



REPUBLIKA E SHQIPËRISË
UNIVERSITETI POLITEKNIK I TIRANËS
FAKULTETI I INXHINIERISË MATEMATIKE DHE INXHINIERISË FIZIKE
DEPARTAMENTI I INXHINIERISË MATEMATIKE

PROGRAMI I STUDIMIT:
STATISTIKË DHE KËRKIME OPERACIONALE

TEZË DOKTORATURE
KONTRIBUTI I ZBËRTHIMIT TË EDGEWORTH NË METODAT E
RIZGJEDHJES

Doktoranti
Msc. Edlira DONEFSKI

Udhëheqës shkencor
Prof. Dr. Lorenc EKONOMI

Tiranë, 2021



REPUBLIKA E SHQIPËRISË
UNIVERSITETI POLITEKNIK I TIRANËS
FAKULTETI I INXHINIERISË MATEMATIKE DHE INXHINIERISË FIZIKE
DEPARTAMENTI I INXHINIERISË MATEMATIKE

DISERTACION

i paraqitur nga

Msc. Edlira DONEFSKI

Udhëhequr nga

Prof. Dr. Lorenc EKONOMI

Për marrjen e gradës shkencore
Doktor në Inxhinieri Matematike

Me Temë:

KONTRIBUTI I ZBËRTHIMIT TË EDGEWORTH NË METODAT E RIZGJEDHJES

Mbrohet më datë ____ / ____ /2021, para jurisë:

1. Prof. _____ Kryetar
2. Prof. _____ Anëtar (Oponent)
3. Prof. _____ Anëtar (Oponent)
4. Prof. _____ Anëtar
5. Prof. _____ Anëtar

MIRËNJOHJE DHE FALENDERIME

Në fund të këtij rrugëtimi të gjatë me punë intensive dhe kolosale drejt marrjes së gradës “Doktor”, dua të shpreh mirënjohjen time të sinqertë për të gjithë personat që më mbështetën dhe besuan tek unë. Kjo punë e finalizuar nuk do të kishte qenë kaq e thjeshtë pa praninë e tyre.

Si fillim falenderoj udhëheqësin tim të nderuar Prof. Dr. Lorenc Ekonomi, për të gjithë mbështetjen profesionale dhe humane përgjatë punës me tezën e doktoraturës.

Mirënjohja dhe falenderimi më i përzemërt shkon për familjen time, për gjithë motivimin, mbështetjen dhe inkurajimin moral dhe shpirtëror përgjatë gjithë rrugëtimin tim.

Gjithashtu dëshiroj të falënderoj kolegët, miqtë, të afërmit dhe gjithë pedagogët e mi për këshillat, udhëzimet dhe suportin e pakushtëzuar.

Faleminderit të gjithëve!

DEKLARATË MBI ORIGJINALITETIN

Deklaroj se kjo tezë përfaqëson punën time origjinale dhe nuk kam përdorur burime të tjera, përveç atyre të cituara dhe të pasqyruara tek Bibliografia.

Tiranë, më ___/___/2021

Kandidati

Edlira Donefski

PËRMBAJTJA

| | |
|----------------------------------------|-----|
| Lista e figurave..... | vii |
| Lista e tabelave..... | ix |
| Shkurtime..... | xi |
| Gjuhët e programimit të përdorura..... | xii |
| Hyrje..... | xiv |

KAPITULLI I

METODAT E RIZGJEDHJES DHE ZBËRTHIMI EDGEWORTH

| | |
|------------------------------------------------------------------------------------------------|----|
| 1.1 Metodatat e rizgjedhjes..... | 1 |
| 1.2 Zbërthimi Edgeworth..... | 2 |
| 1.3 Simulime të Zbërthimit Edgeworth për ndryshoren e rastit të vazhdueshme eksponenciale..... | 8 |
| 1.4 Përafrimi Edgeworth për shpërndarjen hi-katror..... | 13 |

KAPITULLI II

METODA BOOTSTRAP DHE ZBËRTHIMI EDGEWORTH PËR VERSIONIN BOOTSTRAP

| | |
|-----------------------------------------------------------------------------------------------------------------|----|
| 2.1 Metoda bootstrap..... | 18 |
| 2.2 Zbërthimi Edgeworth për versionin bootstrap..... | 21 |
| 2.3 Performanca e përafrimit Edgeworth për versionin me dhe pa bootstrap në tre shpërndarje të rëndësishme..... | 24 |

KAPITULLI III

INTERVALET E BESIMIT BOOTSTRAP

| | |
|------------------------------------------------------------------------|----|
| 3.1 Intervalet e besimit bootstrap..... | 30 |
| 3.2 Zbërthimi Edgeworth për korrektësinë e intervaleve të besimit..... | 32 |
| 3.3 Zbërthimi Edgeworth për intervale të besimit bootstrap..... | 33 |
| 3.4 Krahasimi i intervaleve të besimit bootstrap..... | 34 |
| 3.5 Simulime për shpërndarjen lognormale..... | 35 |

KAPITULLI IV

ZBATIMI I PËRAFRIMIT EDGEWORTH DHE METODËS BOOTSTRAP NË EKONOMETRI. APLIKIMI I PËRAFRIMIT EDGEWORTH DHE BOOTSTRAP NË ANALIZËN E REGRESIT

| | |
|-------------------------------------------------------|----|
| 4.1 Këndvështrim i përgjithshëm mbi Ekonometrinë..... | 39 |
| 4.2 Marrëdhënia e Regresit Linear dhe Kuantil..... | 42 |

| | |
|---------------------------------------------------------------------------------------------------------------------------------|----|
| 4.3 Implementimi i përafrimit Edgeworth për versionin Bootstrap..... | 46 |
| 4.4 Kontributi i Përafrimit Edgeworth në një rast real studimor..... | 50 |
| 4.5 Kontributi i Përafrimit Edgeworth për versionin Bootstrap në prognozimin e trendit të shitjeve për produktet kozmetike..... | 59 |

KAPITULLI V

ZBATIMI I ZBËRTHIMIT EDGEWORTH DHE VERSIONIT BOOTSTRAP NË SERITË KOHORE

| | |
|-----------------------------------------------------------------------------------------------------------------|-----|
| 5.1 Kontributi i përafrimit Edgeworth dhe Metodave të rizgjedhjes Bootstrap në prognozimin e serive kohore..... | 70 |
| 5.2 Zbërthimi Edgeworth për seritë autoregressive stacionare..... | 74 |
| 5.3 Modeli Bootstrap për modelin e Vektorit të Autoregresit (VAR)..... | 75 |
| 5.4 Përpunimi i të dhënave financiare nëpërmjet zbatimit të Metodave Edgeworth, Bootstrap dhe Monte Carlo..... | 76 |
| PËRFUNDIME | 102 |
| BIBLIOGRAFIA | 103 |

LISTA E FIGURAVE

| | |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figura 1.3.1-Figura 1.3.9 Shpërndarja empirike, përafrimi normal, përafrimi Edgeworth me dy terma dhe përafrimi Edgeworth me tre terma për $n=5, 15, 30$ dhe $N=100, 1000, 1000$ | 9 |
| Figura 1.4.1-Figura 1.4.6 Përafrimi Edgeworth për mesataren e n ndryshoreve hi-katror, për $n=10, 100, df=5, 10, 15$ | 14 |
| Figura 2.3.1 Sjellja polinomiale për rastin pa bootstrap dhe rastin bootstrap (Shpërndarja Normale)..... | 25 |
| Figura 2.3.2 Sjellja asimptotike e polinomit të rendit të parë në rastin pa bootstrap dhe rastin bootstrap (Shpërndarja Normale)..... | 25 |
| Figura 2.3.3 Sjellja asimptotike e polinomit të rendit të dytë në rastin pa bootstrap dhe rastin bootstrap (Shpërndarja Normale)..... | 26 |
| Figura 2.3.4 Sjellja polinomiale për rastin pa bootstrap dhe rastin bootstrap (Shpërndarja Eksponenciale)..... | 26 |
| Figura 2.3.5 Sjellja asimptotike e polinomit të rendit të parë në rastin pa bootstrap dhe rastin bootstrap (Shpërndarja Eksponenciale)..... | 27 |
| Figura 2.3.6 Sjellja asimptotike e polinomit të rendit të dytë në rastin pa bootstrap dhe rastin bootstrap (Shpërndarja Eksponenciale)..... | 27 |
| Figura 2.3.7 Sjellja polinomiale për rastin pa bootstrap dhe rastin bootstrap (Shpërndarja Lognormale)..... | 28 |
| Figura 2.3.8 Sjellja asimptotike e polinomit të rendit të parë në rastin pa bootstrap dhe rastin bootstrap (Shpërndarja Lognormale)..... | 28 |
| Figura 2.3.9 Sjellja asimptotike e polinomit të rendit të dytë në rastin pa bootstrap dhe rastin bootstrap (Shpërndarja Lognormale)..... | 29 |
| Figura 3.5.1-Figura 3.5.3 Vija e densitetit empirik e shpërndarjes lognormale për $R=1000$ zgjedhje bootstrap dhe $n=30, 100, 500$ | 36 |
| Figura 4.1.1 Këndvështrim mbi Shkencën e Ekonometrisë..... | 39 |
| Figura 4.2.1 Regresi Linear dhe Kuantil..... | 43 |
| Figura 4.4.1 Shpërndarja e Devijimeve Standarde për 30 kuantile..... | 52 |
| Figura 4.4.2 Shpërndarja e Devijimeve Standarde për 30 kuantile pas zbatimi i Përafrimit Edgeworth..... | 53 |
| Figura 4.4.3 Këndvështrim krahasimor i tendencës së shpërndarjes së Devijimeve Standarde për çdo kuantil..... | 53 |
| Figura 4.4.4 Shpërndarja e devijimeve standarde për secilin kuantil pas zbatimit të versionit Bootstrap..... | 54 |
| Figura 4.4.5 Shpërndarja e devijimeve standarde pas aplikimit të përafrimit Edgeworth të Versionit Bootstrap..... | 56 |
| Figura 4.4.6 Grafik përmbledhës i rezultateve të këtij studimi..... | 56 |

| | |
|------------------------------------------------------------------------------------------------------|-----|
| Figura 4.4.7 Rezultate përmblendhëse për Modelin e vitit 2020..... | 59 |
| Figura 4.5.1 Testi i Normalitetit për Regresin Linear, viti 2019..... | 60 |
| Figura 4.5.2 Testi i Normalitetit për Regresin Linear, viti 2020..... | 60 |
| Figura 4.5.3 Shpërndarja e devijimeve standarde të Modeleve Kuantile për 2019 dhe 2020..... | 64 |
| Figura 4.5.4 Përafrimi Edgeworth i shpërndarjes së gabimeve (devijimi standard)... | 66 |
| Figura 4.5.5 Versioni Bootstrap i shpërndarjes së gabimeve për çdo kuantil..... | 67 |
| Figura 4.5.6 Përafrimi Edgeworth i versionit Bootstrap të Regresit Kuantil (2019-2020)..... | 69 |
| Figura 4.5.7 Përmblendhje rezultatesh..... | 69 |
| Figura 5.1.1 Analizë e Serive Kohore..... | 70 |
| Figura 5.4.1 Jostacionariteti i serisë së Parave dhe Ekuivalentëve të Parave..... | 76 |
| Figura 5.4.2 Jostacionariteti i serisë së Aktivitetit Financiar..... | 77 |
| Figura 5.4.3 Jostacionariteti i serisë së Aktivitetit të Investimit..... | 78 |
| Figura 5.4.4 Jostacionariteti i serisë së Aktivitetit të Shfrytëzimit..... | 79 |
| Figura 5.4.5 Stacionariteti i serisë së Parave dhe Ekuivalentëve të Parave..... | 81 |
| Figura 5.4.6 Stacionariteti i serisë së Aktivitetit Financiar..... | 82 |
| Figura 5.4.7 Stacionariteti i serisë së Aktivitetit të Investimit..... | 83 |
| Figura 5.4.8 Stacionariteti i serisë së Aktivitetit të Shfrytëzimit..... | 84 |
| Figura 5.4.9 Mbetjet e Modelit VAR..... | 89 |
| Figura 5.4.10 (a), (b) Paraqitja grafike e Analizës Asimptotike..... | 93 |
| Figura 5.4.11 (a), (b) Paraqitja grafike e simulimeve Monte Carlo..... | 97 |
| Figura 5.4.12 Rezultate përmblendhëse të Përafrimit Edgeworth dhe Simulimeve Monte Carlo..... | 99 |
| Figura 5.4.13 Mbivendosje e rezultateve të përgjigjeve të çdo serie..... | 100 |

LISTA E TABELAVE

| | |
|---------------------------------------------------------------------------------------------------------------------------------|----|
| Tabela 3.4.1 Intervalet e besimit bootstrap për nivelin 95% | 34 |
| Tabela 3.4.2 Intervalet e besimit bootstrap për nivelin 99%..... | 35 |
| Tabela 3.5.3 Probabilitetet e mbulimit (CP) të intervaleve të besimit për nivelet 95% dhe 99% | 35 |
| Tabela 3.5.1 Intervalet e besimit bootstrap kuantil 95% dhe 99%, për pritjen matematike në shpërndarjen lognormale | 36 |
| Tabela 4.2.1 Autokorrelacioni, Multikolariteti dhe Heteroskedasticiteti..... | 44 |
| Tabela 4.4.1 Modeli i Regresit Kuantil për marrëdhënien Çmim-Kërkesë për vitin 2019..... | 51 |
| Tabela 4.4.2 Përafrimi Edgeworth për Modelin e Regresit Kuantil..... | 52 |
| Tabela 4.4.3 Zbatimi i Bootstrap tek Modeli i Regresit Kuantil..... | 54 |
| Tabela 4.4.4 Përafrimi Edgeworth i Versionit Bootstrap të Modelit..... | 55 |
| Tabela 4.4.5 Modeli i Regresit Kuantil për marrëdhënien Çmim-Kërkesë për vitin 2020..... | 57 |
| Tabela 4.4.6 Përafrimi Edgeworth i Modelit të Regresit Kuantil për vitin 2020..... | 57 |
| Tabela 4.4.7 Zbatimi Bootstrap i Modelit të Regresit Kuantil për vitin 2020..... | 58 |
| Tabela 4.4.8 Përafrimi Edgeworth i versionit Bootstrap i Modelit për vitin 2020..... | 58 |
| Tabela 4.5.1 Testi i korrelacionit serial (LM), viti 2019..... | 61 |
| Tabela 4.5.2 Testi i korrelacionit serial (LM), viti 2020..... | 61 |
| Tabela 4.5.3 Testi i heteroskedasticitetit, viti 2019..... | 62 |
| Tabela 4.5.4 Testi i heteroskedasticitetit, viti 2020..... | 62 |
| Tabela 4.5.5 Modeli i Regresit Kuantil për marrëdhënien Shitje-Fitime, viti 2019..... | 63 |
| Tabela 4.5.6 Modeli i Regresit Kuantil për marrëdhënien Shitje-Fitime, viti 2020..... | 63 |
| Tabela 4.5.7 Përafrimi Edgeworth për Modelin e Regresit Kuantil për vitin 2019..... | 65 |
| Tabela 4.5.8 Përafrimi Edgeworth për Modelin e Regresit Kuantil për vitin 2020..... | 65 |
| Tabela 4.5.9 Versioni Bootstrap i Modelit të Regresit Kuantil për vitin 2019..... | 66 |
| Tabela 4.5.10 Versioni Bootstrap i Modelit të Regresit Kuantil për vitin 2020..... | 67 |
| Tabela 4.5.11 Përafrimi Edgeworth i versionit Bootstrap të Regresit Kuantil, 2019..... | 68 |
| Tabela 4.5.12 Përafrimi Edgeworth i versionit Bootstrap të Regresit Kuantil, 2020..... | 68 |
| Tabela 5.4.1 Testi Unit Root për serinë e Parave dhe Ekuivalentëve të Parave..... | 76 |
| Tabela 5.4.2 Testi Unit Root për serinë e Aktivitetit Financiar..... | 77 |
| Tabela 5.4.3 Testi Unit Root për serinë e Aktivitetit të Investimit..... | 78 |
| Tabela 5.4.4 Testi Unit Root për serinë e Aktivitetit të Shfrytëzimit..... | 79 |

