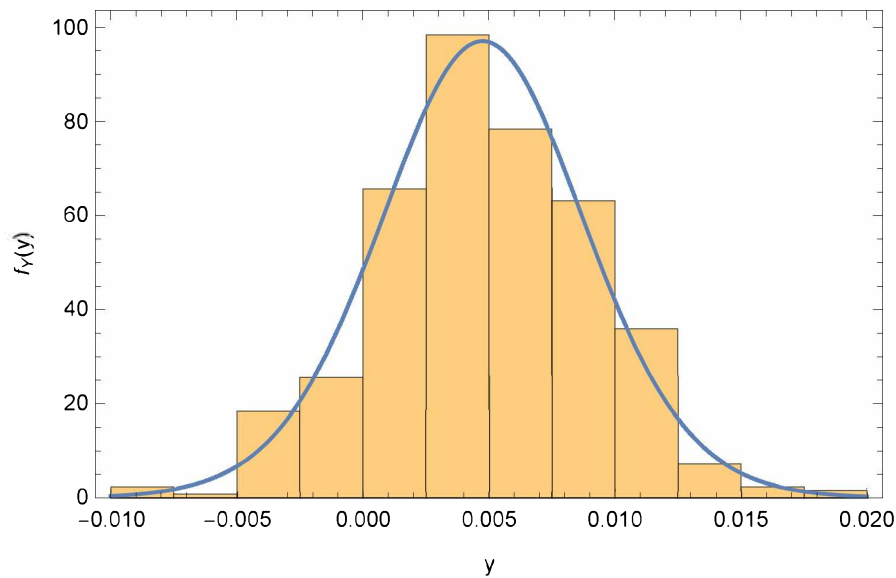
$$Y = c_1 R_{DAX} + c_2 R_{SMI} + c_3 R_{CAC} + c_4 R_{FTSE}, \quad (6.2)$$

gdje su $c_i \in [0, 1]$ za svako $i = 1, 2, 3, 4$ i $c_1 + c_2 + c_3 + c_4 = 1$. Pretpostavljamo da su svi ponderi jednaki, tj. $c_1 = c_2 = c_3 = c_4 = 0,25$. Ova pretpostavka je bila prirodna i njene prednosti su sveobuhvatno opisane u [32]. Takođe u slučaju kada komponente portfolija nijesu jednako ponderisane, tehnika optimizacije može dovesti do beskonačne petlje ili je vrijeme izvršavanje veoma dugo. Može se postaviti pitanje da li se konstante c_1, c_2, c_3, c_4 mogu smatrati nepoznatim i kako ih procijeniti. Naime, konstante se mogu smatrati nepoznatim, ali treba postaviti neka dodatna ograničenja portfolia da bi se dobile njihove procjene. Više detalja o ovom pitanju može se naći u [11].

Po teoremi 4.1 obje funkcije gustine i raspodjele slučajne veličine Y su izražene

preko Foksove H funkcije. Iako se parametri za Foksovu H funkciju mogu lako dobiti korišćenjem rezultata iz tabele 6.3, mi ćemo ocijeniti parametre za Foksovu H funkciju korišćenjem metoda maksimalne vjerodostojnosti. Za ove potrebe koristimo u softveru Mathematica funkciju FindMaximum. Tako da, dobijamo sljedeće ocjene $\mu_1 = 0,0023$; $\mu_2 = 0,0033$; $\mu_3 = 0,0079$; $\mu_4 = 0,0060$ i $\beta_i = 0,0048$ za svako $i = 1, 2, 3, 4$. Vrijednost logaritamske funkcije vjerodostojnosti je $\ell = 2019,11$. Ove ocjene imaju veoma bliske vrijednosti ocjenama iz table 6.3.

Sada preostaje da implementiramo butstrap tehniku u cilju da provjerimo da li funkcija raspodjele (4.3) odgovara slučajnoj veličini Y . Ovaj test bio je veoma vremenski zahtjevan, jer za njegovo izvršavanje bilo je potrebno više od 3 sata u Hewlett Packard mašini sa Intel Core i5 3230M procesorom. Na kraju dobili smo $D_{boot} = 0.01478$ i $p_{boot} = 0.9114$. Dakle, možemo zaključiti da vjerovatnosna raspodjela čija funkcija raspodjele je data sa (4.3) je dobar izbor. Takođe, rezultati prikazani na slici 6.4 poklapaju se sa butstrap rezultatima.



SLIKA 6.4: Histogram i funkcija gustine za slučajnu veličinu Y

Glava 7

Novi testovi saglasnosti sa dvodimenzionalnom logističkom raspodjelom

Test saglasnosti jedan je od ključnih alata u analizi podataka. Ovaj test je veoma dobro izučen u jednodimenzionalnom slučaju, ali kada se radi sa višedimenzionalnim podacima čini se da postoji nedostatak metoda za raspodjele koje nisu Gausove. S druge strane, istraživači iz različitih oblasti postavljaju pitanje prikladnosti Gausove raspodjele u modeliranju određenih empirijskih podataka koji često pokazuju teže repove od onih koji se očekuju po Gausovoj pretpostavci.

Razmotrimo slučajni vektor $X \sim BivLogistic(m_1, m_2, \sigma_1, \sigma_2)$, $m_1, m_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$ i neka je raspodjela od slučajnog vektora X specifikovana sa karakterističnom funkcijom (5.1).

Koristeći (5.1) razvićemo test saglasnosti za testiranje nulte hipoteze

$$H_0 : X \sim BivLogistic(m_1, m_2, \sigma_1, \sigma_2), \text{ za neke } m_1, m_2 \in \mathbb{R} \text{ i } \sigma_1, \sigma_2 > 0.$$

protiv opšte alternativne, tj.

$$H_1 : X \not\sim BivLogistic(m_1, m_2, \sigma_1, \sigma_2) \text{ za } m_1, m_2 \in \mathbb{R} \text{ i } \sigma_1, \sigma_2 > 0.$$

Familija dvodimenzionalnih logističkih raspodjela

$$BL = \{BivLogistic(m_1, m_2, \sigma_1, \sigma_2) : m_1, m_2 \in \mathbb{R}, \sigma_1, \sigma_2 > 0\}$$

je invarijantna u odnosu na affine transformacije primjenjene na slučajni vektor $X = (X_1, X_2)$, tj. ako je $X \in BL \Rightarrow \beta X + d \in BL$, za svaki realni dvodimenzionalni vektor d i realni koeficijent $\beta > 0$. Kako je familija BL zatvorena u odnosu na affine transformacije, razmotrićemo test statistiku koja uključuje standardizovane podatke $Z_j = (Z_j^1, Z_j^2)$, $j = 1, 2, \dots, n$, takve da

$$Z_j := \hat{\Sigma}^{-1} (X_j - \hat{m}), \quad j = 1, 2, \dots, n,$$

gdje su $(\hat{m}, \hat{\Sigma})$ odgovarajuće ocjene parametara (m, Σ) , $m = (m_1, m_2)$ i $\Sigma = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}$. U odnosu na hipotezu H_0 , podaci Z_j , $j = 1, 2, \dots, n$, za veliko n imaju približno *BivLogistic*(0, 0, 1, 1) raspodjelu.

Predlažemo dvije test statistike

$$T_{n,W} = n \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(\left| \frac{\phi_n(t_1, t_2)}{\phi(t_1, t_2)} \right|^2 - 1 \right)^2 W(t_1, t_2) dt_1 dt_2, \quad (7.1)$$

i

$$R_{n,W} = n \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |\phi_n(t_1, t_2) - \phi(t_1, t_2)|^2 W(t_1, t_2) dt_1 dt_2. \quad (7.2)$$

gdje je $\phi_n(t_1, t_2) = n^{-1} \sum_{j=1}^n e^{it_1 Z_j^1 + it_2 Z_j^2}$ empirijska karakteristična funkcija izračunata za standardizovane podatke Z_j , $j = 1, 2, \dots, n$, $\phi(t_1, t_2)$ je karakteristična funkcija za dvodimenzionalnu logističku raspodjelu zadatu sa (5.1) i $W(\cdot)$ označava težinsku funkciju koja se uvodi u cilju konvergencije integrala (7.1). Test statistiku (7.2) izveli smo na osnovu transformacije jednakosti (2.7) za dvodimenzionalni slučaj. Slična test statistiku statistici (7.1) može se naći u radu [15], s tim zbog uprošćavanja numeričkog računa odlučili smo da posmatramo kvadrat količnika empirijske i teorijske karakteristične funkcije.