| | |
|-------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Tabela 5.4.5 Stacionariteti i serisë së Parave dhe Ekuivalentëve të Parave..... | 81 |
| Tabela 5.4.6 Stacionariteti i serisë së Aktivitetit Financiar..... | 82 |
| Tabela 5.4.7 Stacionariteti i serisë së Aktivitetit të Investimit..... | 83 |
| Tabela 5.4.8 Stacionariteti i serisë së Aktivitetit të Shfrytëzimit..... | 84 |
| Tabela 5.4.9 Modeli i Vektorit të Autoregresit (VAR)..... | 85 |
| Tabela 5.4.10 Testimi për korrelacion serial (LM Test)..... | 88 |
| Tabela 5.4.11 Testi i Heteroskedasticitetit për modelin VAR..... | 88 |
| Tabela 5.4.12 Analiza Asimptotike (Përafrimi Edgeworth) për identifikimin e përgjigjes përkatëse të impulsit për çdo seri..... | 90 |
| Tabela 5.4.13 Rezultatet e Metodës Monte Carlo, për 1000 rizgjedhje për identifikimin e përgjigjeve përkatësisht për çdo seri..... | 95 |
| Tabela 5.4.14 Versioni Bootstrap në Modeli VAR..... | 100 |

SHKURTIME

S.E- (Standard Errors)- Devijimet Standarde

S.E (EW)-Edgeworth's Standard Errors- Devijimet Standarde për Edgeworth

S. E (Boot)-Standard Errors for Bootstrap-Devijimet Standarde për Bootstrap

S. E EW(boot)- Standard Errors for Edgeworth Bootstrap Version- Devijimet Standarde për Edgeworth-in e Versionit Bootstrap

VAR-Vector of Autoregression-Vektori i Autoregresit

GJUHËT E PROGRAMIMIT DHE PAKETAT E PËRDORURA

Në punimin e kësaj teze doktore, krahas pjesës teorike e cila ka një rëndësi të konsiderueshme, pjesa që zë një peshë esenciale është zbatimi praktik dhe identifikimi konkret i performancës së Zbërthimit Edgeworth dhe Metodave të rizgjedhjes në realitet. Gjithashtu, për të analizuar rastet praktike studimore, është e nevojshme përdorimi i gjuhëve të programimit dhe paketave të caktuara softuerike, të cilat përpunojnë një sasi të konsiderueshme të dhënash, me saktësi dhe efikasitet duke na kursyer kohë dhe përpjekje. Për të analizuar të dhënat e grumbulluara për gjithë studimet konkrete të kapitujve të kësaj teze, janë përdorur R-Programming dhe Eviews 10.

EViews është një program, i cili mundëson kryerjen e analizave të sofistikuar për të dhënat e grumbulluara, gjenerimin e shpejtë të regresit dhe kryerjen e shpejtë të prognozimeve të të dhënave. Fushat në të cilat Eviews ka zbatim të dobishëm janë: analiza dhe vlerësimi i të dhënave shkencore, realizimi i analizave financiare, parashikimi makroekonomik, simulimet dhe analiza e serive kohore. Krahas komandave të gatshme, Eviews ka dritaren përkatëse të komandave ku mund të shkruhen kodet e nevojshme për të përpunuar të dhënat, [42]. **R** është një gjuhë programimi, e cila përdoret për analizën dhe përpunimin statistikor të të dhënave. Gjuha e programimit R mundëson programimin modular duke implementuar funksione të ndryshme. Gjithashtu ky program lejon integrimin me programet e tjera si: C, C ++, Net, Python ose FORTRAN duke rritur kështu efikasitetin e studimit. Paketa R është lirisht e disponueshme nën Liçensën Publike të Përgjithshme. Krahas dritares përkatëse të komandave, tek kjo paketë ekzistojnë dhe disa skema grafike të disponueshme për përdoruesit , ([85], [86]).

Për simulimet e kryera në kapitujt vijues, janë përdorur Paketat : Boot, EW dhe EQL të Programit R, ku: Boot përfaqëson metodën e rizgjedhjes Bootstrap, EW përdoret për paraqitjen e performancës së polinomeve të Zbërthimit Edgeworth dhe EQL është paketa që përdoret për përafrimin Edgeworth. Tek programi EViews janë përdorur kodet përkatëse të gjenerimit të regresit kuantil dhe linear, si dhe hapat e caktuar për ndërtimin e modelit VAR (Vektori i Autoregresit).

Ashtu sikurse Programi R, edhe EViews mund të shkarkohen dhe të përdoren pa asnjë kosto monetare, dhe ky element përbën një avantazh krahasues. Nëse krahasojmë këto programe të përdorura, konkludohet se EViews është programi, i cili ofron fleksibilitet dhe shpejtësi më të lartë në përpunimet e të dhënave krahasuar me programin R. Ndonëse programi R është më i sofistikuar dhe ofron një diversitet më të lartë të paketave për ofrimin e të dhënave, gjenerimi i rezultateve kërkon shumë herë më tepër kohë sesa simulimi i rezultateve në EViews. Gjithashtu, programi R është shumë herë më tepër kompleks nëse krahasohet me programin EViews, pasi ky i fundit ofron një varietet metodash për gjenerimin e të njëjtës analizë të kërkuar. Për shembull për gjenerimin e regresit të thjeshtë linear mund të përdoret komanda “ls y c x”, komanda Quick Estimation of the Equation ose selektimi i të dhënave dhe zgjedhjes së vlerësimit të ekuacionit. Nga ana tjetër ekzistenca e një varieteti të gjerë paketash krijon konfuzion tek përdoruesi dhe konsiderohet si një gjuhë e vjetër programimi. Pra, në disa aspekte shihet se EViews 10 ka epërsi ndaj Programit R, nëse i referohemi efikasitetit dhe fleksibilitetit që ofron. Pra, sikurse evidentohet këto dy programe kanë avantazhe dhe disavantazhe të cilat kompensohen me njëra-tjetrën, dhe meqënëse asnjëri prej programeve nuk ka epërsi

absolute ndaj tjetrit, zgjidhja është implementimi i kombinuar i Programit R dhe Eviews 10.

HYRJE

Synimi i çdo pune kërkimore konsiston në ndërtimin e modeleve praktike me saktësi të lartë. Korrektësia optimale e një studimi është e arritshme nëpërmjet implementimit të metodave të duhura statistikore. Për të evidentuar nëse kontributi i zbërthimit Edgeworth në metodat e rizgjedhjes Bootstrap arrin të optimizojë modelet e krijuara studimore, kemi marrë në shqyrtim një sërë rastesh teorike dhe praktike, kemi shfrytëzuar burime të gjera literaturash si edhe kemi ndërmarrë studime praktike, për të konkretizuar bazën teorike të përvehtësuar. Tema e zbërthimit Edgeworth në metodat e rizgjedhjes ka patur një zbatueshmëri të limituar në praktikë, duke e kufizuar atë në simulime dhe vërtetime të kuadrit teorik. Në këtë disertacion, përveç kuadrit teorik, theksi vendoset tek implementimi i zbërthimit Edgeworth në raste konkrete, me natyrë ekonometrike për të dhënë rezultate optimale. Gjuhët e programimit të cilat i kemi përdorur në gjenerimin e simulimeve konkrete të punës janë: Programi R dhe Eviews 10. Për të analizuar efektivitetin e kësaj metode, kemi ndjekur një logjikë progresive, ku studimi nis nga vërtetime teorike të efektivitetit të përafrimit Edgeworth dhe metodave Bootstrap. Në vijim, kemi pasqyruar aplikueshmërinë e këtyre metodave në raste konkrete praktike, dhe duke treguar në këtë mënyrë ecurinë e performancës së tyre. Në kapitullin e parë kemi paraqitur një këndvështrim teorik mbi Metodën e rizgjedhjes Bootstrap dhe zbërthimit Edgeworth, ku në fund të këtij kapitulli kemi vizualizuar aspektin teorik me disa grafikë të gjeneruar prej simulimeve në programin R. Për të studiuar shkallën e saktësisë së performimit të zbërthimit Edgeworth, kemi marrë në konsideratë një këndvështrim krahasues mes shpërndarjes normale dhe zbatimit të përafrimit Edgeworth, si edhe performancën e përafrimit Edgeworth në shpërndarje të njohura statistikore. Kapitulli i dytë trajton zbërthimin Edgeworth në metodat e rizgjedhjes Bootstrap. Për të konkretizuar rëndësinë e ndërthurjes së këtyre metodave, në fund të këtij kapitulli kemi trajtuar një çështje e cila pasqyron disa simulime për të vërtetuar këtë rëndësi statistikore, duke parë sjelljen asimptotike të polinomeve pas implementimit të versionit bootstrap. Nga ana tjetër një çështje me interes është vërtetimi i korrektësisë dominuese të Metodave të rizgjedhjes Bootstrap në intervalet e besimit, për të cilat kemi marrë në shqyrtim një rast studimor lidhur me identifikimin e saktësisë në termat e Intervalit të Besimit, si edhe kemi trajtuar rastin sesi implementimi i metodave bootstrap ndikon në korrektësinë e shpërndarjeve me pjerrësi pozitive. Ky trajtim shtjellohet në kapitullin e tretë të kësaj teze disertacioni. Në dy kapitujt e fundit jemi fokusuar totalisht në identifikimin dhe zbulimin e kontributit praktik të zbërthimit dhe përafrimit Edgeworth në metodat e rizgjedhjes Bootstrap. Për të zbuluar zbatimin praktik të këtyre modeleve statistikore, kemi realizuar një lidhje mes statistikës dhe ekonometrisë. Duke qenë se ekonometria përqendrohet në analizën e fenomeneve me natyrë ekonomike duke implementuar modele statistikore, për rastet studimore të kapitullit të tretë dhe të katërt kemi shfrytëzuar të dhëna reale të aspektit ekonomik në përgjithësi dhe financiar në veçanti për biznese konkrete kombëtare dhe ndërkombëtare. Meqënëse thelbi i studimeve konkrete ka qenë rritja e saktësisë së modelit të ndërtuar, me interes është zbulimi se cila prej metodave të: Përafrimit Edgeworth për ndryshoret e pavarura dhe identifikimit të shpërndara, metodave Bootstrap apo përafrimit Edgeworth për versionin Bootstrap, kontribuon më tepër në rritjen e saktësisë së modelit. Për të nxjerrjen e konkluzioneve, kemi studiuar tendencat e devijimeve standarde të cilat nënkuptojnë termin e gabimit, dhe metoda e cila arrin të reduktojë këto devijime konsiderohet metoda më e mirë statistikore.

Gjithashtu, një këndvështrim tjetër studimor që kemi marrë në studim konsiston në identifikimin nëse kontributi i kombinuar i zbërthimit Edgeworth, versionit Bootstrap për mbetjet, si edhe simulimeve Monte Carlo, arrin të japë një prognozë të saktë të serive kohore të analizuara nëpërmjet algoritmit parashikues VAR (Vektori i Autoregresit), për të cilin kemi trajtuar rastin kur dy ose më tepër seri influencojnë njëra-tjetrën. Gjithë rezultatet e studimeve përkatëse të dy kapitujve të fundit i kemi pasqyruar në forma tabelare dhe grafike për të konkretizuar performancën dhe kontributin e zbërthimit Edgeworth në metodat e rizgjedhjes Bootstrap.

KAPITULLI I

METODAT E RIZGJEDHJES DHE ZBËRTHIMI EDGEWORTH

1.1 Metodatat e rizgjedhjes

1.1.1 Metoda Cross-validation

Metoda Cross-validation është një metodë rizgjedhjeje e cila funksionon në këtë mënyrë: jemi të interesuar për një model regresi të përshtatshëm të një bashkësie të dhënash, por nuk jemi të sigurt se cili është modeli i përshtatshëm. P.sh. cilën ndryshore parashikuese të përfshijmë, të bëjmë ose jo transformime logaritmike tek ndryshorja e përgjigjes, nëse po cilat veprimet të përfshijmë, etj. Bashkësia e të dhënave ndahet në dy gjysma dhe gjysma e parë përdoret për modelin e përshtatshëm. Ndërsa tek gjysma e dytë përdoret cross-validation: modeli i regressit i përshtatur tek gjysma e parë e të dhënave është përdorur për të parashikuar gjysmën e dytë. Parashikimet e gjysmës së parë janë tepër optimiste sepse modeli është zgjedhur si i përshtatshëm nga gjysma e parë e të dhënave. Nuk është e nevojshme të bashkësia e të dhënave të ndahet në dy pjesë të barabarta. Zakonisht lihet jashtë nga një e dhënë, përshtatet modeli me të dhënat e mbetura dhe shohim se sa mirë modeli i përshtatur parashikon të dhënë e përjashtuar. Mesatarja e gabimeve të parashikimit, ku çdo e dhënë lihet jashtë nga një herë është matja cross-validation e gabimit të parashikimit (shih [79], [45], [83], [36]).

1.1.2 Metodatat Jackknife

Jackknife si një metodë rizgjedhjeje u zbulua nga Maurice Quenouille (1949), [73] dhe më vonë u zhvillua nga John W. (1958). Kjo metodë rizgjedhjeje përdoret për përmirësimin e zhvendosjes së vlerësimeve, duke e zvogëluar atë, mundësisht deri në zero. Ideja e metodës jackknife është përdorur në fillim për vlerësimin e pritjes matematike e dispersionit dhe më vonë për vlerësime të çfardoshme. Vlerësimet jackknife, alternativa të vlerësimeve parametrike të njohura, janë metoda joparametrike që shërbejnë për vlerësimin e parametrave si zhvendosja, varianca etj. Historikisht fillimet e tyre i takojnë vlerësimit jackknife të Quenouille për zhvendosjen dhe më pas Tukey për variancën.