Po uzoru na radove [15] i [27], kao i zbog optimizacije numeričkih proračuna za težinsku funkciju $W(\cdot)$ uzeli smo specijalni slučaj Gausovske težinske funkcije $W(t_1, t_2) = e^{-a(t_1^2 + t_2^2)}$, $a > 0$. U nastavku disertacije za test statistike (7.1) i (7.2) koristimo oznake $T_{n,a}$ i $R_{n,a}$ redom.

$$\text{Neka je } D_{T_n}(t_1, t_2) = \left| \frac{\phi_n(t_1, t_2)}{\phi(t_1, t_2)} \right|^2 - 1.$$

Test statistiku iz jednakosti (7.1) možemo zapisati kao

$$T_{n,a} = n \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |D_{T_n}(t_1, t_2)|^2 e^{-a(t_1^2 + t_2^2)} dt_1 dt_2. \quad (7.3)$$

Primjenom Ojlerove formule (1.18) dolazimo do

$$\begin{aligned} |\phi(t_1, t_2)|^2 &= |\Gamma(1 - it_1)\Gamma(1 - it_2)\Gamma(1 + it_1 + it_2)|^2 \\ &= \frac{\pi^3 t_1 t_2 (t_1 + t_2)}{\sinh(\pi t_1) \sinh(\pi t_2) \sinh(\pi(t_1 + t_2))}. \end{aligned} \quad (7.4)$$

Koristeći da je

$$|\phi_n(t_1, t_2)|^2 = \operatorname{Re}(\phi_n(t_1, t_2))^2 + \operatorname{Im}(\phi_n(t_1, t_2))^2$$

i

$$\operatorname{Re}(\phi_n(t_1, t_2)) = \frac{1}{n} \sum_{j=1}^n \cos(t_1 Z_j^1 + t_2 Z_j^2)$$

dolazimo do sljedećeg rezultata

$$|\phi_n(t_1, t_2)|^2 = \frac{1}{n^2} \sum_{j,k=1}^n \cos(t_1(Z_j^1 - Z_k^1) + t_2(Z_j^2 - Z_k^2)). \quad (7.5)$$

Kako je

$$|D_{T_n}(t_1, t_2)|^2 = \left| \frac{\phi_n(t_1, t_2)}{\phi(t_1, t_2)} \right|^4 - 2 \left| \frac{\phi_n(t_1, t_2)}{\phi(t_1, t_2)} \right|^2 + 1$$

koristeći (7.4) i (7.5) izraz u (7.3) možemo zapisati kao

$$\begin{aligned} T_{n,a} &= \frac{\pi^6}{n^3} \sum_{j,k=1}^n \sum_{i,h=1}^n \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\cos(t_1(Z_j^1 - Z_k^1) + t_2(Z_j^2 - Z_k^2)) \cos(t_1(Z_i^1 - Z_h^1) + t_2(Z_i^2 - Z_h^2))}{t_1^2 t_2^2 (t_1 + t_2)^2} \\ &\quad \times \sinh^2(\pi t_1) \sinh^2(\pi t_2) \sinh^2(\pi(t_1 + t_2)) e^{-a(t_1^2 + t_2^2)} dt_1 dt_2 - \\ &\quad - \frac{2\pi^3}{n} \sum_{j,k=1}^n \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\cos(t_1(Z_j^1 - Z_k^1) + t_2(Z_j^2 - Z_k^2))}{t_1 t_2 (t_1 + t_2)} \sinh(\pi t_1) \sinh(\pi t_2) \sinh(\pi(t_1 + t_2)) e^{-a(t_1^2 + t_2^2)} dt_1 dt_2 + \\ &\quad + n \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-a(t_1^2 + t_2^2)} dt_1 dt_2. \end{aligned}$$

Da bismo izračunali prvi integral u gornjem izrazu razvili smo funkciju kvadrat hiperboličkog sinusa u red:

$$\sinh^2 x = \sum_{s=1}^{\infty} \frac{2^{2s-1} x^{2s}}{(2s)!}, \quad x \in \mathbb{R}.$$

Dakle, ako označimo $x_1 = Z_j^1 - Z_k^1$, $y_1 = Z_j^2 - Z_k^2$ i $x_2 = Z_i^1 - Z_h^1$, $y_2 = Z_i^2 - Z_h^2$ imamo da je

$$\begin{aligned} I_1 &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\cos(t_1 x_1 + t_2 y_1) \cos(t_1 x_2 + t_2 y_2)}{t_1^2 t_2^2 (t_1 + t_2)^2} \sinh^2(\pi t_1) \sinh^2(\pi t_2) \sinh^2(\pi(t_1 + t_2)) e^{-a(t_1^2 + t_2^2)} dt_1 dt_2 \\ &= \sum_{s=1}^{\infty} \sum_{p=1}^{\infty} \sum_{q=1}^{\infty} \frac{2^{2s-1} \pi^{2s}}{(2s)!} \frac{2^{2p-1} \pi^{2p}}{(2p)!} \frac{2^{2q-1} \pi^{2q}}{(2q)!} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\cos(t_1 x_1 + t_2 y_1) \cos(t_1 x_2 + t_2 y_2)}{t_1^2 t_2^2 (t_1 + t_2)^2} t_1^{2s} t_2^{2p} (t_1 + t_2)^{2q} e^{-a(t_1^2 + t_2^2)} dt_1 dt_2 \\ &= \sum_{s=1}^{\infty} \sum_{p=1}^{\infty} \sum_{q=1}^{\infty} \frac{2^{2s-1} \pi^{2s}}{(2s)!} \frac{2^{2p-1} \pi^{2p}}{(2p)!} \frac{2^{2q-1} \pi^{2q}}{(2q)!} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \cos(t_1 x_1 + t_2 y_1) \cos(t_1 x_2 + t_2 y_2) t_1^{2s-2} t_2^{2p-2} (t_1 + t_2)^{2q-2} e^{-a(t_1^2 + t_2^2)} dt_1 dt_2. \end{aligned}$$

Po binomnoj teoremi, moguće je izraz $(t_1 + t_2)$, stepenovan sa bilo kojim nenegativnim cijelim brojem, predstaviti u formi:

$$(t_1 + t_2)^{2(q-1)} = \sum_{l=0}^{2(q-1)} \binom{2(q-1)}{l} t_1^l t_2^{2q-2-l}.$$

Dakle sada dobijamo sljedeći izraz

$$\begin{aligned} I_1 &= \sum_{s=1}^{\infty} \sum_{p=1}^{\infty} \sum_{q=1}^{\infty} \frac{2^{2s-1} \pi^{2s}}{(2s)!} \frac{2^{2p-1} \pi^{2p}}{(2p)!} \frac{2^{2q-1} \pi^{2q}}{(2q)!} \sum_{l=0}^{2(q-1)} \binom{2(q-1)}{l} \\ &\quad \times \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \cos(t_1 x_1 + t_2 y_1) \cos(t_1 x_2 + t_2 y_2) t_1^{2s-2+l} t_2^{2p+2q-l-4} e^{-a(t_1^2 + t_2^2)} dt_1 dt_2. \end{aligned}$$

U cilju izračunavanja integrala I_1 definišemo sljedeće leme.

Lema 7.1. *Neka je*

$$J_1(c, g) = \int_{-\infty}^{\infty} t^g e^{-at^2} \cos(tc) dt.$$

Tada je

$$J_1(c, g) = [1 + (-1)^g] \frac{1}{2a^{(g+1)/2}} \Gamma\left(\frac{g+1}{2}\right) e^{-c^2/(4a)} {}_1F_1\left(-\frac{g}{2}, \frac{1}{2}, \frac{c^2}{4a}\right),$$

gdje je ${}_1F_1(\cdot, \cdot; \cdot)$ Kumerova konfluentna hipergeometrijska funkcija definisana sa

$${}_1F_1(a, b; z) = \sum_{k=0}^{\infty} \frac{(a)_k}{(b)_k} \frac{z^k}{k!},$$

gdje $(a)_k$ predstavlja k -ti rastući faktorijal ili Pochhammerov simbol.

Dokaz. Dokaz slijedi iz formule 3.952.8 u [18].

□

Lema 7.2. *Neka je*

$$J_2(d, h) = \int_{-\infty}^{\infty} t^h e^{-at^2} \sin(td) dt.$$

Tada je

$$J_2(d, h) = [1 + (-1)^{h+1}] \frac{d}{2a^{(h+2)/2}} \Gamma\left(\frac{h+2}{2}\right) e^{-d^2/(4a)} {}_1F_1\left(\frac{1-h}{2}, \frac{3}{2}; \frac{d^2}{4a}\right).$$

Dokaz. Dokaz slijedi iz formule 3.952.7 u [18]. □

Primjetimo da za g neparno, $J_1(c, g) = 0$ i da za h parno, $J_2(d, h) = 0$.

Teorema 7.1. *Neka je*

$$\mathcal{I} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t_1^{2s-2+l} t_2^{2p+2q-l-4} e^{-a(t_1^2+t_2^2)} \cos(t_1 x_1 + t_2 y_1) \cos(t_1 x_2 + t_2 y_2) dt_1 dt_2.$$

Tada, imamo da je

$$\begin{aligned} \mathcal{I} = & \frac{1}{2} \{ J_1(x_1 - x_2, 2s - 2 + l) J_1(y_1 - y_2, 2p + 2q - l - 4) + J_1(x_1 + x_2, 2s - 2 + l) \\ & \times J_1(y_1 + y_2, 2p + 2q - l - 4) - J_2(x_1 - x_2, 2s - 2 + l) J_2(y_1 - y_2, 2p + 2q - l - 4) \\ & - J_2(x_1 + x_2, 2s - 2 + l) J_2(y_1 + y_2, 2p + 2q - l - 4) \}. \end{aligned}$$

Dokaz. Korišćenjem trigonometrijskih identiteta kao i lema 7.1 i 7.2 dolazimo da željenog rezultata. □

Teorema 7.2. *Neka je*

$$\mathcal{J} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t_1^{2s+l} t_2^{2(p+q)-l} e^{-a(t_1^2+t_2^2)} \cos[x_1 t_1 + y_1 t_2] dt_1 dt_2.$$

Tada, imamo da je

$$\mathcal{J} = J_1(x_1, 2s + l) J_1(y_1, 2(p + q) - l) - J_2(x_1, 2s + l) J_2(y_1, 2(p + q) - l).$$

Dokaz. Korišćenjem trigonometrijskih identiteta kao i Lema 7.1 i 7.2 dolazimo da traženog rezultata. □

Dakle, iz teoreme 7.1 slijedi da integral I_1 možemo zapisati u obliku

$$I_1 = \sum_{s=1}^{\infty} \sum_{p=1}^{\infty} \sum_{q=1}^{\infty} \frac{2^{2s-2} \pi^{2s}}{(2s)!} \frac{2^{2p-1} \pi^{2p}}{(2p)!} \frac{2^{2q-1} \pi^{2q}}{(2q)!} \sum_{l=0}^{2(q-1)} \binom{2(q-1)}{l} \\ \times \{J_1(x_1 - x_2, 2s - 2 + l) J_1(y_1 - y_2, 2p + 2q - l - 4) + J_1(x_1 + x_2, 2s - 2 + l) \\ \times J_1(y_1 + y_2, 2p + 2q - l - 4) - J_2(x_1 - x_2, 2s - 2 + l) J_2(y_1 - y_2, 2p + 2q - l - 4) \\ - J_2(x_1 + x_2, 2s - 2 + l) J_2(y_1 + y_2, 2p + 2q - l - 4)\}.$$

Sljedeći integral

$$I_2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\cos(t_1(Z_j^1 - Z_k^1) + t_2(Z_j^2 - Z_k^2))}{t_1 t_2 (t_1 + t_2)} \sinh(\pi t_1) \sinh(\pi t_2) \sinh(\pi(t_1 + t_2)) e^{-a(t_1^2 + t_2^2)} dt_1 dt_2$$

izračunamo na sličan način. Nakon primjene sličnih tehnika kao u prethodnom slučaju dobijamo

$$I_2 = \sum_{s=0}^{\infty} \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{\pi^{2s+1}}{(2s+1)!} \frac{\pi^{2p+1}}{(2p+1)!} \frac{\pi^{2q+1}}{(2q+1)!} \sum_{l=0}^{2q} \binom{2q}{l} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \cos(x_1 t_1 + y_1 t_2) t_1^{2s+l} t_2^{2p+2q-l} e^{-a(t_1^2 + t_2^2)} dt_1 dt_2.$$

Dalje, primjenom teoreme 7.2 slijedi da je

$$I_2 = \sum_{s=0}^{\infty} \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{\pi^{2s+1}}{(2s+1)!} \frac{\pi^{2p+1}}{(2p+1)!} \frac{\pi^{2q+1}}{(2q+1)!} \sum_{l=0}^{2q} \binom{2q}{l} \{J_1(x_1, 2s+l) J_1(y_1, 2(p+q)-l) \\ - J_2(x_1, 2s+l) J_2(y_1, 2(p+q)-l)\}.$$

Na kraju, posljednji integral iz izraza za test statistiku $T_{n,a}$ ima vrijednost:

$$I_3 = n \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-a(t_1^2 + t_2^2)} dt_1 dt_2 = \frac{n\pi}{a}.$$

Standardizovani podaci Z_j , $j = \overline{1, n}$ zavise od ocjena $(\hat{m}, \hat{\Sigma})$ parametara m_1, m_2 i σ_1, σ_2 . U nastavku razmatraćemo samo ocjene dobijene metodom momenata. Odnosno, ako $X = (X_1, X_2) \sim BivLogistic(m_1, m_2, \sigma_1, \sigma_2)$ tada $\mathbb{E}(X_1) = m_1$, $\mathbb{E}(X_2) = m_2$ i $\mathbb{D}(X_1) = \frac{\sigma_1^2 \pi^2}{3}$ i $\mathbb{D}(X_2) = \frac{\sigma_2^2 \pi^2}{3}$. Dakle, ocjene momenata su date sa uzoračkom sredinom i uzoračkom disperzijom, tj.

$$\hat{m} = (\overline{X_n^1}, \overline{X_n^2}) \text{ i } \hat{\Sigma} = \begin{pmatrix} \frac{\sqrt{3} S_n^1}{\pi} & 0 \\ 0 & \frac{\sqrt{3} S_n^2}{\pi} \end{pmatrix},$$

gdje su $\overline{X_n^j} = n^{-1} \sum_{i=1}^n X_i^j$ i $(S_n^j)^2 = n^{-1} \sum_{i=1}^n (X_i^j - \overline{X_n^j})^2$, za $j = \{1, 2\}$. Ove ocjene zadovoljavaju lokacijsko-skalarnu i skalarnu invarijantnost, respektivno, tj.

$$\hat{m}(AX_1 + \beta, \dots, AX_n + \beta) = A\hat{m}(X_1, X_2, \dots, X_n) + \beta$$

i

$$\hat{\Sigma}(AX_1 + \beta, \dots, AX_n + \beta) = A\hat{\Sigma}(X_1, X_2, \dots, X_n)A^T,$$

gdje je A dijagonalna regularna matrica dimenzije 2×2 i $\beta \in \mathbb{R}^2$. Da bi postigli afinu invarijantnost posmatrali smo test statistke od transformisanih podataka $Z_j = \hat{\Sigma}^{-1}(X_j - \hat{m})$, $j = 1, 2, \dots, n$. Primjetimo da svaka test statistika koja zavisi od X_j isključivo preko Z_j je afino invarijantna.

Teorema 7.3. *Neka je slučajni vektor $X = (X_1, X_2) \sim \text{BivLogistic}(0, 0, 1, 1)$. Tada kada $n \rightarrow \infty$*

$$\frac{T_{n,a}}{n} \rightarrow \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(\left| \frac{\phi(C^{-\frac{1}{2}}(t_1, t_2)^T)}{\phi(t_1, t_2)} \right|^2 - 1 \right)^2 e^{-a(t_1^2 + t_2^2)} dt_1 dt_2 \quad (7.6)$$

skoro sigurno, gdje je $\phi(t_1, t_2)$ karakteristična funkcija slučajnog vektora X i C je matrica kovarijacije definisana u (5.5).

Dokaz. Iz (7.3) imamo da je

$$\frac{T_{n,a}}{n} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |D_{T_n}(t_1, t_2)|^2 e^{-a(t_1^2 + t_2^2)} dt_1 dt_2.$$

Kako je

$$|D_{T_n}(t_1, t_2)|^2 = \left| \frac{\phi_n(t_1, t_2)}{\phi(t_1, t_2)} \right|^4 - 2 \left| \frac{\phi_n(t_1, t_2)}{\phi(t_1, t_2)} \right|^2 + 1,$$

koristeći izraz (7.5) dobijamo sljedeću ocjenu

$$|D_{T_n}(t_1, t_2)|^2 \leq \left(1 + \frac{1}{|\phi(t_1, t_2)|^2} \right)^2.$$

Na osnovu teoreme 2.7.8 iz [52] imamo da je gornja ocjena za karakterističnu funkciju data sa

$$|\phi(t)| \geq 1 - \frac{tCt^T}{2}, \quad t \in \mathbb{R}^2.$$

Tada iz posljedice 5.1 slijedi da je

$$|\phi(t_1, t_2)| \geq 1 - \frac{\pi^2}{6}(t_1^2 + t_1 t_2 + t_2^2).$$

Dakle, dobijamo da je

$$\frac{T_{n,a}}{n} \leq \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(1 + \frac{1}{\left(1 - \frac{\pi^2}{6}(t_1^2 + t_1 t_2 + t_2^2) \right)^2} \right)^2 e^{-a(t_1^2 + t_2^2)} dt_1 dt_2. \quad (7.7)$$

Integral iz nejednakosti (7.7) rješavamo uvođenjem polarnih koordinata. Za $r \in (0, \infty)$ i $\varphi \in (0, 2\pi)$ dobijamo

$$\begin{aligned} & \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(1 + \frac{1}{\left(1 - \frac{\pi^2}{6}(t_1^2 + t_1 t_2 + t_2^2) \right)^2} \right)^2 e^{-a(t_1^2 + t_2^2)} dt_1 dt_2 = \\ & = \int_0^{2\pi} \int_0^{+\infty} \left(1 + \frac{1}{\left(1 - \frac{\pi^2}{6}r^2(1 + \sin \varphi \cos \varphi) \right)^2} \right)^2 r e^{-ar^2} dr d\varphi. \end{aligned}$$

Daljim rješavanjem dolazimo do sljedećeg rezultata

$$\int_0^{2\pi} \left(1 + \frac{1}{\left(1 - \frac{\pi^2}{6}r^2(1 + \sin \varphi \cos \varphi) \right)^2} \right)^2 d\varphi = 2\pi.$$

Dakle,

$$\begin{aligned} & \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(1 + \frac{1}{\left(1 - \frac{\pi^2}{6}(t_1^2 + t_1 t_2 + t_2^2) \right)^2} \right)^2 e^{-a(t_1^2 + t_2^2)} dt_1 dt_2 = \\ & = 2\pi \int_0^{\infty} r e^{-ar^2} dr = \frac{\pi}{a}, \quad a > 0. \end{aligned} \quad (7.8)$$

Označimo sa S_n uzoračku kovaracionu matricu. Tada

$$S_n^{-\frac{1}{2}} \rightarrow C^{-\frac{1}{2}}, \quad \text{kada } n \rightarrow \infty,$$

skoro sigurno. Iz teoreme 3.2.1 iz [52] empirijska karakteristička funkcija $\phi_n(t_1, t_2)$ konvergira uniformno ka $\phi(t_1, t_2)$ na svakom kompaktnom podskupu od \mathbb{R}^2 . Dakle, imamo

da za $t \in \mathbb{R}^2$

$$\phi_n(t) = \exp\left\{-iS_n^{-\frac{1}{2}}t^T\overline{X_n}\right\}\frac{1}{n}\sum_{j=1}^n\exp\left\{iS_n^{-\frac{1}{2}}t^TX_j\right\}\rightarrow\exp\left\{-iC^{-\frac{1}{2}}t^T\mathbb{E}(X)\right\}\phi(C^{-\frac{1}{2}}t^T), \quad (7.9)$$

skoro sigurno kada $n \rightarrow \infty$.