1.1.3 Metoda Bootstrap

Metoda "Bootstrap" nënkupton se një zgjedhje në dispozicion krijon shumë nënzgjedhje të tjera me të njëjtin vëllim me përsëritje. Krahasuar me teknikën e Jackknife, strategjia e rizgjedhjeve Bootstrap është më e plotë në aspektin e madhësisë së replikimit. Në metodën Jackknife, numri i zgjedhjeve është i përcaktuar nga numri i vrojtimit (n-1). Duke përdorur resurset kompjuterike ne mund të marrim rizgjedhje sa të duam ne nga zgjedhja në shqyrtim.

Duke përdorur metodat Bootstrap, ne jemi në gjendje të vlerësojmë intervalet e besimit për parametrin e interesit. Supozojmë se zgjedhja jonë origjinale është në mënyrë të arsyeshme përfaqësuese e popullatës nga e cila vjen. Ne rastësisht marrim një zgjedhje me zëvendësim nga rezultatet e vrojtuar për të prodhuar një zgjedhje të re të të njëjtës madhësi me zgjedhjen tonë origjinale. Tani mund të llogaritim statistikën e interesit (p.sh., mesatarja) nga zgjedhja e re e përftuar. Me një numër të madh të zgjedhjeve të reja, të paktën 10000, ne gjenerojmë një shpërndarje empirike të zgjedhjeve të përftuar për statistikën e interesit dhe më pas mund të përcaktojmë intervalet e besimit për këtë statistikë. Nëse kemi dy grupe, ne mund të gjenerojmë zgjedhje bootstrap nga secili grup veçmas dhe të llogaritim statistikën e interesit (p.sh.

diferenca mes medianave, një vlerë t-test ose një ndryshim në variancë). Me replikime të shumëfishta, ne gjenerojmë një shpërndarje të zgjedhjeve për statistikën e interesit (shih [33], [34], [36], [37]).

Principi bazë i bootstrap

Le të jenë dhënë X_1, \dots, X_n , n ndryshore rasti të pavarura, me shpërndarje të njëjtë por të panjohur F . Algoritmi i rizgjedhjes bootstrap është si më poshtë:

Hapi I Nga zgjedhja e marrë llogarisim vlerësuesin $\hat{\theta}$ për parametrin e panjohur θ .

Hapi II Ndërtojmë shpërndarjen empirike \hat{F} , duke i vendosur probabilitetin $\frac{1}{n}$ secilës prej vlerave të zgjedhjes.

Hapi III Nga shpërndarja \hat{F} , marrim zgjedhjen X_1^*, \dots, X_n^* , që quhet rizgjedhje bootstrap (kjo është zgjedhje me kthim).

Hapi IV Përafrojmë shpërndarjen e $\hat{\theta}$ me shpërndarjen e $\hat{\theta}^*$, që merret nga zgjedhja bootstrap.

Hapi V Përsërisim dy hapat e mësipërm një numër B -herësh duke përfutur vlerësimet bootstrap $\hat{\theta}_1^*, \dots, \hat{\theta}_B^*$.

Hapi VI Mbështetur tek metoda Monte Carlo, vlerësimi pikësor bootstrap për parametrin e panjohur θ ka trajtën:

$$\hat{\theta}^* = \frac{1}{B} \sum_{b=1}^B \hat{\theta}_b^* .$$

Parimisht të trija metodat cross-validation, bootstrap dhe jackknife janë shumë të ngjashme por metodat bootstrap janë më të përdorshme për shkak se ato janë më të plota, në kuptimin që mund të marrim më shumë nënzgjedhje krahasuar me metodat e tjera. Mooney dhe Duval (1993), [69] sugjeruan se Jackknife është me interes të madh historik por nuk gjen shumë përdorim në ditët e sotme.

Sidoqoftë, mbështetur tek metodat e prezantuara në këtë material, jackknife dhe bootstrap spikasin si dy metoda rizgjedhje ku janë fokusuar pjesa më e madhe e studiuesve, kjo për mjaft arsye, por sidomos për lehtësinë e përdorimit të këtyre metodave dhe përparimin dhe mundësitë që ofrohen nga programe të ndryshme dhe gjuhë programimi për zbatimin e këtyre teknikave. Ne gjatë simulimeve të realizuara në kapitujt e mëposhtëm do të fokusohemi vetëm tek metodat bootstrap.

1.2 Zbërthimi Edgeworth

Në këtë pjesë do të shohim zbërthimin Edgeworth si një përafrim për shpërndarjen e vlerësuesit $\hat{\theta}$ të parametrin të panjohur θ_0 . Shkurtimisht, nëse $\hat{\theta}$ është ndërtuar nga një zgjedhje me vëllim n , dhe nëse $n^{1/2}(\hat{\theta} - \theta_0)$ asimptotikisht ka shpërndarje Normale me pritje matematike zero dhe variancë σ^2 , atëherë funksioni i shpërndarjes i $n^{1/2}(\hat{\theta} - \theta_0)$ mund të zbërthehet si një seri e fuqisë $n^{-1/2}$,

$$P(n^{1/2}(\hat{\theta} - \theta_0) / \sigma \leq x) = \Phi(x) + n^{-1/2} p_1(x)\phi(x) + \dots + n^{-j/2} p_j(x)\phi(x) + \dots, \quad (1.2.1)$$

ku $\phi(x) = (2\pi)^{-1/2} e^{-x^2/2}$ është funksioni i densitetit Normal Standard dhe $\Phi(x) = \int_{-\infty}^x \phi(u) du$ është funksioni i shpërndarjes Normale Standarde. Formula (1.2.1)

njihet si zbërthimi Edgeworth, [50]. Funkzionet p_j janë polinome me koeficientë që varen nga kumulantët e $\hat{\theta} - \theta_0$. Konceptualisht, ne e mendojmë zbërthimin (1.2.1) si fillim pas përafrimit Normal $\Phi(x)$, kështu që $n^{-1/2} p_1(x)\phi(x)$ është termi i parë dhe jo i dyti i këtij zbërthimi.

Ekuacioni (1.2.1) mund të kthehet në një tjetër formë, duke treguar kështu që zgjidhja e $x = u_\alpha$ e ekuacionit

$$P\{n^{1/2}(\hat{\theta} - \theta_0) / \sigma \leq x\} = \alpha$$

lejon një zbërthim

$$u_\alpha = z_\alpha + n^{-1/2} p_{11}(z_\alpha) + n^{-1} p_{21}(z_\alpha) + \dots + n^{-j/2} p_{j1}(z_\alpha) + \dots, \quad (1.2.2)$$

ku p_{j1} janë polinomet e derivuara nga polinomet p_j dhe z_α është zgjidhja e ekuacionit $\Phi(z_\alpha) = \alpha$. Formula inverse si ajo e (1.2.2) njihet si zbërthimi Cornish-Fisher ([15], [44]).

Në këtë diskutim të shkurtër, ne u fokusuar në shpërndarjen e statistikës jo të Studentizuar, $n^{1/2}(\hat{\theta} - \theta_0) / \sigma$. Por, në mjaft raste me interes praktik, varianca asimptotike σ^2 mund të jetë e panjohur dhe të vlerësohet nga një funksion $\hat{\sigma}^2$ i të dhënave. Në këtë rast, jemi të interesuar për funksionin e shpërndarjes të statistikës së Studentizuar $n^{1/2}(\hat{\theta} - \theta_0) / \hat{\sigma}$. Trajtimi ynë i jep theks këtij rasti, i cili lejon një zbërthim si ai (1.2.1), por me polinome p_j të ndryshme.

Për t'i dalluar këto dy llojet e polinomeve, ne duhet të përdorim polinomet p_j në rastin jo të Studentizuar dhe polinomet q_j në rastin e Studentizuar.

Shuma e ndryshoreve të rastit të pavarura

Chebyshev [13] dhe Edgeworth ([31], [32]) krijuan idenë e zbërthimit të një funksioni shpërndarjeje sipas zbërthimit Edgeworth. Ata diskutuan ndërtimin formal të zbërthimit në rastin e shumës së ndryshoreve të rastit të pavarura. Ndërkohë, Cramer ([16], [17]) ishte ai që zhvilloi një teori rigoroze dhe të detajuar për atë rast.

Le të jenë X_1, X_2, \dots, X_n ndryshore rasti të pavarura dhe identikisht të shpërndara me pritje μ dhe variancë σ^2 . Nga Teorema Qendrore Limite [48],

$$S_n = \frac{\sum_{i=1}^n X_i / n - \mu}{\sigma / \sqrt{n}}$$

asimptotikisht ka shpërndarje normale me pritje matematike 0 dhe variancë 1. Ajo për të cilën jemi të interesuar të studiojmë është sjellja asimptotike e diferencës ndërmjet shpërndarjes normale $\Phi(x)$ dhe funksionit të shpërndarjes $F_n(x)$ i S_n . Me fjalë të tjera, duam të përshkruajmë gabimin e përafrimit normal dhe një nga mënyrat për ta bërë këtë është nëpërmjet funksionit karakteristik, [12].

Nëse funksioni karakteristik $\psi_X(t)$ plotëson kushtin

$$\lim_{|t| \rightarrow \infty} \sup |\psi_X(t)| < 1, \quad (C)$$

të quajtur kushti i Cramér-it, atëherë funksioni i shpërndarjes ka zbërthimin që vijon

$$F_n(x) = \Phi(x) + \sum_{j=1}^s \frac{p_j(x)}{n^{j/2}} e^{-\frac{x^2}{2}} + O\left(\frac{1}{n^{(s+1)/2}}\right), \quad s \geq 1$$

ku $p_j(x)$ është polinom i x .

Shënim. Kur $s = 1, O(\frac{1}{n}) \ll o(\frac{1}{\sqrt{n}})$, që do të thotë se kur (C) plotësohet, gabimi është shumë herë më i vogël sesa kur ai nuk plotësohet.

Vërejtje 1. Ne qoftë se funksioni i shpërndarjes së X është absolutisht i vazhdueshëm, pra X është një ndryshore rasti e vazhdueshme, atëherë ka vend kushti (C).

Vërejtje 2. Kolmogorov [46] vuri në dukje se ekzistojnë shpërndarje diskrete të cilat nuk janë laticë. Për shembull, nëse X_j merr vetëm vlerat ± 1 dhe $\pm\sqrt{3}$, secila me probabilitet $\frac{1}{4}$, atëherë shpërndarja e saj nuk është laticë sepse sistemi i ekuacioneve

$$\begin{aligned} -\sqrt{3} &= a + k_1 h \\ 1 &= a + k_2 h \\ 1 &= a + k_3 h \\ -\sqrt{3} &= a + k_4 h \end{aligned}$$

ku $k_i \in \{\pm 1, \pm 2, \dots\}$, nuk ka zgjidhje.

Meqë S_n është asimptotikisht Normale $N(0,1)$, funksioni karakteristik ψ_n i S_n konvergjon tek $N(0,1)$ kur $n \rightarrow \infty$,

$$\psi_{S_n}(t) = E[\exp(itS_n)] \rightarrow E[\exp(itN)] = e^{-t^2/2}, \quad -\infty < t < \infty. \quad (1.2.3)$$

Tani,

$$\psi_{S_n}(t) = (\psi_Y(t/n^{1/2}))^n, \quad (1.2.4)$$

ku ψ_Y është funksioni karakteristik i $Y = (X - \mu)/\sigma$.

Gjithashtu, ne mund ta shkruajmë funksionin karakteristik si një zbërthim

$$\log \psi_Y(t) = \sum_{j=1}^{\infty} \frac{k_j(it)^j}{j!}.$$

Atëherë

$$\psi_Y(t) = \exp\left\{k_1 it + \frac{1}{2} k_2 (it)^2 + \dots + \frac{1}{j!} k_j (it)^j + \dots\right\}. \quad (1.2.5)$$

Veç kësaj, nga funksioni karakteristik i zhvilluar tek seria Maclaurin për vlera të vogla, kemi që

$$\psi_Y(t) = 1 + E(Y)it + \frac{1}{2} E(Y^2)(it)^2 + \dots + \frac{1}{j!} E(Y^j)(it)^j + \dots$$

Kumulantët k_j mund t'i përcaktojmë duke përdorur identitetet formale

$$\begin{aligned}\sum_{j \geq 1} \frac{1}{j!} k_j (it)^j &= \log \left\{ 1 + \sum_{j \geq 1} \frac{1}{j!} E(Y^j) (it)^j \right\} \\ &= \sum_{k \geq 1} (-1)^{k+1} \frac{1}{k} \left\{ \sum_{j \geq 1} \frac{1}{j!} E(Y^j) (it)^j \right\}^k.\end{aligned}$$

Duke barazuar koeficientët e $(it)^j$, nxjerrim që

$$\left. \begin{aligned}k_1 &= E(Y), \\ k_2 &= E(Y^2) - (EY)^2 = \text{var}(Y), \\ k_3 &= E(Y^3) - 3E(Y^2)E(Y) + 2(EY)^3 = E(Y - EY)^3, \\ k_4 &= E(Y^4) - 4E(Y^3)E(Y) - 3(EY^2)^2 + 12E(Y^2)(EY)^2 - 6(EY)^4 \\ &= E(Y - EY)^4 - 3(\text{var } Y)^2,\end{aligned} \right\} \quad (1.2.6)$$

e kështu me rradhë, formula për k_{10} përmban 41 terma të tilla. Shënojmë që, k_j është polinom homogjen i momenteve të shkallës j . Po kështu, edhe momentet mund të shprehen si polinome homogjene të kumulantëve, ([12], [80]).

Ne standartizuar ndryshoren e rastit Y , kështu që $E(Y) = k_1 = 0$ dhe $\text{var}(Y) = k_2 = 1$. Pra nga (1.2.4) dhe (1.2.5), duke përdorur zberthimin e serisë të funksionit eksponencial,

$$\begin{aligned}\psi_{S_n}(t) &= \exp \left\{ \frac{-t^2}{2} + n^{-1/2} \frac{k_3(it)^3}{3!} + \dots + n^{-(j-2)/2} \frac{k_j(it)^j}{j!} + \dots \right\}, \\ \psi_{S_n}(t) &= e^{-t^2/2} \left\{ 1 + n^{-1/2} r_1(it) + n^{-1} r_2(it) + \dots + n^{-j/2} r_j(it) + \dots \right\},\end{aligned} \quad (1.2.7)$$

ku r_j është një polinom i shkallës $3j$ që varet nga kumulantët por jo nga n , dhe ky zberthim jep të njëjtën konvergjencë me atë të Teoremës Qendrore Limite (1.2.3). Polinomi r_j është çift kur j është çift dhe është tek kur j është tek. Nga (1.2.7) është e qartë se

$$r_1(u) = \frac{1}{6} k_3 u^3 \quad (1.2.8)$$

dhe

$$r_2(u) = \frac{1}{24} k_4 u^4 + \frac{1}{72} k_3^2 u^6. \quad (1.2.9)$$

E rishkruajmë (1.2.7) në formën

$$\psi_{S_n}(t) = e^{-t^2/2} + n^{-1/2} r_1(it) e^{-t^2/2} + n^{-1} r_2(it) e^{-t^2/2} + \dots + n^{-j/2} r_j(it) e^{-t^2/2} + \dots \quad (1.2.10)$$

Meqenëse

$$\psi_{S_n}(t) = \int_{-\infty}^{\infty} e^{itx} dP(S_n \leq x)$$

dhe

$$e^{-t^2/2} = \int_{-\infty}^{\infty} e^{itx} d\Phi(x), \quad (1.2.11)$$

ku Φ tregon funksionin e shpërndarjes Normale Standarde, në qoftë se ne aplikojmë metodën inverse të transformimit Fourier-Stieltjes, tek (1.2.12) do të kemi

$$P(S_n \leq x) = \Phi(x) + n^{-1/2}R_1(x) + n^{-1}R_2(x) + \dots + n^{-j/2}R_j(x) + \dots, \quad (1.2.12)$$

ku $R_j(x)$ paraqet një funksion, transformimi Fourier-Stieltjes i të cilit është i barabartë me $r_j(it)e^{-t^2/2}$

$$\int_{-\infty}^{\infty} e^{itx} dR_j(x) = r_j(it)e^{-t^2/2}.$$

Hapi i radhës në këtë argument është llogaritja e R_j .