Sada (7.6) slijedi iz (7.7), (7.8) i (7.9), primjenom Lebegove teoreme o dominantnoj konvergenciji. \square

Razmotrimo sada test statistiku

$$R_{n,a} = n \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |\phi_n(t_1, t_2) - \phi(t_1, t_2)|^2 e^{-a(t_1^2 + t_2^2)} dt_1 dt_2. \quad (7.10)$$

Neka je $D_{R_n}(t_1, t_2) = |\phi_n(t_1, t_2) - \phi(t_1, t_2)|$. Sličan rezultat kao u Teoremi 7.3 možemo izvesti i za test statistiku datu sa (7.10).

Teorema 7.4. *Neka je slučajni vektor $X = (X_1, X_2) \sim BivLogistic(0, 0, 1, 1)$. Tada kada $n \rightarrow \infty$*

$$\frac{R_{n,a}}{n} \rightarrow \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left| \phi\left(C^{-\frac{1}{2}}(t_1, t_2)^T\right) - \phi(t_1, t_2) \right|^2 e^{-a(t_1^2 + t_2^2)} dt_1 dt_2 \quad (7.11)$$

skoro sigurno, gdje je $\phi(t_1, t_2)$ karakteristična funkcija slučajnog vektora X i C je matrica kovarijacije definisana u (5.5).

Dokaz. Iz izraza (7.10) imamo

$$\frac{R_{n,a}}{n} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |\phi_n(t_1, t_2) - \phi(t_1, t_2)|^2 e^{-a(t_1^2 + t_2^2)} dt_1 dt_2, \quad (7.12)$$

takođe primijetimo da je

$$|\phi_n(t_1, t_2) - \phi(t_1, t_2)|^2 \leq 4.$$

Dalje je,

$$\frac{R_{n,a}}{n} \leq \frac{4\pi}{a}, \quad a > 0. \quad (7.13)$$

Sada kao što je pokazano u dokazu teoreme 7.3, iz (7.9), (7.12) i (7.13), primjenom Lebegove teoreme o dominantnoj konvergenciji dolazimo do rezultata (7.11). \square

7.1 Kritične vrijednosti

Pod pretpostavkom da je hipoteza H_0 tačna raspodjele test statistika definisanih sa (7.3) i (7.10) su komplikovane za analitičko izračunavanje. Zbog toga, kritične vrijednosti test statistika su izračunate pomoću Monte Karlo simulacije. Koristićemo sljedeće korake da odredimo kritične vrijednosti predloženih test statistika:

korak 1: Generisati uzorak X_1, X_2, \dots, X_n obima n iz standardne dvodimenzionalne logističke raspodjele, tj. $X_i \sim \text{BivLog}(0, 0, 1, 1)$, $i = 1, 2, \dots, n$, koristeći algoritam koji smo opisali u poglavlju 5.4;

korak 2: Izračunati vrijednosti predloženih test statistika na osnovu uzorka (X_1, X_2, \dots, X_n) . Koristeći jednakost (7.3) kreirati odgovarajuće algoritam za numeričko izračunavanje test statistike, kao što je prikazano Matlabovim kodom,

```
F = @(t,aa) (abs(cf2DEmpirical(t,data) ./ ...
    cf2DLogistic(t)).^2 - 1).^2 .* WeightFunction(t,aa);
fun = @(x,y) reshape(F([x(:),y(:)],a),size(x));
T = n * integral2(fun,t1min,t1max,t2min,t2max);
```

gdje **data** označava $(n \times 2)$ dimenzionalnu matricu u kojoj su dati podaci u kojima izračunavamo test statistiku i funkcija **WeightFunction** izračunava Gausovu težinsku funkciju. Algoritmi **Cf2DEmpirical** i **cf2DLogistic** koji izračunava dvodimenzionalnu empirijsku karaktersitičnu funkciju i dvodimenzionalnu logističku karakterističnu funkciju datu sa (5.1), respektivno, implementirani su u MATLAB repozitorijumu CharFunTool, pogledati u [55]. Na sličan način numerički izračunavamo test statistiku definisanu sa (7.10).

korak 3: Ponavljati krake 1 - 2 10000 puta i odrediti $(1 - \alpha)100\%$ percentil za predložene test statistike.

Kritične vrijednosti za test statistiku $T_{n,a}$ zasnovane na 10000 uzoraka obima $10 \leq n \leq 1000$ za nivo značajnosti $\alpha \in \{1\%, 5\%, 10\%\}$ date su u tabali 7.1. Vrijednost parametra a kontroliše brzinu opadanja težinske funkcije $\exp(-a(t_1^2 + t_2^2))$. Za velike vrijednosti parametra a težinska funkcija brže teži ka nuli i posljedično u vrijednostima test statistika $T_{n,a}$ i $R_{n,a}$ u okolini $(t_1, t_2) = (0, 0)$ dominira ponašanje $D_{T_n}(t_1, t_2)$ i $D_{R_n}(t_1, t_2)$, respektivno. Dok za male vrijednosti parametra a doprinos $D_{T_n}(t_1, t_2)$ i $D_{R_n}(t_1, t_2)$ u vrijednostima test statistika izvan okoline $(t_1, t_2) = (0, 0)$ takođe je značajan. Za nivo značajnosti 5% i obim uzorka $n \in \{30, 40, 50, 70, 100\}$ u tabeli 7.2 date su kritične vrijednosti za test statistiku $T_{n,a}$ za vrijednost parametra $a \in \{1, 5, 10\}$ i date su kritične vrijednosti za test statistiku $R_{n,a}$ za različite vrijednosti parametra a .

TABELA 7.1: Kritične vrijednosti za test statistiku $T_{n,a}$ za nivo značajnosti $\alpha \in \{0,1; 0,05; 0,01\}$, $10 \leq n \leq 1000$ i $a \in \{2; 3\}$

n	$a = 2$			$a = 3$		
	$\alpha = 0,1$	$\alpha = 0,05$	$\alpha = 0,01$	$\alpha = 0,1$	$\alpha = 0,05$	$\alpha = 0,01$
10	1,0494e + 09	1,4146e + 09	2,4670e + 09	1,9603e + 04	2,7617e + 04	4,7295e + 04
20	5,3753e + 08	7,2385e + 08	1,2767e + 09	1,0063e + 04	1,4224e + 04	2,6279e + 04
30	3,5141e + 08	4,9243e + 08	8,9836e + 08	6,7431e + 03	9,4212e + 03	1,7415e + 04
40	2,6883e + 08	3,7628e + 08	6,6255e + 08	5,1486e + 03	7,2930e + 03	1,3962e + 04
50	2,1549e + 08	3,0296e + 08	5,3364e + 08	4,1048e + 03	5,7187e + 03	1,1023e + 04
60	1,8146e + 08	2,5196e + 08	4,6611e + 08	3,4556e + 03	4,9712e + 03	9,2745e + 03
70	1,5921e + 08	2,1660e + 08	3,7539e + 08	2,9853e + 03	4,2233e + 03	8,4093e + 03
80	1,3733e + 08	1,9129e + 08	3,3694e + 08	2,5474e + 03	3,6561e + 03	6,8856e + 03
90	1,2221e + 08	1,7054e + 08	3,1818e + 08	2,2850e + 03	3,2839e + 03	6,1419e + 03
100	1,0962e + 08	1,5051e + 08	2,8153e + 08	2,0754e + 03	2,8884e + 03	5,4581e + 03
200	5,3149e + 07	7,2924e + 07	1,3923e + 08	1,0857e + 03	1,5527e + 03	2,7503e + 03
300	3,7184e + 07	5,1468e + 07	9,0299e + 07	7,2308E + 02	1,0293e + 03	1,9153e + 03
400	2,7018e + 07	3,7854e + 07	6,6999e + 07	5,4675e + 02	7,6063e + 02	1,3306e + 03
500	2,2055e + 07	3,1479e + 07	5,5925e + 07	4,7178e + 02	6,5140e + 02	1,1545e + 03
600	1,8258e + 07	2,5864e + 07	4,6206e + 07	3,8651e + 02	5,3093e + 02	9,7225e + 02
700	1,5771e + 07	2,1658e + 07	4,0796e + 07	3,5577e + 02	4,9510e + 02	8,9771e + 02
800	1,3814e + 07	1,9066e + 07	3,5297e + 07	3,1023e + 02	4,1394e + 02	7,6054e + 02
900	1,2405e + 07	1,7433e + 07	3,1749e + 07	2,9559e + 02	3,9697e + 02	7,2266e + 02
1000	1,0872e + 07	1,5150e + 07	2,7819e + 07	2,7137e + 02	3,7032e + 02	6,3918e + 02

Uporedićemo dvije predložene statistike, na taj način što ćemo generisati slučajne brojeve iz različitih dvodimenzionalnih raspodjela. U simulaciji razmatrali smo 10000 uzoraka obima $n \in \{30, 40, 50, 70, 100\}$ iz različitih dvodimenzionalnih raspodjela. Te raspodjele su: standardna normalna raspodjela (N), standardna Laplasova raspodjela (L) i Studentova t -raspodjela za različite stepene slobode.

Za generisanje slučajnih brojeva iz dvodimenzionalne standardne normalne raspodjele koristili smo ugrađenu funkciju u Matlabu `mvnrnd`. Za dati vektor X koji dolazi iz dvodimenzionalne standardne normalne raspodjele, ostale raspodjele Y simulirali smo na sljedeći način:

1. standardna Laplasova raspodjela: $Y_L = \sqrt{w}X$;
2. Studentova t_m raspodjela: $Y_m = \left(\frac{\sqrt{S}}{m}\right)^{-1} X$ označena sa $t(m)$;

gdje w dolazi iz standardne eksponencijalne raspodjele i S dolazi iz χ_m^2 raspodjele.