Duke zbatuar integrimin me pjesë j herë tek (1.2.11), marrim

$$\begin{aligned} e^{-t^2/2} &= (-it)^{-1} \int_{-\infty}^{\infty} e^{itx} d\Phi^{(1)}(x) \\ &= (-it)^{-2} \int_{-\infty}^{\infty} e^{itx} d\Phi^{(2)}(x) \\ &\vdots \\ &= (-it)^{-j} \int_{-\infty}^{\infty} e^{itx} d\Phi^{(j)}(x) \end{aligned}$$

ku $\Phi^{(j)}(x) = (d/dx)^j \Phi(x)$. Prandaj

$$\int_{-\infty}^{\infty} e^{itx} d\{(-D)^j \Phi(x)\} = (it)^j e^{-t^2/2}, \quad (1.2.13)$$

ku D është operatori diferencial d/dx . Edhe $r_j(-D)$ është gjithashtu operator diferencial në vetvete. Nga (1.2.13)

$$\int_{-\infty}^{\infty} e^{itx} d\{r_j(-D)\Phi(x)\} = r_j(it)e^{-t^2/2},$$

dhe $r_j(-D)\Phi(x)$ është funksioni që po kërkojmë këtu

$$R_j(x) = r_j(-d/dx)\Phi(x). \quad (1.2.14)$$

Për $j \geq 1$, $(-D)^j \Phi(x) = -He_{j-1}(x)\phi(x)$, ku funksionet He_j janë polinomet Hermitiane:

$$\begin{aligned} He_0(x) &= 1, & He_1(x) &= x, & He_2(x) &= x^2 - 1, \\ He_3(x) &= x(x^2 - 3), & He_4(x) &= x^4 - 6x^2 + 3, \\ He_5(x) &= x(x^4 - 10x^2 + 15), \dots \end{aligned}$$

Shih ([68], [1], [71]).

Shohim se polinomet Hermitiane janë ortogonale në lidhje me $\phi(x)$ dhe He_j është një polinom i shkallës j dhe është çift kur j është çift dhe është tek kur j është tek. Prej (1.2.8), (1.2.9) dhe (1.2.14) nxjerrim se

$$R_1(x) = -\frac{1}{6}k_3(x^2 - 1)\phi(x)$$

dhe

$$R_2(x) = -x \left\{ \frac{1}{24}k_4(x^2 - 3) + \frac{1}{72}k_3^2(x^4 - 10x^2 + 15) \right\} \phi(x).$$

$$R_j(x) = p_j(x)\phi(x), \text{ për } j \geq 1.$$

Polinomi p_j është i shkallës $3j-1$ dhe është tek për j çift, dhe është çift për j tek. Kjo rrjedh nga fakti se r_j është i shkallës $3j$ dhe është çift për j çift, tek për j tek. ([4], [11])

Kështu që

$$p_1(x) = -\frac{1}{6}k_3(x^2 - 1) \quad (1.2.15)$$

dhe

$$p_2(x) = -x \left\{ \frac{1}{24}k_4(x^2 - 3) + \frac{1}{72}k_3^2(x^4 - 10x^2 + 15) \right\}. \quad (1.2.16)$$

Tani formula (1.12) merr formën

$$P(S_n \leq x) = \Phi(x) + n^{-1/2}p_1(x)\phi(x) + n^{-1}p_2(x)\phi(x) + \dots + n^{-j/2}p_j(x)\phi(x) + \dots, \quad (1.2.17)$$

e cila quhet zbërthimi Edgeworth i funksionit të shpërndarjes $P(S_n \leq x)$.

Kumulanti i tretë k_3 dhe i katërt k_4 i referohen pjerrësisë dhe sheshtësisë, respektivisht. Termi i rendit $n^{-1/2}$ tek (1.2.17) korrekton përafrimin Normal duke u fokusuar kryesisht tek përmirësimi i efektit të pjerrësisë, ndërsa termi i rendit n^{-1} korrekton kryesisht efektin e sheshtësisë dhe tërthorazi efektin e pjerrësisë.

Zbërthimi (1.2.17), jo gjithmonë konvergjon si seri e pafundme, [50]. Vërtet, nëse X ka shpërndarje absolutisht të vazhdueshme, atëherë kushti strikt që duhet vendosur për të siguruar konvergjencën është $E \left\{ \exp\left(\frac{1}{4}Y^2\right) \right\} < \infty$, [18].

Zakonisht (1.2.17) jepet si një seri asimptotike ose si një zbërthim asimptotik, që do të thotë se n.q.s seria ndalon pas një numri të dhënë termash, atëherë termi mbetës është i rendit më të vogël se termi i fundit i përfshirë,

$$P(S_n \leq x) = \Phi(x) + n^{-1/2}p_1(x)\phi(x) + n^{-1}p_2(x)\phi(x) + \dots + n^{-j/2}p_j(x)\phi(x) + o(n^{-j/2}). \quad (1.2.18)$$

Ky zbërthim është i vlefshëm për j të fiksuar, kur $n \rightarrow \infty$. Në fakt, termi mbetës $o(n^{-j/2})$ është shumë i vogël, domethënë $O(n^{-\frac{j+1}{2}})$, [41].

Kufizimet tek (1.2.18) janë

$$E(|X|^{j+2}) < \infty \text{ dhe } \lim_{|t| \rightarrow \infty} \sup |\psi(t)| < 1.$$

Vërtetimin e këtij fakti mund ta gjeni tek [50].

Zbërthimin Edgeworth e përftuam nga një zbërthim për logartimin e funksionit karakteristik të S_n . Kushti Cramér (C) na siguron që funksioni karakteristik mund të zbërthehet në mënyrën e kërkuar. Zbërthimi i F_n u përftua nga e anasjellta e transformimit Fourier i zbërthimit për funksionin karakteristik.

1.3 Simulime të Zbërthimit Edgeworth për ndryshoren e rastit të vazhdueshme eksponenciale

Le të shohim tani saktësinë e zbërthimit Edgeworth. Do ta aplikojmë atë tek ndryshoret e rastit eksponenciale që janë ndryshore rasti të vazhdueshme dhe të krahasojmë mesataren e standardizuar të ndryshores së rastit eksponenciale me përafrimin normal, shih [23].

Kujtojmë që ndryshorja e rastit e vazhdueshme X ka shpërndarje eksponenciale nëse densiteti i saj ka formën

$$f(x) = \begin{cases} \lambda \exp(-\lambda x), & x \geq 0 \\ 0 & , \quad x < 0 \end{cases}, \text{ për } \lambda > 0$$

dhe funksioni i shpërndarjes

$$F(x) = \begin{cases} 1 - \exp(-\lambda x), & x \geq 0 \\ 0 & , \quad x < 0 \end{cases}.$$

Ne kemi zgjedhur përkatësisht 5, 15, 30 ndryshore rasti të pavarura dhe identikisht të shpërndara (eksponencialisht), me $\lambda = 2$, pra pritja është $\frac{1}{2}$ dhe kemi bërë simulime 100, 1000, 10000 herë duke përdorur metodën Monte Carlo për të gjeneruar rezultatet eksponenciale të rastit. Simulimet i kemi bërë në gjuhën e programimit R, ku edhe kemi ndërtuar kodin përkatës. Prej figurave më poshtë mund të shohim se përafrimi Edgeworth jep vlerësim më të mirë për funksionin e shpërndarjes sesa përafrimi normal. Në këto figura, vijat me ngjyra blu, të gjelbër, të kuqe dhe të verdhë tregojnë shpërndarjen empirike, përafrimin normal, zbërthimin Edgeworth me dy terma dhe zbërthimin Edgeworth me tre terma, respektivisht.

Gjithashtu, mund të shohim nga figurat e mëposhtme se për vëllim të njëjtë n të zgjedhjes dhe numër simulimesh të ndryshme N , përafrimi Edgeworth është më i mirë se përafrimi normal. Përafrimi Edgeworth me tre terma është më i mirë se përafrimi Edgeworth me dy terma, me rritjen e numrit të simulimeve.

Për numër të njëjtë simulimesh, por vëllim të ndryshëm të zgjedhjes, përafrimi Edgeworth me tre terma është më i mirë se përafrimi Edgeworth me dy terma, dhe përafrimi Edgeworth me dy terma është më i mirë se përafrimi normal me rritjen e vëllimit të zgjedhjes. Pra, nga të gjitha këto figura nxjerrim përfundimin se me rritjen e numrit të simulimeve N dhe me rritjen e vëllimit të zgjedhjes n , përafrimi Edgeworth është përafrimi më i mirë që përftojmë sesa përafrimi normal.

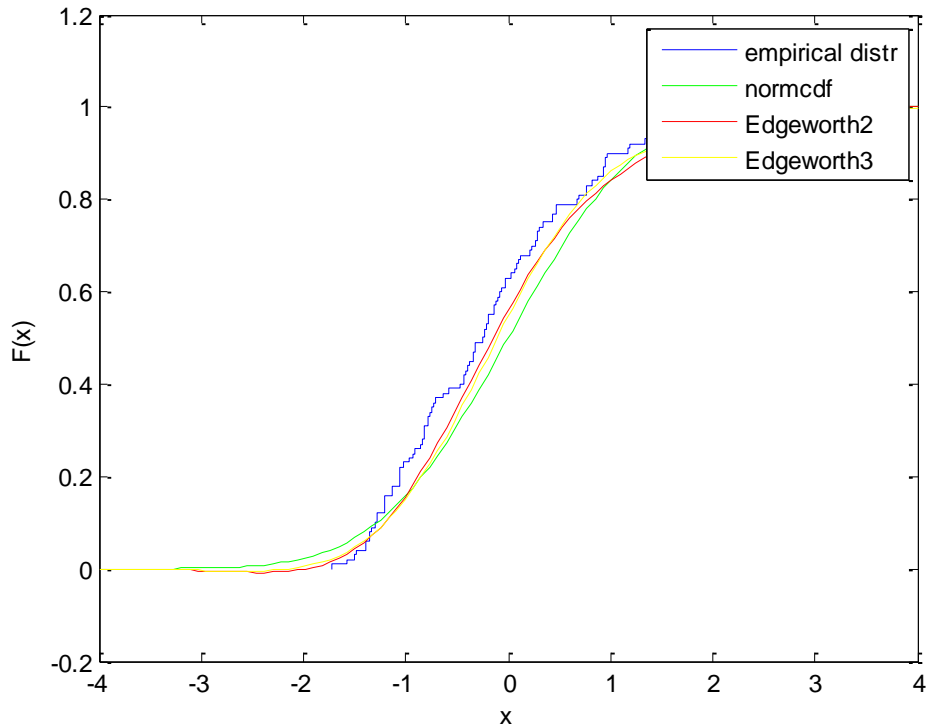


Figura 1.3.1 Shpërndarja empirike, përafrimi normal, përafrimi Edgeworth me dy terma dhe përafrimi Edgeworth me tre terma për $n=5$, $N=100$

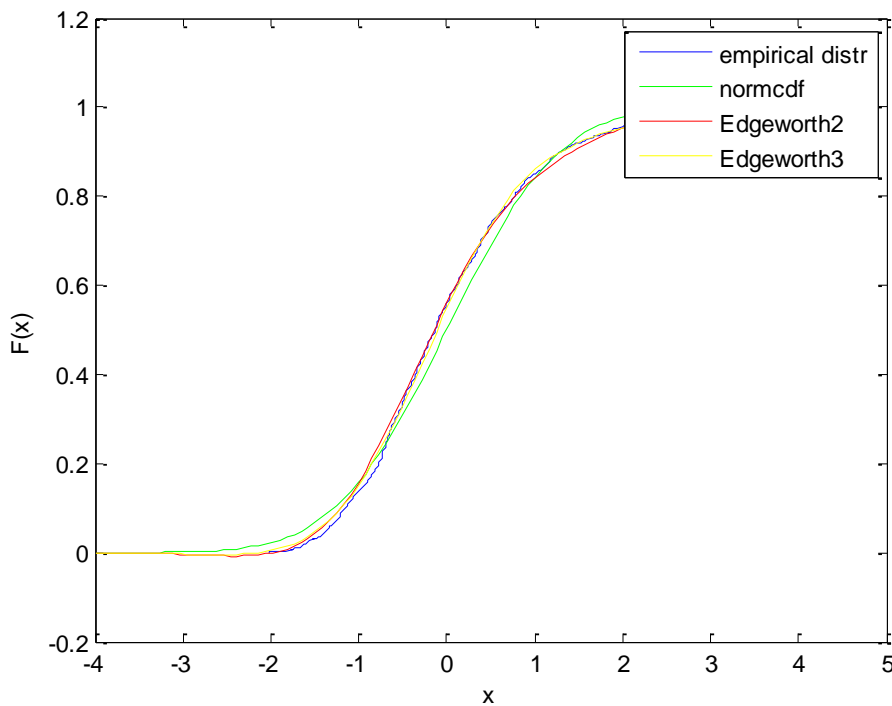


Figura 1.3.2 Shpërndarja empirike, përafrimi normal, përafrimi Edgeworth me dy terma dhe përafrimi Edgeworth me tre terma për $n=5$, $N=1000$