U tabelama B.3 i B.4 iz dodatka B dati su odgovarajući procenti odbacivanja (zaokruženi na najbliži cio broj) za obje test statistike i simulirane dvodimenzionalne raspodjele. Uočavamo da za uzorke malog obima $n \leq 40$ i $a \geq 5$ test $T_{n,a}$ se pokazao moćnijim.

Rezultati iz glave 7 su dio rada [33].

TABELA 7.2: Kritične vrijednosti za test statistike $T_{n,a}$ i $R_{n,a}$ za nivo značajnosti $\alpha = 0.05$, $n \in \{30, 40, 50, 100\}$ i $a \in \{0, 5; 1; 2; 3; 5; 10\}$

n	$T_{n,a}$				$R_{n,a}$					
	$a = 0, 5$	$a = 1$	$a = 5$	$a = 10$	$a = 0, 5$	$a = 1$	$a = 2$	$a = 3$	$a = 5$	$a = 10$
30	$1,5024e + 32$	$1,4023e + 22$	12,6039	0,6638	9,2362	4,4534	1,9666	1,1670	0,5562	0,1936
40	$1,1067e + 32$	$9,9868e + 21$	11,1609	0,6589	9,0992	4,5092	1,9875	1,1735	0,5645	0,1974
50	$8,8256e + 31$	$8,4856e + 21$	10,4015	0,6697	9,4758	4,5552	2,0038	1,1763	0,5846	0,2059
70	$6,3392e + 31$	$5,9803e + 21$	10,1497	0,7071	9,5676	4,7328	2,1180	1,2594	0,6046	0,2151
100	$4,7487e + 31$	$4,1849e + 21$	9,5613	0,7365	9,7549	4,8548	2,2060	1,3146	0,6757	0,2274

Zaključak

Karakteristične funkcije imaju veoma značajnu ulogu u teoriji vjerovatnoće i matematičkoj statistici. Postoje brojne knjige i naučni radovi koji se bave matematičkom teorijom karakterističnih funkcija i njihovim primjenama. S druge strane, jednodimenzionalna logistička raspodjela je intenzivno proučavana, posebno jer je često posmatrana kao alternativa za Gausovu raspodjelu. Sve su to bili osnovni motivi zbog kojih smo odlučili da se bavimo ovom temom i damo doprinos u daljem razvitku teorije i primjene.

Počevši od jednodimenzionalne logističke raspodjele izveli smo funkciju raspodjele i funkciju gustine za linearnu kombinaciju n nezavisnih logistički raspodijeljenih slučajnih veličina. U tu svrhu razvili smo algoritam za numeričko računanje vrijednosti Foksove H funkcije koji se po prvi put javlja u literaturi. U cilju da provjerimo efikasnost i preciznost našeg algoritma za Foksovu H funkciju razmatrali smo još jedan pristup. Odlučili smo se za numeričko invertovanje karakteristične funkcije, dizajnirali smo odgovarajući algoritam i uporedili smo rezultate.

Za razliku od obimne literature koja je posvećena dvodimezionalnoj i višedimenzionalnoj normalnoj raspodjeli, malobrojne rezultate možemo pronaći za višedimenzionalnu logističku raspodjelu. Što nam je bio podstrek da u nastavku disertacije izvedemo neke od značajnih rezultata za dvodimenzionalnu logističku raspodjelu.

Izveli smo izraz za karakterističnu funkciju dvodimenzionalne logističke raspodjele i njenim invertovanjem došli do njene funkcije raspodjele i funkcije gustine. U disertaciji smo opisali algoritam kako se funkcije raspodjele i gustine mogu numerički izračunati u dvodimenzionalnom slučaju. Takođe smo prikazali algoritam za generisanje slučajnih brojeva iz dvodimezionalne raspodjele. Ovaj algoritam je zasnovan na uslovnoj karakterističnoj funkciji i taj pristup se prvi put javlja u literaturi. Što dalje omogućava praktičnu primjenu kompleksnih raspodjela koje su specifikovane svojim karakterističnim funkcijama, putem numeričkog izračunavanja njihovih funkcija raspodjela. Na taj način uspostavljamo konekciju između čisto teorijskih rezultata i primjena.

Dalje, proučavali smo empirijsku karakterističnu funkciju i njene primjene u statistici. Konkretno, bavili smo se pristupom empirijskih karakterističnim funkcija u testu

saglasnosti sa dvodimenzionalnom logističkom raspodjelom. Predložili smo dvije test statistike i izračunali smo odgovarajuće kritične vrijednosti primjenom Monte Karlo simulacije za 5% nivo značajnosti. U nastavku smo sproveli simulacionu studiju i uporedili moć predloženih test statistika.

Na kraju, modelirali smo stope povrata sa logističkom raspodjelom na stvarnim podacima. Zatim, procijenili smo očekivanu stopu prinosa i volatilitet portfolija.

U districiji smo obradili mnogobrojne značajne teme iz oblasti teorije vjerovatnoća, matematičke statistike i numeričke analize, što je omogućilo da teorijske rezultate pretočimo u praktične primjere. Nadam se da će naš doprinos biti podstrek za dalji razvoj ove oblasti.

Dodatak A

Algoritam za Foksovu H funkciju

```
Off[NIntegrate::precw];
(*FoxH Implementation Starts*)(*FoxH Error Messages*)
FoxH::NegativeTol = "The tolerance is not greater than zero.";
FoxH::InconsistentCoeffs = "The coefficients are inconsistent.";
(*FoxH Default Option Values*)
FoxHDefaultFractionTolerance = 0.001;
FoxHDefaultDuplicationLimit = 50;
(*FoxHDefaultWorkingPrecision=$MachinePrecision;*)
FoxHDefaultWorkingPrecision = $MachinePrecision;
FoxHDefaultMaxRecursion = 40;
(*FoxH Options*)
Options[FoxH] = {FoxHMaxRecursion -> FoxHDefaultMaxRecursion,
  FoxHFractionTolerance -> FoxHDefaultFractionTolerance,
  FoxHDuplicationLimit -> FoxHDefaultDuplicationLimit,
  FoxHWorkingPrecision -> FoxHDefaultWorkingPrecision};
(*FoxH Function Module*)
FoxH[a_, b_, z_,
  OptionsPattern[{FoxHMaxRecursion -> FoxHDefaultMaxRecursion,
    FoxHFractionTolerance -> FoxHDefaultFractionTolerance,
    FoxHDuplicationLimit -> FoxHDefaultDuplicationLimit,
    FoxHWorkingPrecision -> FoxHDefaultWorkingPrecision}]] :=
Module[{evalmaxrec, evaltol, evalprec, evalduplim, ra, rb, aa, ab,
  ba, bb, raa, rab, rba, rbb, s, I, K, L, M, Pa, Pb, Qa, Qb, j, n, m,
  p, q, Z, G, R, A, B, T, Rmax, Rmin, value},
  evalmaxrec = OptionValue[FoxHMaxRecursion];
  evaltol = OptionValue[FoxHFractionTolerance];
  evalprec = OptionValue[FoxHWorkingPrecision];
  evalduplim = OptionValue[FoxHDuplicationLimit];
  value = {}; A = B = {{}, {}};
  ra = N[a]; rb = N[b]; aa = a[[1]]; ba = b[[1]]; n = Length[aa];
  m = Length[ba];
  ab = a[[2]]; bb = b[[2]];
  raa = N[aa]; rab = N[ab]; rba = N[ba]; rbb = N[bb];
  Rmin = -Infinity; Rmax = Infinity;
  If[evaltol < 0, Message[FoxH::NegativeTol];
    value = Indeterminate;
    Return[value];];
  (*Rationalize the coefficients in order to use Legendre duplication \
formula*)
  raa[[All, 2]] = Rationalize[#, evaltol] & /@ aa[[All, 2]];
```

```

rba[[All, 2]] = Rationalize[#, evaltol] & /@ ba[[All, 2]];
If[rab == {}, ra = raa; p = 0, p = Length[ab];
rab[[All, 2]] = Rationalize[#, evaltol] & /@ ab[[All, 2]];
ra[[All, All, 2]] = Join[{raa[[All, 2]]}, {rab[[All, 2]]} ];

If[rbb == {}, rb = rba; q = 0, q = Length[bb];
rbb[[All, 2]] = Rationalize[#, evaltol] & /@ bb[[All, 2]];
rb[[All, All, 2]] = Join[{rba[[All, 2]]}, {rbb[[All, 2]]} ];
(*Find K factor which is the least common multiple of rationalized \
coefficients*)

If[rab == {} || rbb == {},
K = Apply[LCM,
Flatten[{Denominator[#] & /@ raa[[All, 2]],
Denominator[#] & /@ rba[[All, 2]]}],
K = Apply[LCM,
Flatten[{Denominator[#] & /@ ra[[All, All, 2]],
Denominator[#] & /@ rb[[All, All, 2]]}]];

(*Find the coefficient and the input value according to the K \
factor*)
G = K; Z = Power[z, K];
(*Compute the upper Alpha coefficients*)
If[p == 0, ra = Join[{ra}, {}]];
If[q == 0, rb = Join[{rb}, {}]];
Do[L = ra[[1, j, 2]] K;
Z = Z Power[L, L];
G = G Sqrt[Power[L, 1 - 2 ra[[1, j, 1]]/Power[2 Pi, L - 1]]];
A[[1]] =
FlattenAt[{A[[1]],
1 + ra[[1, j, 1]]/L - Range[1, L]/L}, {{1}, {2}}, {j, 1, n}];
(*Compute the upper Alpha Mellin function*)
Pa = Function[u,
Product[SetAccuracy[Gamma[1 - a[[1, j, 1]] - u a[[1, j, 2]]],
evalprec], {j, 1, n}]];
(*Compute the lower Alpha coefficients*)
If[p > 0, Do[L = ra[[2, j, 2]] K;
Z = Z Power[L, L];
G = G/Sqrt[Power[L, 2 ra[[2, j, 1]] - 1]/Power[2 Pi, L - 1]]];
A[[2]] =
FlattenAt[{A[[2]],
ra[[2, j, 1]]/L + Range[0, L - 1]/L}, {{1}, {2}}, {j, 1, p}];
(*Compute the lower Alpha Mellin function*)
Qa = Function[u,
Product[SetAccuracy[Gamma[a[[2, n, 1]] + u a[[2, n, 2]]],
evalprec], {n, 1, p}], A[[2]] = {}];
(*Compute the upper Beta coefficients*)
Do[L = rb[[1, j, 2]] K;
Z = Z/Power[L, L];
G = G Sqrt[Power[L, 2 rb[[1, j, 1]] - 1]/Power[2 Pi, L - 1]]];
B[[1]] =
FlattenAt[{B[[1]],
rb[[1, j, 1]]/L + Range[0, L - 1]/L}, {{1}, {2}}, {j, 1, m}];
(*Compute the upper Beta Mellin function*)
Pb = Function[u,
Product[SetAccuracy[Gamma[b[[1, j, 1]] + u b[[1, j, 2]]],
evalprec], {j, 1, m}]];
(*Compute the lower Beta coefficients*)
If[q > 0, Do[L = rb[[2, j, 2]] K;
Z = Z/Power[L, L];

```

```

G = G/Sqrt[Power[L, 1 - 2 rb[[2, j, 1]]/Power[2 Pi, L - 1]]];
B[[2]] =
  FlattenAt[{B[[2]]},
    1 + rb[[2, j, 1]]/L - Range[1, L]/L, {{1}, {2}}, {j, 1, q}];
(*Compute the lower Beta Mellin function*)
Qb = Function[u,
  Product[SetAccuracy[Gamma[1 - b[[2, j, 1]] - u b[[2, j, 2]]],
    evalprec], {j, 1, q}], B[[2]] = {}];
(*Numerical value check*)
A = N[A]; B = N[B]; G = N[G]; Z = N[Z];
(*Compute the overall Mellin function*)
If[p == 0, M = Function[u, Pa[u] Pb[u]/Qb[u]]];
If[q == 0, M = Function[u, Pa[u] Pb[u]/Qa[u]]];
If[(p == 0) && (q == 0), M = Function[u, Pa[u] Pb[u]]];
If[(p > 0) && (q > 0), M = Function[u, Pa[u] Pb[u]/Qa[u]/Qb[u]]];
(*Compute the FoxH function*)
T = Length[A[[1]]] + Length[A[[2]]] + Length[B[[1]]] +
  Length[B[[2]]];
If[T <= evalduplim, Print["Computing by Meijer's G Function"];
  Print[T]; value = G MeijerG[A, B, Z],
  Print["Computing by Contour Integration"]; {Rmin,
    Rmax} = {Max[-Min[b[[1, All, 1]]/b[[1, All, 2]]], -Infinity],
    Min[Min[1 - a[[1, All, 1]]/a[[1, All, 2]]], Infinity]};
If[Rmin == -Infinity, Rmin = Rmax - 2];
If[Rmax == Infinity, Rmax = Rmin + 2];
If[Rmin == Rmax, Message[FoxH::InconsistentCoeffs];
  value = Indeterminate;
  Return[value];]; R = Mean[{Rmax, Rmin}];
R = SetAccuracy[R, evalprec];
Z = N[Z]; Z = SetAccuracy[Z, evalprec];
I = N[Sqrt[-1]]; I = SetAccuracy[I, evalprec];
If[Abs[R] < Infinity,
  value = 1/2/Pi NIntegrate[
    M[R + Sqrt[-1] s] Power[Z, -R - Sqrt[-1] s], {s, -Infinity,
      Infinity}, WorkingPrecision -> evalprec,
    MaxRecursion -> evalmaxrec], value = Indeterminate];];
]; Return[value];]

```

Dodatak B

Tablice

TABELA B.1: PDF/CDF od Y , linearne kombinacije nezavisnih logistički raspodijeljenih slučajnih veličina izračunate u zadatim tačkama y , korišćenjem CharFunTool

y	PDF	CDF	y	PDF	CDF
-5,0	3,09117201939e-05	4,82412463059e-05	13,0	1,18948832902e-01	6,23556847003e-01
-4,5	4,19718453427e-05	6,63161196174e-05	13,5	1,10879078825e-01	6,81120619514e-01
-4,0	5,71372398187e-05	9,08951262301e-05	14,0	1,00641393987e-01	7,34073997132e-01
-3,5	7,78817150776e-05	1,24380262381e-04	14,5	8,90718179242e-02	7,81540120019e-01
-3,0	1,06214720571e-04	1,70037616831e-04	15,0	7,69936675533e-02	8,23061690135e-01
-2,5	1,44869636423e-04	2,32309929796e-04	15,5	6,51203796801e-02	8,58568893568e-01
-2,0	1,97556758454e-04	3,17238721265e-04	16,0	5,39971135518e-02	8,88308132418e-01
-1,5	2,69301434286e-04	4,33033239807e-04	16,5	4,39822550448e-02	9,12751718848e-01
-1,0	3,66893266677e-04	5,90835042253e-04	17,0	3,52605906858e-02	9,32506708364e-01
-0,5	4,99478459238e-04	8,05741490179e-04	17,5	2,78757821501e-02	9,48235680179e-01
0,0	6,79333745761e-04	1,09816908634e-03	18,0	2,17701956048e-02	9,60596039392e-01
0,5	9,22865852253e-04	1,49565824027e-03	18,5	1,68232132002e-02	9,70199144391e-01
1,0	1,25188309161e-03	2,03524386700e-03	19,0	1,28829491390e-02	9,77587077865e-01
1,5	1,69518204444e-03	2,76653897014e-03	19,5	9,78953585543e-03	9,83223190846e-01
2,0	2,29047680188e-03	3,75569661048e-03	20,0	7,39031656758e-03	9,87492221961e-01
2,5	3,08666259342e-03	5,09042154393e-03	20,5	5,54839631918e-03	9,90706280409e-01
3,0	4,14633808959e-03	6,88618345839e-03	21,0	4,14633808960e-03	9,93113816542e-01
3,5	5,54839631918e-03	9,29371959058e-03	21,5	3,08666259342e-03	9,94909578456e-01
4,0	7,39031656758e-03	1,25077780395e-02	22,0	2,29047680188e-03	9,96244303390e-01
4,5	9,78953585543e-03	1,67768091545e-02	22,5	1,69518204444e-03	9,97233461030e-01
5,0	1,28829491390e-02	2,24129221351e-02	23,0	1,25188309161e-03	9,97964756133e-01
5,5	1,68232132002e-02	2,98008556093e-02	23,5	9,22865852253e-04	9,98504341760e-01
6,0	2,17701956048e-02	3,94039606085e-02	24,0	6,79333745761e-04	9,98901830914e-01
6,5	2,78757821501e-02	5,17643198209e-02	24,5	4,99478459238e-04	9,99194258510e-01
7,0	3,52605906858e-02	6,74932916356e-02	25,0	3,66893266677e-04	9,99409164958e-01
7,5	4,39822550448e-02	8,72482811521e-02	25,5	2,69301434286e-04	9,99566966760e-01
8,0	5,39971135518e-02	1,11691867582e-01	26,0	1,97556758454e-04	9,99682761279e-01
8,5	6,51203796801e-02	1,41431106432e-01	26,5	1,44869636423e-04	9,99767690070e-01
9,0	7,69936675533e-02	1,76938309865e-01	27,0	1,06214720571e-04	9,99829962383e-01
9,5	8,90718179242e-02	2,18459879981e-01	27,5	7,78817150776e-05	9,99875619738e-01
10,0	1,00641393987e-01	2,65926002868e-01	28,0	5,71372398187e-05	9,99909104874e-01
10,5	1,10879078825e-01	3,18879380486e-01	28,5	4,19718453427e-05	9,99933683880e-01
11,0	1,18948832902e-01	3,76443152997e-01	29,0	3,09117201939e-05	9,99951758754e-01
11,5	1,24123860262e-01	4,37344457841e-01	29,5	2,28792218243e-05	9,99965098593e-01
12,0	1,25907707043e-01	5,00000000000e-01	30,0	1,70900941747e-05	9,99975010672e-01
12,5	1,24123860262e-01	5,6265542159e-01			