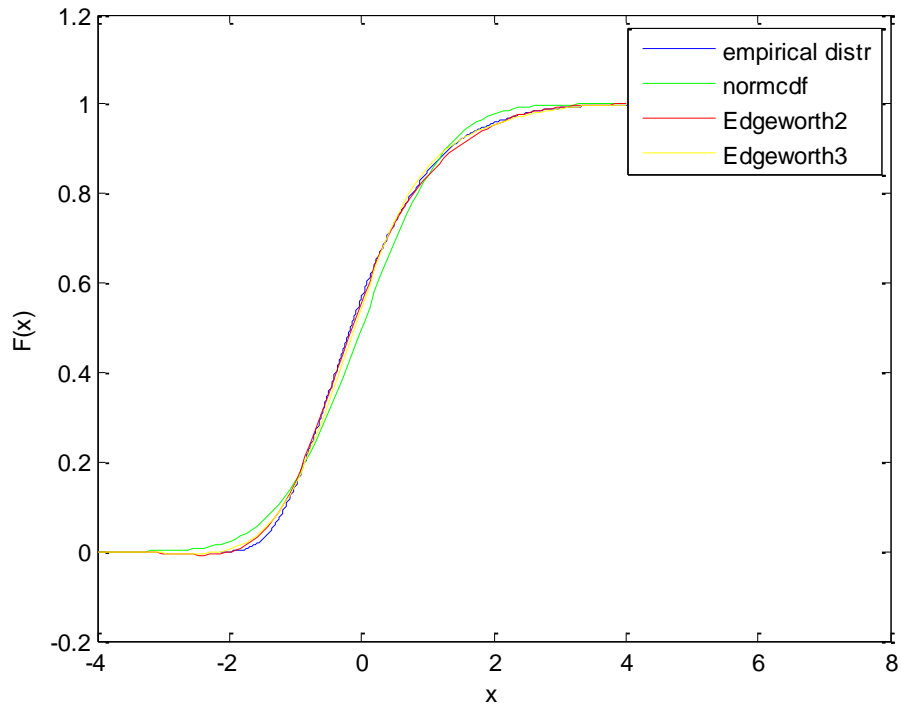


Figura 1.3.3 Shpërndarja empirike, përafrimi normal, përafrimi Edgeworth me dy terma dhe përafrimi Edgeworth me tre terma për $n=5$, $N=10000$

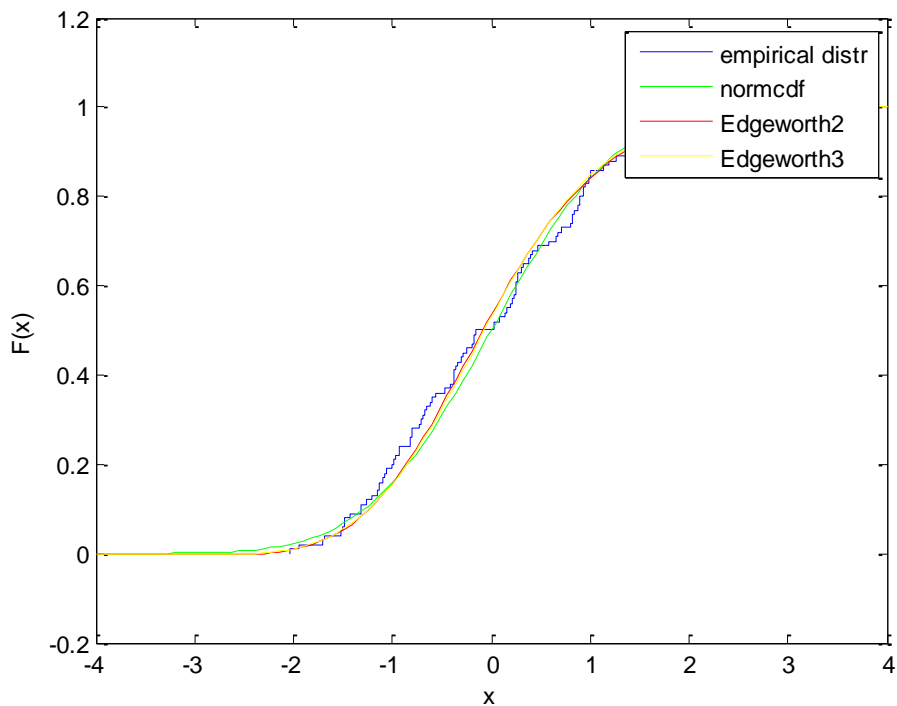


Figura 1.3.4 Shpërndarja empirike, përafrimi normal, përafrimi Edgeworth me dy terma dhe përafrimi Edgeworth me tre terma për $n=15$, $N=100$

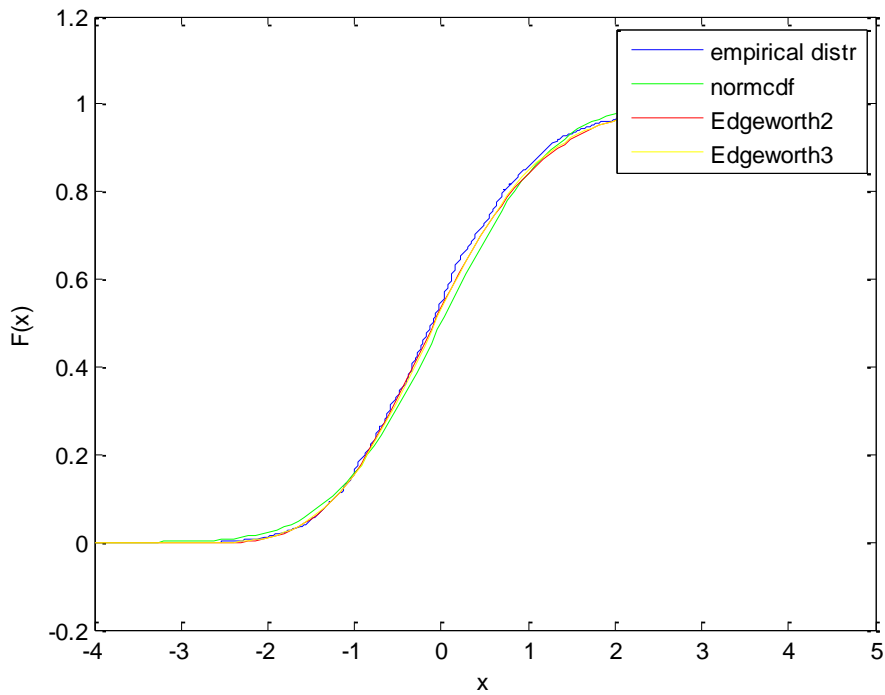


Figura 1.3.5 Shpërndarja empirike, përafrimi normal, përafrimi Edgeworth me dy terma dhe përafrimi Edgeworth me tre terma për $n=15$, $N=1000$

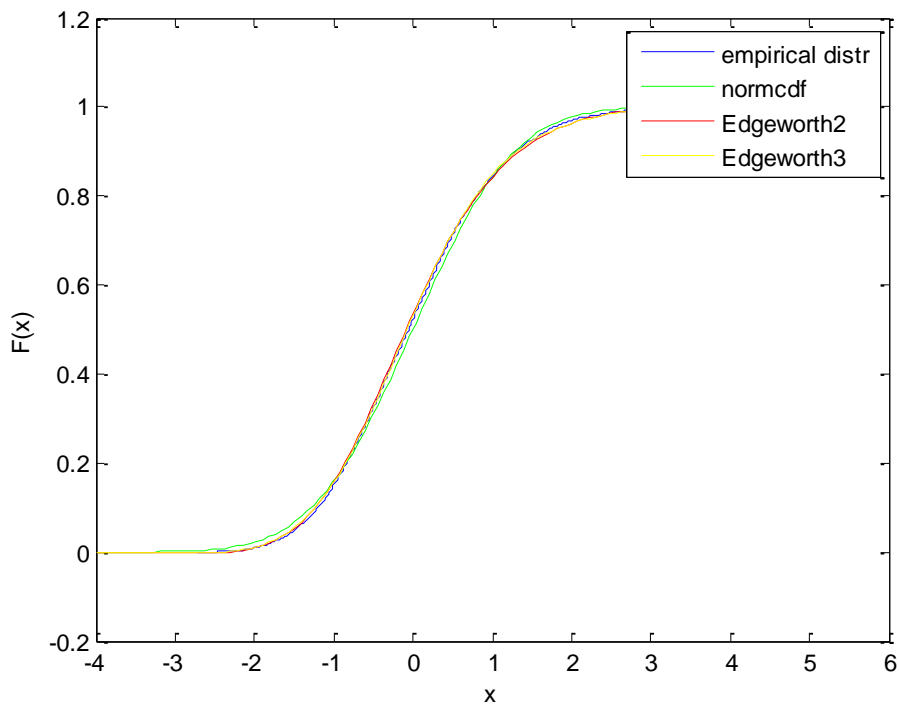


Figura 1.3.6 Shpërndarja empirike, përafrimi normal, përafrimi Edgeworth me dy terma dhe përafrimi Edgeworth me tre terma për $n=15$, $N=10000$

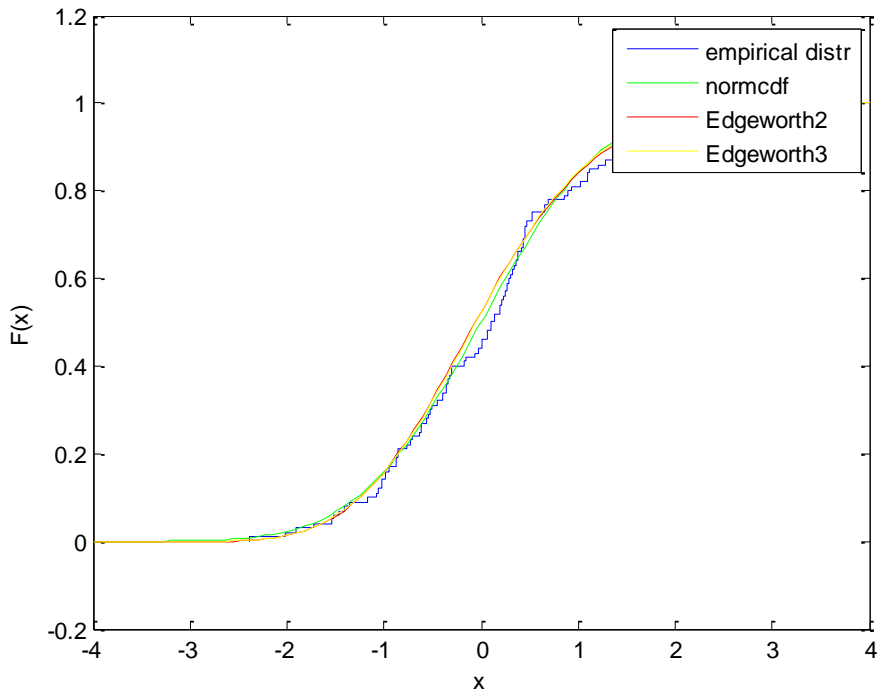


Figura 1.3.7 Shpërndarja empirike, përafrimi normal, përafrimi Edgeworth me dy terma dhe përafrimi Edgeworth me tre terma për $n=30$, $N=100$

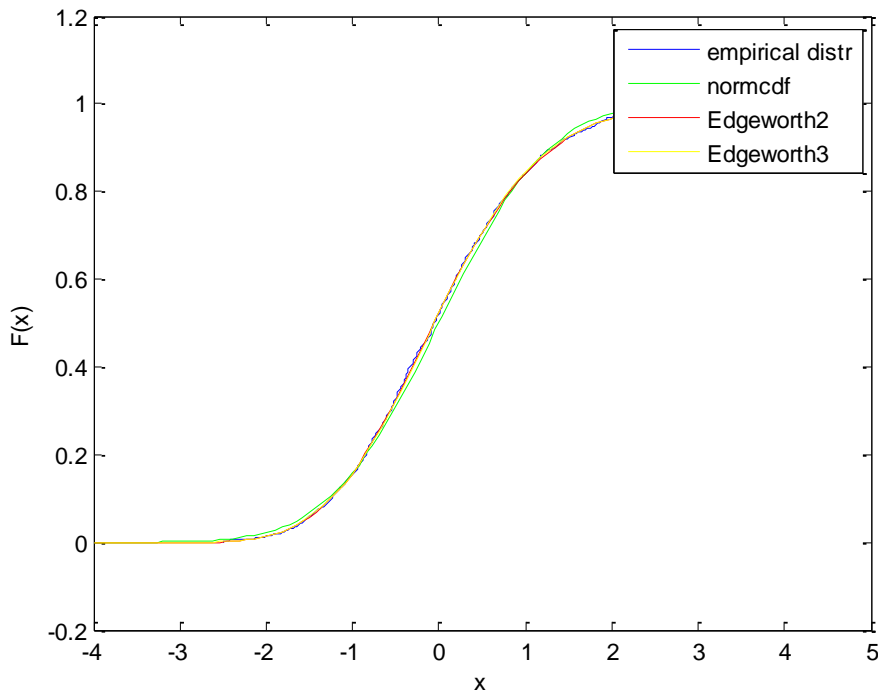


Figura 1.3.8 Shpërndarja empirike, përafrimi normal, përafrimi Edgeworth me dy terma dhe përafrimi Edgeworth me tre terma për $n=30$, $N=1000$

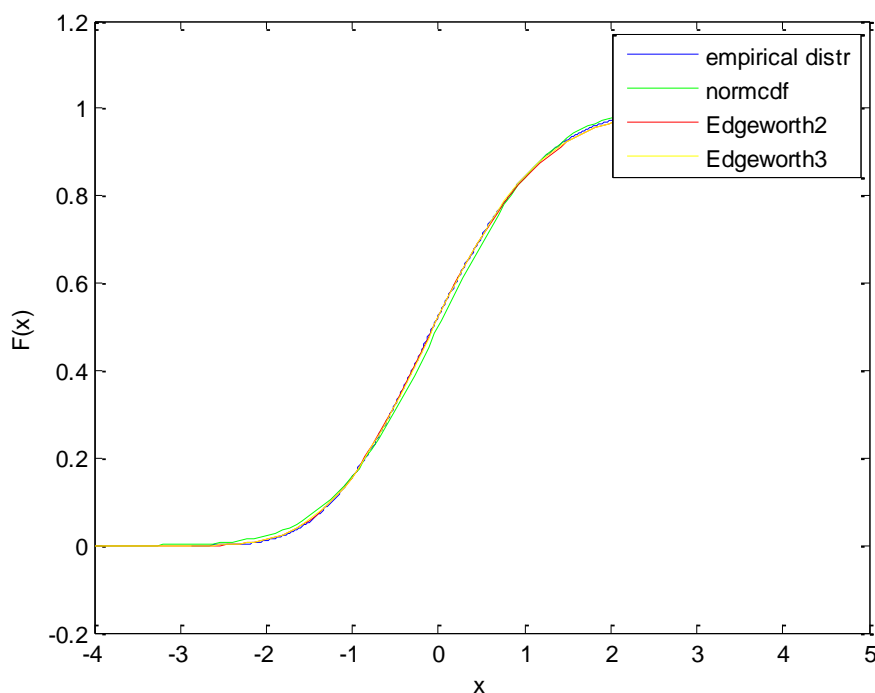


Figura 1.3.9 Shpërndarja empirike, përafrimi normal, përafrimi Edgeworth me dy terma dhe përafrimi Edgeworth me tre terma për $n=30$, $N=10000$

1.4 Përafrimi Edgeworth për shpërndarjen hi-katror

Zbërthimi Edgeworth është një teknikë përafrimi që na ndihmon të vlerësojmë funksionin e shpërndarjes të mesatares së standardizuar, mesatares apo edhe të shumës së ndryshoreve të rastit të pavarura dhe identikisht të shpërndara, [26]. Le ta shohim këtë përafrim për një shpërndarje të rëndësishme statistikore, shpërndarjen hi-katror.

Kujtojmë që shpërndarja hi-katror (shënohet χ^2) me k shkallë lirie është shpërndarja e shumës së k ndryshoreve të rastit normale standarde të pavarura në katror. Ky është një rast i veçantë i shpërndarjes gamma dhe është nga shpërndarjet e përdorura më gjerësisht në statistikë, si p.sh. kontroll hipotezash apo ndërtim intervalesh besimi (shih [58], [70]).

Në qoftë se X_1, X_2, \dots, X_k janë k ndryshore rasti të pavarura me shpërndarje normale, atëherë themi se ndryshorja V ka shpërndarje hi-katror me k shkallë lirie, shënohet

$$V = X_1^2 + X_2^2 + \dots + X_k^2 \sim \chi^2(k).$$

Kjo shpërndarje ka pritje k dhe variancë $2k$.

Duke përdorur gjuhën e programimit R dhe paketën “EQL” kemi ndërtuar kodin specifik dhe kemi bërë simulimet e nevojshme të këtij përafrimi për mesataren e n ndryshoreve hi-katror të pavarura dhe identikisht të shpërndara. Kemi shqyrtuar rastet për vëllimin e zgjedhjes $n=10$ dhe $n=100$ dhe për këto raste kemi marrë shkallët e lirisë $df=5, 10, 15$. Rezultatet e simulimeve mund t’i shohim tek figurat e mëposhtme:

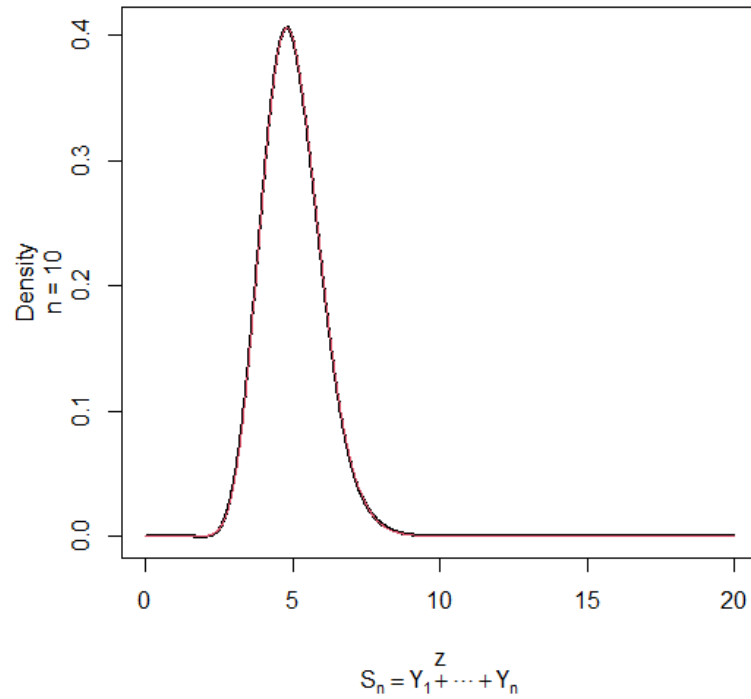


Figura 1.4.1 Përafrimi Edgeworth për mesataren e n ndryshoreve hi-katror, për $n=10$, $df=5$

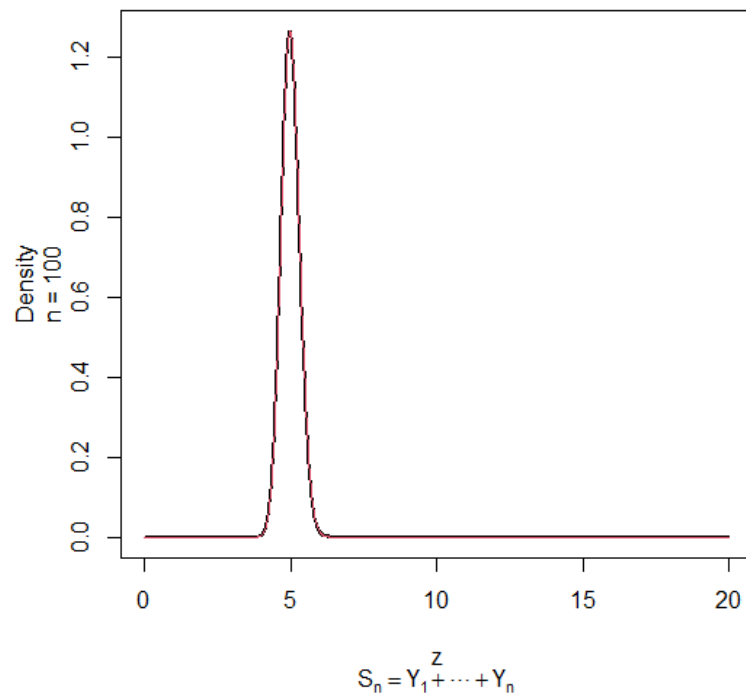


Figura 1.4.2 Përafrimi Edgeworth për mesataren e n ndryshoreve hi-katror, për $n=100$, $df=5$

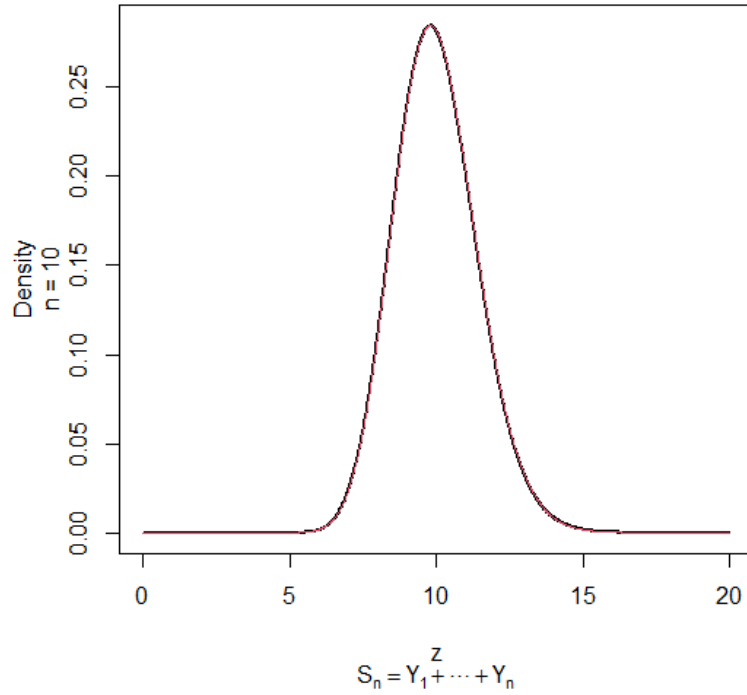


Figura 1.4.3 Përafrimi Edgeworth për mesataren e n ndryshoreve hi-katror, për $n=10$, $df=10$

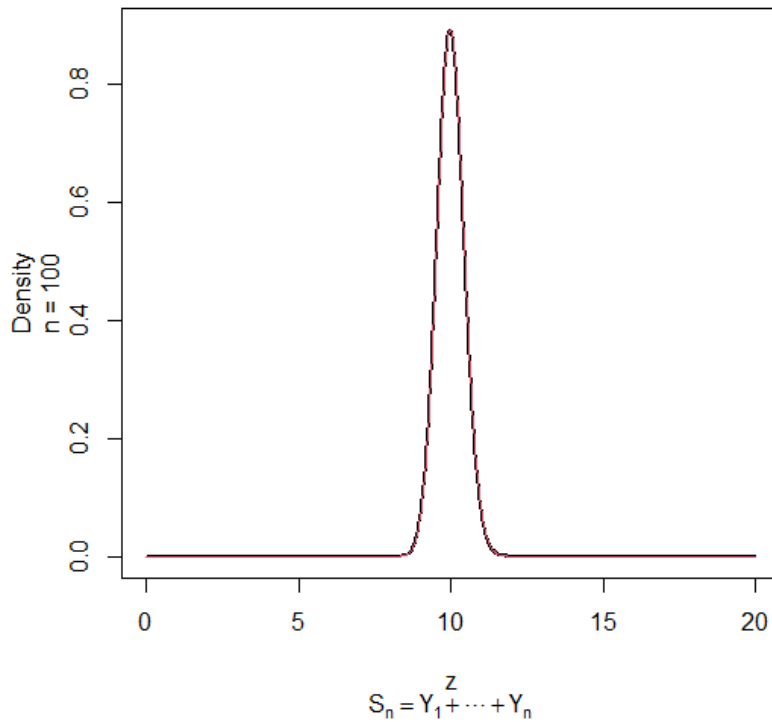


Figura 1.4.4 Përafrimi Edgeworth për mesataren e n ndryshoreve hi-katror, për $n=100$, $df=10$

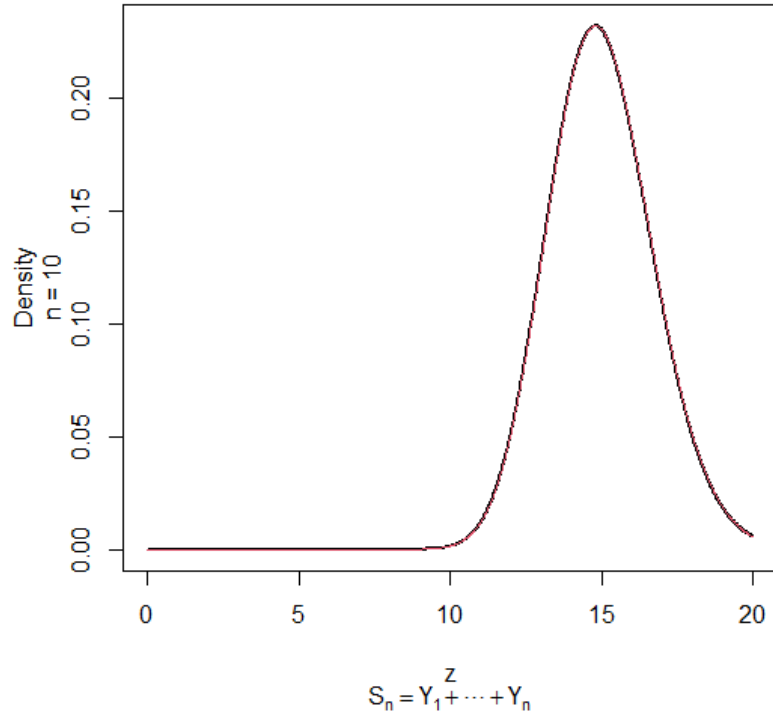


Figura 1.4.5 Përafrimi Edgeworth për mesataren e n ndryshoreve hi-katror, për $n=10$, $df=15$

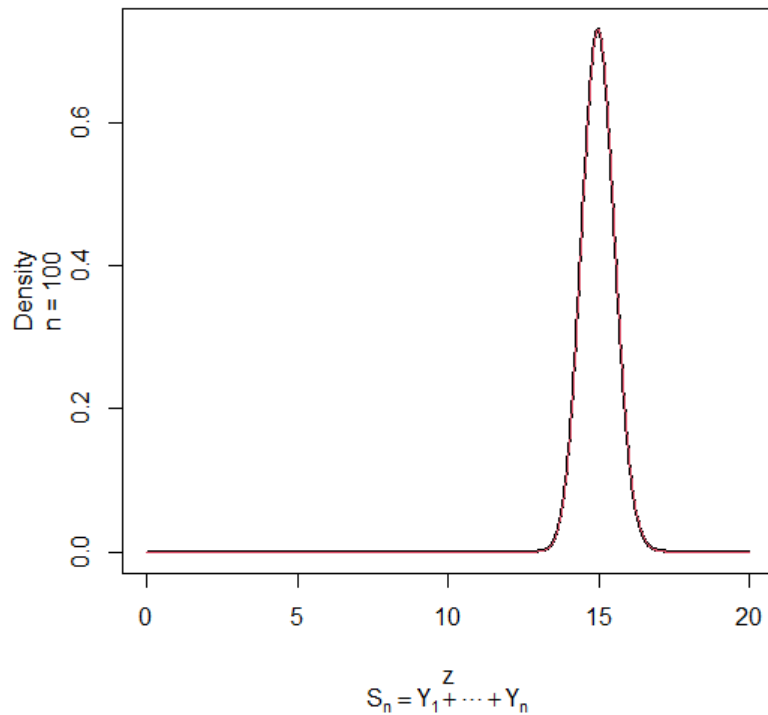


Figura 1.4.6 Përafrimi Edgeworth për mesataren e n ndryshoreve hi-katror, për $n=100$, $df=15$

Vija me ngjyrë të zezë tregon përafrimin Edgeworth, kurse vija me ngjyrë të kuqe tregon mesataren e n ndryshoreve hi-katror.

Nga figurat më sipër shohim se me rritjen e shkallës së lirisë orientimi i grafikut anon djathtas dhe me rritjen e vëllimit n , forma e grafikut ngushtohet. Për më tepër, tek të

gjitha figurat shohim se përafrimi Edgeworth nuk varet nga ndryshimi i vëllimit apo i shkallës së lirisë sepse mbetet një përafrim mjaft i mirë ndaj mesatares së ndryshoreve hi-katror.

KAPITULLI II

METODA BOOTSTRAP DHE ZBËRTHIMI EDGEWORTH PËR VERSIONIN BOOTSTRAP

2.1 Metoda bootstrap

Metodat e rizgjedhjes janë ndër metodat kryesore të statistikës që kanë marrë një zhvillim mjaft të madh kohët e fundit veçanërisht për shkak të suksesit të tyre, thjeshtësisë së përdorimit dhe kushteve paraprake që kërkohen për përdorimin e tyre. Këto metoda konsistojnë në rizgjedhjet e njëpasnjëshme mbështetur tek zgjedhja ose nënzgjedhja e marrë. Zhvillimi i teknologjive kompjuterike në dekadat e fundit bëri që të realizohej një zhvillim i menjëhershëm dhe i rëndësishëm i përdorimit të metodave të rizgjedhjes. Një ndër metodat e rizgjedhjes, që ka pasur një përdorim mjaft të madh, sidomos në dekadat e fundit, është dhe metoda Bootstrap.

Bootstrap është një teknikë e fuqishme për vlerësimin e saktësisë së një vlerësuesi për parametrat e panjohur në situata kur teknikat e zakonshme nuk janë efektive.

Zbatime të metodës bootstrap përfshijnë vlerësimin apo reduktimin e zhvendosjes së një vlerësuesi, ndërtimin e intervaleve të besimit pa supozime normaliteti etj.

Metoda bootstrap është prezantuar nga Efron ([34], [35]) si një mënyrë për të llogaritur intervale besimi për parametrat në situatat kur metodat standarde nuk mund të aplikohen. Për këtë qëllim bootstrap është përdorur në mjaft situata për të zgjidhur probleme të komplikuar në lidhje me analizën statistikore. Me anë të kësaj metode ne mund të rizgjedhim rastësisht vrojtimet si dhe të llogarisim vlerësimet.

Sipas Efron, nëse F është shpërndarja e popullimit, që gjeneron një zgjedhje të rastit të madhësisë T , atëherë si vlerësuesi $\hat{\beta}$ ashtu dhe shpërndarja e zgjedhjes së tij, G , mund të konsiderohen funksione të F . Efron sugjeroi zëvendësimin e F me një vlerësues të qëndrueshëm \hat{F} .