TABELA B.2: PDF/CDF od Y , linearne kombinacije nezavisnih logistički raspodijeljenih slučajnih veličina izračunate korišćenjem Foksove H funkcije

y	PDF	CDF	y	PDF	CDF
-5,0	3,05148709909e-05	4,8876230587e-05	13,0	1,18948832879e-01	6,23556846983e-01
-4,5	4,1681498523e-05	6,67806892284e-05	13,5	1,10879078796e-01	6,8112061948e-01
-4,0	5,69248142399e-05	9,12350158272e-05	14,0	1,0064139395e-01	7,34073997083e-01
-3,5	7,77262992987e-05	1,24628932647e-04	14,5	8,90718178762e-02	7,81540119948e-01
-3,0	1,06101014861e-04	1,7021954854e-04	15,0	7,69936674889e-02	8,23061690037e-01
-2,5	1,44786446916e-04	2,32443034715e-4	15,5	6,51203795931e-02	8,58568893432e-01
-2,0	1,97495895346e-04	3,17336103256e-04	16,0	5,39971134335e-02	8,88308132231e-01
-1,5	2,69256905675e-04	4,33104486053e-04	16,5	4,39822548837e-02	9,12751718592e-01
-1,0	3,66860688728e-04	5,90887167322e-04	17,0	3,52605904658e-02	9,32506708014e-01
-0,5	4,9945462463e-04	8,0577962576e-04	17,5	2,78757818498e-02	9,482356797e-01
0,0	6,79316307947e-04	1,09819698691e-03	18,0	2,17701951945e-02	9,60596038736e-01
0,5	9,22853094449e-04	1,49567865279e-03	18,5	1,68232126395e-02	9,70199143494e-01
1,0	1,25187375779e-03	2,03525880113e-03	19,0	1,28829483727e-02	9,77587076639e-01
1,5	1,69517521566e-03	2,76654989616e-03	19,5	9,78953480811e-03	9,8322318917e-01
2,0	2,29047180582e-03	3,75570460413e-03	20,0	7,39031513612e-03	9,8749221967e-01
2,5	3,08665893821e-03	5,09042739219e-03	20,5	5,54839436265e-03	9,90706277279e-01
3,0	4,14633541537e-03	6,88618773704e-03	21,0	4,14633541537e-03	9,93113812263e-01
3,5	5,54839436265e-03	9,29372272087e-03	21,5	3,08665893821e-03	9,94909572608e-01
4,0	7,39031513612e-03	1,25077803296e-02	22,0	2,29047180582e-03	9,96244295396e-01
4,5	9,78953480811e-03	1,67768108299e-02	22,5	1,69517521566e-03	9,97233450104e-01
5,0	1,28829483727e-02	2,24129233608e-02	23,0	1,25187375779e-03	9,97964741199e-01
5,5	1,68232126395e-02	2,98008565059e-02	23,5	9,22853094449e-04	9,98504321347e-01
6,0	2,17701951945e-02	3,94039612643e-02	24,0	6,79316307947e-04	9,98901803013e-01
6,5	2,78757818498e-02	5,17643203005e-02	24,5	4,9945462463e-04	9,99194220374e-01
7,0	3,52605904658e-02	6,74932919862e-02	25,0	3,66860688728e-04	9,99409112833e-01
7,5	4,39822548837e-02	8,72482814082e-01	25,5	2,69256905675e-04	9,99566895514e-01
8,0	5,39971134335e-02	1,11691867769e-01	26,0	1,97495895346e-04	9,99682663897e-01
8,5	6,51203795931e-02	1,41431106568e-01	26,5	1,44786446916e-04	9,99767556965e-01
9,0	7,69936674889e-02	1,76938309963e-01	27,0	1,06101014861e-04	9,99829780451e-01
9,5	8,90718178762e-02	2,18459880052e-01	27,5	7,77262992987e-05	9,99875371067e-01
10,0	1,0064139395e-01	2,65926002917e-01	28,0	5,69248142399e-05	9,99908764984e-01
10,5	1,10879078796e-01	3,1887938052e-01	28,5	4,1681498523e-05	9,99933219311e-01
11,0	1,18948832879e-01	3,76443153017e-01	29,0	3,05148709909e-05	9,99951123769e-01
11,5	1,24123860242e-01	4,37344457851e-01	29,5	2,23368068897e-05	9,99964230685e-01
12,0	1,25907707024e-01	5,00000000000e-01	30,0	1,6348724472e-05	9,99973824404e-01
12,5	1,24123860242e-01	5,6265542149e-01			

TABELA B.3: Procenat odbacivanja nulte hipoteze za test statistiku $R_{n,a}$ na nivou značajnosti $\alpha = 5\%$

$R_{n,a}$		$a = 0.5$	$a = 1$	$a = 2$	$a = 3$	$a = 5$	$a = 10$
N	$n = 30$	93	97	97	96	86	0
	$n = 40$	99	99	99	99	99	8
	$n = 50$	100	100	100	100	100	99
	$n = 70$	100	100	100	100	100	100
	$n = 100$	100	100	100	100	100	100
L	$n = 30$	99	100	100	100	99	0
	$n = 40$	100	100	100	100	100	8
	$n = 50$	100	100	100	100	100	100
	$n = 70$	100	100	100	100	100	100
	$n = 100$	100	100	100	100	100	100
$t(1)$	$n = 30$	99	100	100	100	99	0
	$n = 40$	100	100	100	100	100	8
	$n = 50$	100	100	100	100	100	100
	$n = 70$	100	100	100	100	100	100
	$n = 100$	100	100	100	100	100	100
$t(2)$	$n = 30$	99	100	100	100	99	0
	$n = 40$	100	100	100	100	100	8
	$n = 50$	100	100	100	100	100	100
	$n = 70$	100	100	100	100	100	100
	$n = 100$	100	100	100	100	100	100
$t(5)$	$n = 30$	99	100	100	100	99	0
	$n = 40$	100	100	100	100	100	8
	$n = 50$	100	100	100	100	100	100
	$n = 70$	100	100	100	100	100	100
	$n = 100$	100	100	100	100	100	100

TABELA B.4: Procenat odbacivanja nulte hipoteze za test statistiku $T_{n,a}$ na nivou značajnosti $\alpha = 5\%$

$T_{n,a}$		$a = 0.5$	$a = 1$	$a = 2$	$a = 3$	$a = 5$	$a = 10$
N	$n = 30$	5	6	7	14	99	100
	$n = 40$	5	6	7	18	100	100
	$n = 50$	5	6	7	22	100	100
	$n = 70$	5	6	7	34	100	100
	$n = 100$	5	6	7	63	100	100
L	$n = 30$	5	6	7	11	100	100
	$n = 40$	5	6	7	13	100	100
	$n = 50$	5	6	7	17	100	100
	$n = 70$	5	6	7	28	100	100
	$n = 100$	5	6	7	60	100	100
$t(1)$	$n = 30$	5	5	7	10	100	100
	$n = 40$	5	5	7	12	100	100
	$n = 50$	5	5	7	18	100	100
	$n = 70$	5	5	7	28	100	100
	$n = 100$	5	5	7	61	100	100
$t(2)$	$n = 30$	5	6	7	11	100	100
	$n = 40$	5	6	7	13	100	100
	$n = 50$	5	6	7	18	100	100
	$n = 70$	5	6	7	28	100	100
	$n = 100$	5	6	7	61	100	100
$t(5)$	$n = 30$	5	6	7	11	100	100
	$n = 40$	5	6	7	13	100	100
	$n = 50$	5	6	7	18	100	100
	$n = 70$	5	6	7	28	100	100
	$n = 100$	5	6	7	61	100	100

Literatura

- [1] Abate, J., Whitt, W.: The fourier-series method for inverting transforms of probability distributions. *Queueing Syst.* **10**, 5–87 (1992)
- [2] Anderson, D., Sweeney, D., Williams, T., Camm, J., Cochran, J.: *Essentials of Modern Business Statistics*. South-Western, Cengage Learning, Boston (2018)
- [3] Andrews, G.E., Askey, R., Roy, R.: *Special Functions*, 2 edn. Cambridge University Press, Cambridge, United Kingdom (2000)
- [4] Antle, C.E., Bain, L.J.: Short notes: A property of maximum likelihood estimators of location and scale parameters. *SIAM Review* 1969-apr vol. 11 iss. 2 **11** (1969)
- [5] Arnold, B.C.: Distributions with logistic marginals and/or conditionals. Rüschendorf, L., Schweizer, B., Taylor, M.D. (Eds.) **28**, 15–32 (1996)
- [6] Balakrishnan, N.: *Handbook of the Logistic Distribution*. Marcel Dekker, New York (1992)
- [7] Balakrishnan, N., Ma, C., Wang, R.: Logistic vector random fields with logistic direct and cross covariances. *Journal of Statistical Planning and Inference* **161**, 109–118 (2015)
- [8] Bartlett, M.: The characteristic function of a conditional statistic. *Journal of the London Mathematical Society* (1938)
- [9] Berrut, J.P.: Baryzentrische formeln zur trigonometrischen interpolation (i). *Zeitschrift für angewandte Mathematik und Physik ZAMP* **35**, 91–105 (1984)
- [10] Castro, R.: *Lectures 2 and 3 - goodness-of-fit (gof) tests* (2013)
- [11] Chalabi, Y., Würtz, D.: Portfolio allocation. *Computational Actuarial Science with R* pp. 447–470 (2015)
- [12] Davies, R.: Numerical inversion of a characteristic function. *Biometrika* **60(2)**, 415–417 (1973)

- [13] Feng, L., Lin, X.: Inverting analytic characteristic functions and financial applications. *SIAM J Financ Math.* **4**, 372–398 (2013)
- [14] Ford, W.: *Numerical Linear Algebra with Applications Using MATLAB*. Academic Press (2015)
- [15] Fragiadakis, K., Meintanis, S.G.: Goodness-of-fit tests for multivariate laplace distributions. *Mathematical and Computer Modelling*, Elsevier (2011)
- [16] Gautschi, W.: *Numerical Analysis*, 2 edn. Springer, New York (2010)
- [17] Gil-Pelaez, J.: Note on the inversion theorem. *Biometrika* **38**, 481–482 (1995)
- [18] Gradshteyn, I., Ryzhik, I.: *Table of Integrals, Series, and Products* (7th. ed.). Academic Press, San Diego (2007)
- [19] Henric, P.: *Essentials of Numerical Analysis*. Wiley, New York (1982)
- [20] Kotz, S., Balakrishnan, N., Johnson, N.: *Continuous Multivariate Distributions*. John Wiley and Sons, New York (2000)
- [21] Lagrange, J.L., McCormack, T.J.: *Lectures on Elementary Mathematic*. Nabu Press (2010)
- [22] Lange, K.: *Numerical Analysis for Statisticians*. Springe (2010)
- [23] Leuangthong, O., Khan, D., Deutsch, C.: *Solved Problem in Geostatistics*. Wiley Interscience (2008)
- [24] Lifshits, M.: *Random processes by example*. World Scientific Publishing Company (2014)
- [25] Marques, F.J.: On the linear combination of independent logistic random variables. *Statistics, Optimization and Information Computing* **6(3)**, 383–397 (2018)
- [26] Mathai, A., Saxena, R.K., Haubold, H.J.: *The H-Function: Theory and Applications*, 1 edn. Springer (2010)
- [27] Meintanis, S.G.: Goodness-of-fit tests for the logistic distribution based on empirical transforms. *Indian Statistical Institute* (2004)
- [28] Mijanović, A., Popović, B., Witkovský, V.: A numerical inversion of the bivariate characteristic function. *Applied Mathematics and Computation* **4**, 372–398 (2023)
- [29] Monahan, J.F.: *Numerical Methods of Statistics*. Cambridge University Press, New York (2011)