Mjaft aplikime dhe punime teorike përdorin funksionin e shpërndarjes empirike me probabilitet $\frac{1}{T}$ në çdo vrojtim të zgjedhjes së rastit. Nëse β është një moment i

popullatës me shpërndarje F , atëherë sipas kësaj analogjie, $\hat{\beta}$ jepet nga momenti korrespondues i zgjedhjes. Meqë shumë rrallë është e mundur të vlerësohet direkt shpërndarja e vlerësuesit të zgjedhjes si funksion i funksionit të shpërndarjes empirike, Efron ([34], [35]) propozoi më tej vlerësimin e shpërndarjes bootstrap,

$G(\hat{F})$, nga simulimet Monte Carlo. Kur e zbatojmë për vlerësimin e mesatares së popullatës nga një zgjedhje e rastësishme, metoda bootstrap sipas Efronit, konsiston në rizgjedhjet e përsëritura me zëvendësim nga vrojtimet, duke llogaritur mesataren nga secila zgjedhje "bootstrap" dhe pastaj duke përdorur shpërndarjen e këtyre mesatareve të zgjedhjeve bootstrap për të përafëruar shpërndarjen e mesatares së zgjedhjes origjinale.

Të shohim disa teorema të rëndësishme, mbi veti asimptotike të bootstrap, [37].

Është dhënë X_1, \dots, X_n zgjedhje e rastit nga ndryshorja e rastit X me funksion shpërndarjeje F . Supozojmë se X ka pritje matematike μ dhe variancë σ^2 . Vlerësimi i zakonshëm për pritjen matematike është mesatarja aritmetike, që e shënojmë $\bar{X}_{(n)}$,

ndërsa për të vlerësuar variancën ndërtojmë $s_{(n)}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_{(n)})^2$. Nga Teorema

Qendrore Limite, shpërndarja e $Q_{(n)} = \frac{\sqrt{n}(\bar{X}_{(n)} - \mu)}{s_{(n)}}$ konvergjon dobët tek shpërndarja $N(0,1)$. Të shohim sesi do të sillet vlerësimi bootstrap në këtë rast.

Le të jetë $\hat{F}_{(n)}$ funksioni i shpërndarjes i ndryshores së rastit, që merr vlerat $X_i = x_i$ me probabilitete të barabarta me $\frac{1}{n}$. Nga kjo shpërndarje, marrim një zgjedhje të rastit

me vëllim m . Ndërtojmë ndryshoret e rastit $Q_{(m)}^* = \frac{\sqrt{m}(\bar{X}_{(m)}^* - \bar{X}_{(n)})}{s_{(m)}^*}$ ku

$\bar{X}_{(m)}^* = \frac{1}{m} \sum_{i=1}^m X_i^*$ dhe $s_{(m)}^{2*} = \frac{1}{m} \sum_{i=1}^m (X_i^* - \bar{X}_{(m)}^*)^2$. Ka vend teorema:

Teoremë 2.1 Supozojmë se X_1, \dots, X_n është zgjedhje e rastit nga ndryshorja e rastit X me funksion shpërndarje F . Supozojmë se X ka pritje matematike μ dhe variancë σ^2 . Atëherë, kur m dhe n tentojnë në ∞ kanë vend këto pohime:

a) Shpërndarja me kusht e $Q_{(m)}^* = \frac{\sqrt{m}(\bar{X}_{(m)}^* - \bar{X}_{(n)})}{s_{(m)}^*}$ konvergjon dobët tek

$N(0, \sigma^2)$.

b) $s_{(m)}^* \rightarrow \sigma$ sipas probabilitetit me kusht; d.m.th. për çdo

$\varepsilon > 0, p(|s_{(m)}^* - \sigma| > \varepsilon | X_1, \dots, X_n) \rightarrow 0$ p.k. [7]

Metodat bootstrap mund të përdoren edhe për ndërtimin e vlerësimeve intervalore për medianën. Kështu ka vend teorema:

Teorema 2.2 Le të jetë m^* mediana e zgjedhjes bootstrap X^* dhe m mediana e zgjedhjes së rastit X . Nëqoftëse F ka medianë të vetme μ dhe f ka derivat pozitiv të vazhdueshëm në një fqinjësi të μ , atëherë për pothuajse të gjitha zgjedhjet X_1, X_2, \dots ,

të marra nga (X_1, \dots, X_n) , $\sqrt{n}(m^* - m)$ konvergjon dobët tek $N\left(0, \frac{1}{4f^2(\theta)}\right)$, e cila

është shpërndarja limite e $\sqrt{n}(m - \mu)$. [7]

Pavarësisht nga rezultati i arritur tek Teorema 2.1, po në të njëjtën kohë Singh [78], u arrit një rezultat i ngjashëm me të. Po e japim atë me anë të teoremave të mëposhtme:

Teorema 2.3 Le të jetë $\mu = E(X)$, $\sigma^2 = \text{var}(X)$, $\rho = E(|X - \mu|^3)$,

$s_{(n)} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_{(n)})^2$, p^* probabilitetet që lidhen me shpërndarjen \hat{F} dhe $\|\cdot\|_\infty$

norma e përcaktuar si për $\sup_{x \in R} |\cdot|$. Janë të vërtetë pohimet e mëposhtme:

a) Nëqoftëse $E(X^2) < \infty$, atëherë

$\left\| p\left(\sqrt{n}(\bar{X}_{(n)} - \mu) \leq x\right) - p^*\left(\sqrt{n}(\bar{X}_{(n)}^* - \bar{X}_{(n)}) \leq x\right) \right\|_\infty \rightarrow 0$ p.k.

b) Nëqoftëse $E(X^4) < \infty$, atëherë

$$\limsup_{n \rightarrow \infty} \sqrt{n} (\log \log n)^{-\frac{1}{2}} \left\| p \left(\sqrt{n}(\bar{X}_{(n)} - \mu) \leq x \right) - p^* \left(\sqrt{n}(\bar{X}_{(n)}^* - \bar{X}_{(n)}) \leq x \right) \right\|_{\infty} = \\ = \left(2\sigma^2 \sqrt{2\pi e} \right)^{-1} \sqrt{2 \operatorname{var}((X - \mu)^2)} \quad p.k.$$

c) Nëqoftëse $E(|X|^3) < \infty$, atëherë

$$\limsup_{n \rightarrow \infty} \rho \sigma^{-3} \sqrt{n} \left\| p \left(\frac{\sqrt{n}(\bar{X}_{(n)} - \mu)}{\sigma} \leq x \right) - \right. \\ \left. - p^* \left(\frac{\sqrt{n}(\bar{X}_{(n)}^* - \bar{X}_{(n)})}{s_{(n)}} \leq x \right) \right\|_{\infty} \leq 2K, \quad p.k.$$

ku K është konstantja në kufirin e Berry-Essen

d) Nëqoftëse $E(|X|^3) < \infty$ dhe F nuk ka laticë, atëherë

$$p^* \left(\frac{\sqrt{n}(\bar{X}_{(n)}^* - \bar{X}_{(n)})}{s_{(n)}} \leq x \right) = \Phi(x) + \frac{\mu_3(1-x^2)}{6\sigma^3 \sqrt{n}} \varphi(x) + o\left(\frac{1}{\sqrt{n}}\right)$$

uniformisht në lidhje me x p.k. ku $\Phi(x), \varphi(x)$ janë funksioni dhe densiteti i shpërndarjes normale, për më tepër në këtë rast

$$\sqrt{n} \left\| p \left(\frac{\sqrt{n}(\bar{X}_{(n)} - \mu)}{\sigma} \leq x \right) - p^* \left(\frac{\sqrt{n}(\bar{X}_{(n)}^* - \bar{X}_{(n)})}{s_{(n)}} \leq x \right) \right\|_{\infty} \rightarrow 0 \quad p.k.$$

e) Nëqoftëse $E(|X|^3) < \infty$ dhe F ka laticë me periudhë h

$$p^* \left(\frac{\sqrt{n}(\bar{X}_{(n)}^* - \bar{X}_{(n)})}{s_{(n)}} \leq x \right) = \Phi(x) + \frac{\mu_3(1-x^2)}{6\sigma^3 \sqrt{n}} \varphi(x) + \frac{h}{\sigma \sqrt{n}} g \left(\frac{\sqrt{n}s_{(n)}}{h} x \right) \varphi(x) + o\left(\frac{1}{\sqrt{n}}\right)$$

uniformisht në lidhje me x p.k. ku $g(y) = [y] - y + \frac{1}{2}$ për $\forall y \in \mathbb{R}$. Gjithashtu në këtë rast, [78]

$$\limsup_{n \rightarrow \infty} \sqrt{n} \left\| p^* \left(\frac{\sqrt{n}(\bar{X}_{(n)}^* - \bar{X}_{(n)})}{s_{(n)}} \leq x \right) - p \left(\frac{\sqrt{n}(\bar{X}_{(n)} - \mu)}{\sigma} \leq x \right) \right\|_{\infty} = \frac{h}{\sqrt{2\pi\sigma^2}}, \quad p.k.$$

Teoremë 2.4 Nëqoftëse F ka derivat të dytë të kufizuar në një fqinjësi të $F^{-1}(t)$ dhe $F'(F^{-1}(x)) > 0$, atëherë p.k.

$$\limsup_{n \rightarrow \infty} n^{\frac{1}{4}} (\log \log n)^{\frac{1}{2}} \left\| p \left(\sqrt{n}(\hat{F}_{(n)}^{-1}(t) - F^{-1}(t)) \leq x \right) - \right. \\ \left. - p^* \left(\sqrt{n}(G_{(n)}^{-1}(t) - \hat{F}_{(n)}^{-1}(t)) \leq x \right) \right\|_{\infty} = K_{t,F}$$

ku $K_{t,F}$ konstate që varet vetëm nga t dhe F , $G_{(n)}(x) = \frac{nr(X_i^* \leq x; 1 \leq i \leq n)}{n}$, [78].

2.2 Zbërthimi Edgeworth për versionin bootstrap

Kujtojmë që një statistikë është asimptotikisht pivote nëse limiti i shpërndarjes së saj nuk varet nga parametri i panjohur. Në disa raste bootstrap jep një vlerësim më të mirë të shpërndarjes kur kemi një statistikë pivote sesa kur kemi një statistikë jopivote. Avantazhi i pivotimit mund të shpjegohet mjaft lehtë nëpërmjet zbërthimit Edgeworth, siç vijon. N.q.s një statistikë pivote T ka asimptotikisht shpërndarje normale, atëherë funksioni i shpërndarjes së tij mund të zbërthehet si më poshtë

$$G(x) = P(T \leq x) = \Phi(x) + n^{-1/2}q(x)\phi(x) + O(n^{-1}), \quad (2.2.1)$$

ku q është një polinom kuadratik çift dhe Φ, ϕ janë respektivisht funksioni dhe densiteti i shpërndarjes Normale Standarde. Si një shembull mund të marrim $T = n^{1/2}(\hat{\theta} - \theta_0) / \hat{\sigma}$, ku $\hat{\theta}$ është një vlerësim i parametrin të panjohur θ_0 , dhe $\hat{\sigma}^2$ është një vlerësim i variancës asimptotike σ^2 të $n^{1/2}\hat{\theta}$. Vlerësimi bootstrap i G lejon një zbërthim analog,

$$G(x) = P(T^* \leq x | X) = \Phi(x) + n^{-1/2}\hat{q}(x)\phi(x) + O_p(n^{-1}), \quad (2.2.2)$$

ku T^* është versioni bootstrap i T , i njehsuar nga një rizgjedhje X^* në vend të zgjedhjes X , dhe polinomi \hat{q} është përftuar nga q duke zëvendësuar të panjohurat, të tilla si pjerrësia, me vlerësimet e tyre bootstrap. (Shënimi " $O_p(n^{-1})$ " tregon një ndryshore rasti të rendit n^{-1} "sipas probabilitetit"). Shpërndarja e T^* me kusht X quhet shpërndarja bootstrap e T^* . Vlerësimet e koeficientëve të \hat{q} ndryshojnë nga vlerat e tyre respektive të q me madhësinë $O_p(n^{-1/2})$, kështu $\hat{q} - q = O_p(n^{-1/2})$. Prandaj, duke zbritur (2.2.2) nga (2.2.1) nxjerrim se

$$P(T^* \leq x | X) - P(T < x) = O_p(n^{-1}).$$

Pra, përafrimi bootstrap i G është një gabim i rendit n^{-1} . Ky është një përafrim i konsiderueshëm në përafrimin Normal, $G \approx \Phi$, i cili nga (2.2.1) është një gabim i rendit $n^{-1/2}$.

Nga ana tjetër, kur përdorim bootstrap për të përafuar shpërndarjen e një statistike jopivote U , të tillë që $U = n^{1/2}(\hat{\theta} - \theta_0)$, do të bëjmë një gabim të rendit $n^{-1/2}$, në vend të n^{-1} . Për ta vlerësuar pse, vëzhgojmë se analoget e (2.2.1) dhe (2.2.2) në këtë rast janë, respektivisht

$$\begin{aligned} H(x) &= P(U \leq x) = \Phi(x/\sigma) + n^{-1/2}p(x/\sigma)\phi(x/\sigma) + O(n^{-1}), \\ H(x) &= P(U^* \leq x | X) = \Phi(x/\hat{\sigma}) + n^{-1/2}\hat{p}(x/\hat{\sigma})\phi(x/\hat{\sigma}) + O_p(n^{-1}) \end{aligned} \quad (2.2.3)$$

ku p është një polinom, \hat{p} është përftuar nga p duke zëvendësuar të panjohurat me vlerësimet e tyre bootstrap, σ^2 është e njëjtë me variancën asimptotike të U , $\hat{\sigma}^2$ është vlerësimi bootstrap i σ^2 , dhe U^* është versioni bootstrap i U . Sërish, $\hat{p} - p = O_p(n^{-1/2})$ dhe gjithashtu $\hat{\sigma} - \sigma = O_p(n^{-1/2})$, prej nga marrim

$$H(x) - H(x) = \Phi(x/\hat{\sigma}) - \Phi(x/\sigma) + O_p(n^{-1}). \quad (2.2.4)$$

Diferenca ndërmjet $\hat{\sigma}$ dhe σ zakonisht saktëson rendin $n^{-1/2}$.

Vërtet, limiti i $n^{1/2}(\hat{\sigma} - \sigma)$ ka shpërndarje Normale $N(0, \zeta^2)$, për $\zeta > 0$. Kështu, $\Phi(x/\hat{\sigma}) - \Phi(x/\sigma)$ zakonisht është e rendit $n^{-1/2}$, jo n^{-1} . Pra nga (2.2.4), përafrimi bootstrap i H është një gabim i rendit $n^{-1/2}$, jo n^{-1} . Kjo performancë relativisht e dobët është për shkak të prezencës së σ në limitin e funksionit të shpërndarjes $\Phi(x/\sigma)$, dmth për faktin që U është statistikë jopivote.