- [30] Nelsen, R.B.: An introduction to copulas, 2nd. edn. Springer series in statistics. Springer (2006)
- [31] Petrović, L.: Teorija verovatnoća. Ekonomski fakultet Univerziteta u Beogradu (2021)
- [32] Pfeiffer, T., Cozmanciuc, E., Colin, F.: When size doesn't matter: Equal weighting (vs. market cap weighting) (2018). URL <https://www.solactive.com/wp-content/uploads/2018/04/Solactive-Equal-Weighting-vs.-Market-Cap-Weighting.pdf>
- [33] Popović, B., Mijanović, A., Witkovský, V.: The goodness of fit test in case of bivariate logistic distribution. Statistical Papers (u fazi recenzije) (2023)
- [34] Popović, B., Mijanović, A., Genc, A.: On linear combination of generalized logistic random variables with an application to financial returns. Applied Mathematics and Computation **381** (2020)
- [35] Popović, B., Mijanović, A., Witkovský, V.: Computing the exact distribution of a linear combination of generalized logistic random variables and its applications. Journal of Statistical Computation and Simulation **92**, 1015–1033 (2022)
- [36] Poularikas, A.D.: Transforms and applications handbook (2010)
- [37] Rathie, P.N., Ozelim, L., Otiniano, C.: Exact distribution of the product and the quotient of two stable lévy random variables. Communications in Nonlinear Science and Numerical Simulation **36**, 204–218 (2016)
- [38] Ristić, M., Nastić, A., Popović, B.: Matematička statistika. Univerzitet u Nišu, PMF (2021)
- [39] Rutishauser, H.: Lectures on Numerical Mathematics, vol. 1. Birkhäuser Boston (1990)
- [40] Schwarz, H.R.: Numerical Analysis: A Comprehensive Introduction. Wiley, New York (1909)
- [41] Shephard, N.: From characteristic function to distribution function: A simple framework for the theory. Econometric Theory **7**, 519–529 (1991)
- [42] Shephard, N.: Numerical integration rules for multivariate inversions. Journal of Statistical Computation and Simulation **39**, 37–46 (1991)
- [43] Slater, L.J.: Generalized hypergeometric functions. ambridge University Press (1966)

- [44] Srivastava, H.M., Manocha, H.L.: A treatise on generating functions. Ellis Horwood Limited, London (1984)
- [45] Stein, E.M., Shakarchi, R.: Fourier analysis: an introduction. Princeton Lectures in Analysis. Princeton University Press (2003)
- [46] Strawderman, R.L.: Computing tail probabilities by numerical fourier inversion: the absolutely continuous case. *Stat Sin.* **14**, 175–201 (2004)
- [47] Szego, G.: Orthogonal Polynomials. American Mathematical Society, Providence (1978)
- [48] Team, R.C.: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2019). URL <http://www.R-project.org/>
- [49] Towhidi, M., Salmanpour, M.: A new goodness-of-fit test based on the empirical characteristic function. Taylor & Francis (2014)
- [50] Trefethen, L.N.: Spectral Methods in MATLAB. SIAM, Philadelphia (2000)
- [51] Trefethen, L.N., Weideman, J.A.C.: Two results on polynomial interpolation in equally spaced points. *Journal of Approximation Theory* **65**, 247–260 (1991)
- [52] Ushakov, N.: Selected Topics in Characteristic Functions. Utrecht (1999)
- [53] Waller, L.A., Turnbull, B.W., Hardin, J.M.: Obtaining distribution functions by numerical inversion of characteristic functions with applications. *Am Stat.* **49**, 346–350 (1995)
- [54] Wang, C., Zeng, B., Shao, J.: Application of bootstrap method in kolmogorov–smirnov test. International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering, Xi'an **161**, 287–291 (2011)
- [55] Witkovský, V.: Charfuntool: The characteristic functions toolbox (matlab). *SIAM J Financ Math.* (2020). URL <https://github.com/witkovsky/CharFunTool>
- [56] Yilmaz, F., Alouini, M.S.: Product of the powers of generalized nakagami-m variates and performance of cascaded fading channels. *Proceedings of IEEE Global Telecommunications Conference* p. 6178–6185 (2009)

Biografija

Andela Mijanović je rođena u Nikšiću, 17. septembra 1994. godine. Osnovnu školu i gimnaziju završila je u rodnom gradu. Dobitnica je diplome „Luča” za odličan uspjeh u svim razredima osnovnog i srednjeg školovanja.

Trogodišnje studije na Prirodno-matematičkom fakultetu u Podgorici, smjer Matematika i računarske nauke, završila je 2016. godine, sa prosječnom ocjenom 9,27. Iste godine upisuje specijalističke studije na istom smjeru i diplomira u julu 2017. godine, pod mentorstvom prof. dr Siniše Stamatovića brani rad pod nazivom „Maksimum Vinerovog procesa i arkus sinusni zakon”.

U toku studiranja bila je dobitnik stipendije koju Vlada Republike Crne Gore dodjeljuje talentovanim učenicima i studentima i slične stipendije Opštine Nikšić.

Magistarski rad pod nazivom „Analiza rekorda u atletici primjenom teorije ekstremnih vrijednosti”, odbranila je u decembru 2018. godine, pod mentorstvom prof. dr Siniše Stamatovića. U oktobru 2019. godine upisuje doktorske studije na Prirodno-matematičkom fakultetu na smjeru Matematika.

Na Prirodno-matematičkom fakultetu, Univerziteta Crne Gore, zaposlena je na mjestu saradnika u nastavi od oktobra 2018. godine. Izvodi vježbe na predmetima: Teorija vjerovatnoće, Vjerovatnoća i statistika, Statistika, Slučajni procesi, Lanci Markova, Matematika 1, Statistika u pedagogiji, Statistička analiza, Osnovi statistike, Statistika u sociologiji, Kvantitativne statističke metode (posljednjih pet predmeta na Filozofskom fakultetu), Statistika u farmaciji i Matematika na Medicinskom fakultetu.

Oblasti njegovog naučnog interesovanja su: statistika, teorija vjerovatnoće, slučajni procesi.

Izjava o autorstvu

Potpisana Anđela Mijanović

Broj indeksa/upisa 1/2019

Izjavljujem

da je doktorska disertacija pod naslovom „**Višedimenzionalna logistička raspodjela specifikovana svojom karakterističnom funkcijom - doprinos teoriji i praksi**”

- rezultat sopstvenog istraživačkog rada,
- da predložena disertacija ni u cjelini ni u djelovima nije bila predložena za dobijanje bilo koje diplome prema studijskim programima drugih ustanova visokog obrazovanja,
- da su rezultati korektno navedeni, i
- da nijesam povrijedila autorska i druga prava intelektualne svojine koja pripadaju trećim licima.

Potpis doktoranda

U Podgorici, 10.06.2023.

Anđela Mijanović

Izjava o istovjetnosti štampane i elektronske verzije doktorskog rada

Ime i prezime autora: Anđela Mijanović

Broj indeksa/upisa: 1/2019

Studijski program: Matematika

Naslov rada: „Višedimenzionalna logistička raspodjela specifikovana svojom karakterističnom funkcijom - doprinos teoriji i praksi”

Mentor: prof. dr Božidar Popović

Potpisana Anđela Mijanović

Izjavljujem da je štampana verzija mog doktorskog rada istovjetna elektronskoj verziji koju sam predala za objavljivanje u Digitalni arhiv Univerziteta Crne Gore.

Istovremeno izjavljujem da dozvoljavam objavljivanje mojih ličnih podataka u vezi sa dobijanjem akademskog naziva doktora nauka, odnosno zvanja doktora umjetnosti, kao što su ime i prezime, godina i mjesto rođenja, naziv disertacije i datum odbrane rada.

Potpis doktoranda

U Podgorici, 10.06.2023.

Anđela Mijanović

Izjava o korišćenju

Ovlašćujem Univerzitetsku biblioteku da u Digitalni arhiv Univerziteta Crne Gore pohrani moju doktorsku disertaciju pod naslovom: „**Višedimenzionalna logistička raspodjela specifikovana svojom karakterističnom funkcijom - doprinos teoriji i praksi**”, koja je moje autorsko djelo.

Disertaciju sa svim prilogima predala sam u elektronskom formatu pogodnom za trajno arhiviranje.

Moju doktorsku disertaciju pohranjenu u Digitalni arhiv Univerziteta Crne Gore mogu da koriste svi koji poštuju odredbe sadržane u odabranom tipu licence Kreativne zajednice (Creative Commons) za koju sam se odlučila.

☒ 1. Autorstvo

2. Autorstvo – nekomercijalno

3. Autorstvo – nekomercijalno – bez prerade

4. Autorstvo – nekomercijalno – dijeliti pod istim uslovima

5. Autorstvo – bez prerade

6. Autorstvo – dijeliti pod istim uslovima

U Podgorici, 10.06.2023.

Potpis doktoranda

Stefana Mujaković