Le të jenë

$$P\left\{n^{1/2}(\hat{\theta} - \theta_0)/\sigma \leq x\right\} = \Phi(x) + n^{-1/2}p_1(x)\phi(x) + \dots + n^{-j/2}p_j(x)\phi(x) + \dots \quad (2.2.5)$$

dhe

$$P\left\{n^{1/2}(\hat{\theta} - \theta_0)/\hat{\sigma} \leq x\right\} = \Phi(x) + n^{-1/2}q_1(x)\phi(x) + \dots + n^{-j/2}q_j(x)\phi(x) + \dots \quad (2.2.6)$$

zberthimet Edgeworth për statistikat e Studentizuara dhe jo të Studentizuara, respektivisht. Këtu, p_j dhe q_j janë polinome kryesisht të shkallës $3j-1$ dhe janë funksione çift apo tek sipas j nëse është tek apo çift. Koeficientët e polinomeve p_j dhe q_j varen nga kumulantët e $(\hat{\theta} - \theta_0)/\sigma$ dhe $(\hat{\theta} - \theta_0)/\hat{\sigma}$, respektivisht. Sipas “model funksionit të lemuar”, të përshkruar në kapitullin 2 të [50], ku $\hat{\theta}$ dhe $\hat{\sigma}$ janë funksione të lëmuara të pritjes matematike të zgjedhjes d -përmasore $X = (X_1, \dots, X_n)$, këto koeficientë janë polinome të momenteve të vektorit X . Kur kalojmë nga zberthimet “bazuar në popullatë” (2.2.5) dhe (2.2.6) tek versionet e tyre “bazuar në zgjedhje”, roli i funksionit të shpërndarjes së popullatës F_0 është zëvendësuar nga funksioni i shpërndarjes së zgjedhjes F_1 , dhe kështu momentet e popullatës në koeficientët e polinomeve p_j dhe q_j ndryshohen me momentet e zgjedhjes. Le të shënojmë \hat{p}_j dhe \hat{q}_j polinomet e përftuar nga polinomet p_j dhe q_j duke zëvendësuar momentet e popullatës me momentet e zgjedhjes, dhe shprehim $\hat{\theta}^* = \theta(F_2)$ dhe $\hat{\sigma}^* = \sigma(F_2)$. Atëherë, versionet e zgjedhjes të zberthimeve (2.2.5) dhe (2.2.6), respektivisht, do të jenë

$$P\left\{n^{1/2}(\hat{\theta}^* - \hat{\theta})/\hat{\sigma} \leq x \mid X\right\} = \Phi(x) + n^{-1/2}\hat{p}_1(x)\phi(x) + \dots + n^{-j/2}\hat{p}_j(x)\phi(x) + \dots \quad (2.2.7)$$

dhe

$$P\left\{n^{1/2}(\hat{\theta}^* - \hat{\theta})/\hat{\sigma}^* \leq x \mid X\right\} = \Phi(x) + n^{-1/2}\hat{q}_1(x)\phi(x) + \dots + n^{-j/2}\hat{q}_j(x)\phi(x) + \dots \quad (2.2.8)$$

Konkretisht, këto zberthime janë përftuar nga (2.2.5) dhe (2.2.6) duke zëvendësuar çiftin (F_0, F_1) me (F_1, F_2) . Vërtetimet rigoroze të (2.2.7) dhe (2.2.8) mund t’i shikoni tek Seksioni 5.2 i [50].

Të dyja këto zberthime mund të interpretohen si seri asimptotike. P.sh. duke patur parasysh supozimin e lëmueshmërisë dhe momentit të pështatshëm,

$$\begin{aligned}
P\{n^{1/2}(\hat{\theta}^* - \hat{\theta}) / \hat{\sigma} \leq x | X\} \\
= \Phi(x) + n^{-1/2} \hat{p}_1(x) \phi(x) + \dots \\
+ n^{-j/2} \hat{p}_j(x) \phi(x) + o_p(n^{-j/2}).
\end{aligned} \tag{2.2.9}$$

Për më tepër, termi mbetës $o_p(n^{-j/2})$ është i këtij rendi, uniformisht në lidhje me x .

Në fakt, sipas kushteve të përshtatshme të rregullsisë, ai është i njëjtë me $o(n^{-j/2})$ pothuajse me siguri, uniformisht në lidhje me x :

$$\begin{aligned}
n^{-j/2} \sup_{-\infty < x < \infty} |P\{n^{1/2}(\hat{\theta}^* - \hat{\theta}) / \hat{\sigma} \leq x | X\} \\
- \{\Phi(x) + n^{-1/2} \hat{p}_1(x) \phi(x) + \dots + n^{-j/2} \hat{p}_j(x) \phi(x) + o_p(n^{-j/2})\}| \rightarrow 0
\end{aligned}$$

pothuajse me siguri kur $n \rightarrow \infty$.

Le të shqyrtojmë tani rastin e përafrimit të shpërndarjes së mesatares së zgjedhjes. Le të jetë $X = \{X_1, \dots, X_n\}$ një zgjedhje e ndryshoreve të rastit të pavarura dhe identikisht të shpërndara nga një popullatë me pritje matematike zero, variancë 1 dhe moment të fundëm të rendit të tretë, dhe supozojmë që X merr vetëm vlera të trajtës $a + jb$, për $j = 0, \pm 1, \pm 2, \dots$, ku $b > 0$ përcakton hapin maksimal të laticës. Përcaktojmë $R(x) = [x] - x + \frac{1}{2}$ dhe shënojmë \bar{X} mesataren e zgjedhjes. Atëherë,

$$\begin{aligned}
P(n^{1/2} \bar{X} \leq x) \\
= \Phi(x) + n^{-1/2} p_1(x) \phi(x) + \\
+ n^{-1/2} b R\{(n^{1/2} x - na) / b\} \phi(x) + o(n^{-1/2}).
\end{aligned}$$

Për të ndërtuar përafrimin bootstrap, le të shënojmë $X^* = \{X_1^*, \dots, X_n^*\}$ një rizgjedhje të rastësishme, me rikthim, nga X . Shënojmë \bar{X}^* mesataren e rizgjedhjes dhe $\hat{\sigma}^2$ variancën e zgjedhjes. Meqë $(\sum X_i^* - n\bar{X}) / \hat{\sigma}$ merr vlera vetëm të trajtës $jb / \hat{\sigma}$, për $j = 0, \pm 1, \pm 2, \dots$, atëherë versioni bootstrap i zbërthimit Edgeworth të mësipërm përfshin zëvendësimin e (a, b) me $(0, b / \hat{\sigma})$, si edhe ndryshimin e p_1 me \hat{p}_1 . Kështu,

$$\begin{aligned}
P\{n^{1/2}(\bar{X}^* - \bar{X}) / \hat{\sigma} \leq x | X\} \\
= \Phi(x) + n^{-1/2} \hat{p}_1(x) \phi(x) + \\
+ n^{-1/2} \hat{\sigma}^{-1} b R(n^{1/2} \hat{\sigma} x / b) \phi(x) + o_p(n^{-1/2}).
\end{aligned}$$

Duke marrë diferencën e dy zbërthimeve dhe duke vënë re se $\hat{p}_1 - p_1 = O_p(n^{-1/2})$ dhe $\hat{\sigma} - 1 = O_p(n^{-1/2})$, shohim që

$$\begin{aligned}
P(n^{1/2} \bar{X} \leq x) - P\{n^{1/2}(\bar{X}^* - \bar{X}) / \hat{\sigma} \leq x | X\} \\
= n^{-1/2} b \left[R\{(n^{1/2} x - na) / b\} - R(n^{1/2} \hat{\sigma} x / b) \right] \phi(x) + o_p(n^{-1/2}).
\end{aligned}$$

Ana e djathtë këtu nuk është e barabartë me $o_p(n^{-1/2})$. Vërtet, për ndonjë interval të hapur jo bosh $I \subseteq \mathbb{R}$, kemi

$$\limsup_{n \rightarrow \infty} \sup_{x \in I} |R\{(n^{1/2}x - na) / b\} - R(n^{1/2} \hat{\sigma}x / b)| \phi(x) \rightarrow \sup_{x \in I} \phi(x)$$

me probabilitet një. Prandaj, përafrimi bootstrap i shpërndarjes së \bar{X} është një gabim i rendit $n^{-1/2}$, dhe në këtë kuptim nuk përmirëson përafrimin Normal. Nga ana tjetër, në rastin kur shpërndarja e zgjedhjes është e lëmuar, nga (2.2.5) dhe (2.2.9) kemi që

$$P(n^{1/2} \bar{X} \leq x) - P\{n^{1/2}(\bar{X}^* - \bar{X}) / \hat{\sigma} \leq x | X\} = O_p(n^{-1}),$$

kështu që, përafrimi bootstrap është i një gabimi të rendit n^{-1} , në vend të rendit $n^{-1/2}$.

Singh [78] drejtoi problemin e përafrimeve bootstrap për shpërndarjen e pritjes matematike të një zgjedhjeje nga një popullatë laticë, dhe Hall [49] trajtoi korrektimet e vazhdueshmërisë për përmirësimin e këtyre përafrimeve.

2.3 Performanca e përafrimit Edgeworth për versionin me dhe pa bootstrap në tre shpërndarje të rëndësishme

Në këtë pjesë do të shohim sjelljen asimptotike të polinomeve të rendit të parë dhe të rendit të dytë të zbërthimit Edgeworth. Simulimet i kemi kryer për të dyja rastet kur përdorim dhe kur nuk përdorim versionin bootstrap, në tre shpërndarje të rëndësishme: normale, eksponenciale dhe lognormale, shih [25].

Simulimet i kemi kryer duke përdorur paketën “EW” në gjuhën e programimit R [75]. Kjo paketë na ndihmon për të gjeneruar një funksion polinomial p_j i cili përafron termin e gabimit të rendit $o(n^{-j/2})$ tek diferenca ndërmjet shpërndarjes normale $\Phi(x)$ dhe funksionit të shpërndarjes $F_n(x)$ të S_n . Aplikimi i kësaj pakete nxjerr një figurë në të cilën tregohet sjellja asimptotike e rendit të polinomit. Kemi ndërtuar dhe kodin përkatës bootstrap për ta implementuar tek kodi i mësipërm për të parë performancën edhe për rastin bootstrap.

Së pari kemi aplikuar sjelljen polinomiale në të dy rastet me dhe pa bootstrap, dhe më pas kemi prezantuar sjelljen asimptotike të polinomeve të rendit të parë dhe të rendit të dytë, gjithashtu në të dyja këto raste. Kemi kryer 100 replikime bootstrap dhe grafikët janë të gjithë shumë të ngjashëm dhe të mbivendosur njeri mbi tjetrin. Këtu kemi paraqitur vetëm njërin prej replikimeve për të shmangur kompleksitetin e figurave.

Për të tre shpërndarjet e marra në studim, nxjerrim përfundimin se performancat asimptotike të polinomeve të rendit të parë dhe të dytë për versionin bootstrap dhe versionin pa bootstrap janë shumë të ngjashme. Përfundime këto, që janë të ilustruara mjaft qartë nga figurat e mëposhtme:

Shpërndarja Normale

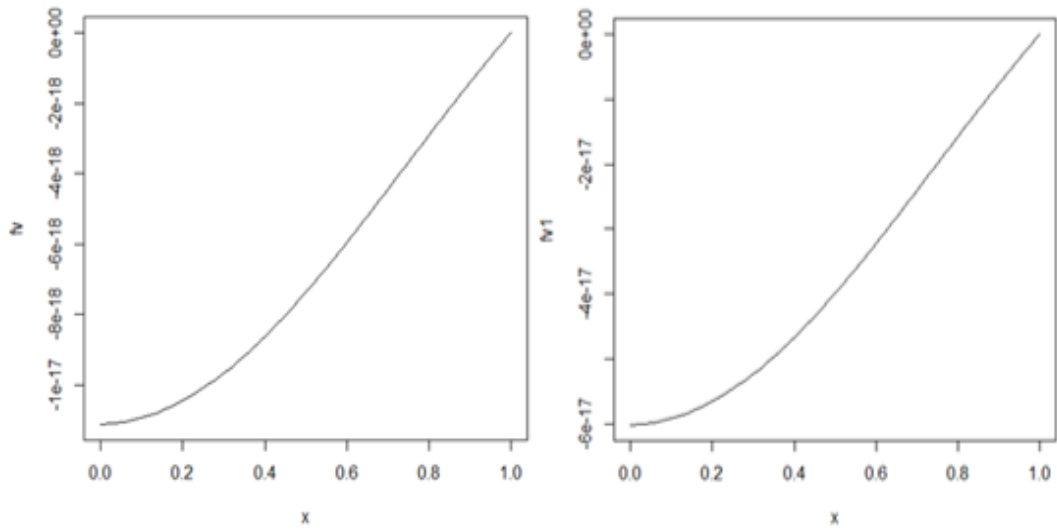


Figura 2.3.1 Sjellja polinomiale për rastin pa bootstrap dhe rastin bootstrap (Shpërndarja Normale)

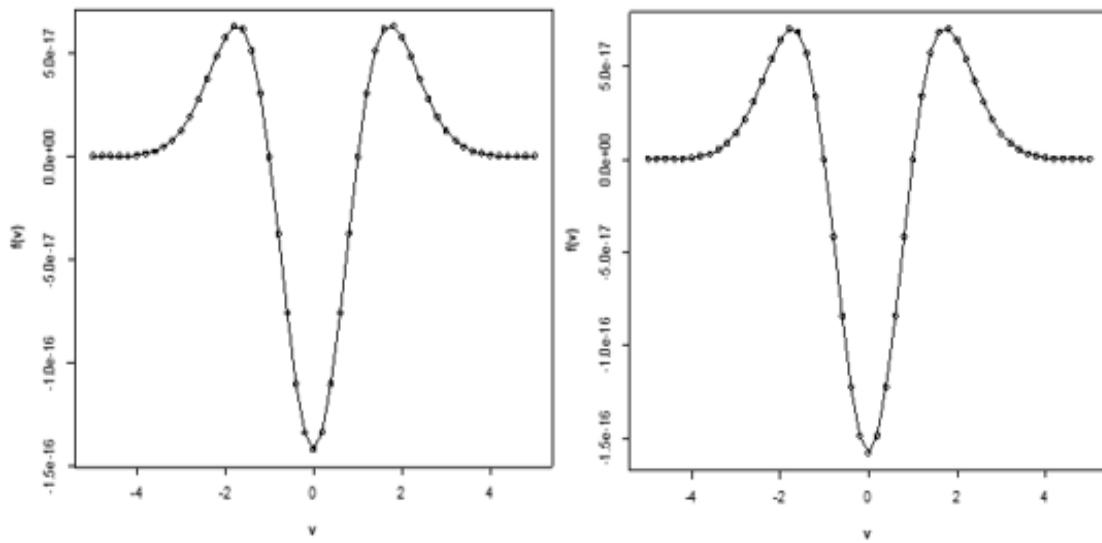


Figura 2.3.2 Sjellja asimptotike e polinomit të rendit të parë në rastin pa bootstrap dhe rastin bootstrap (Shpërndarja Normale)

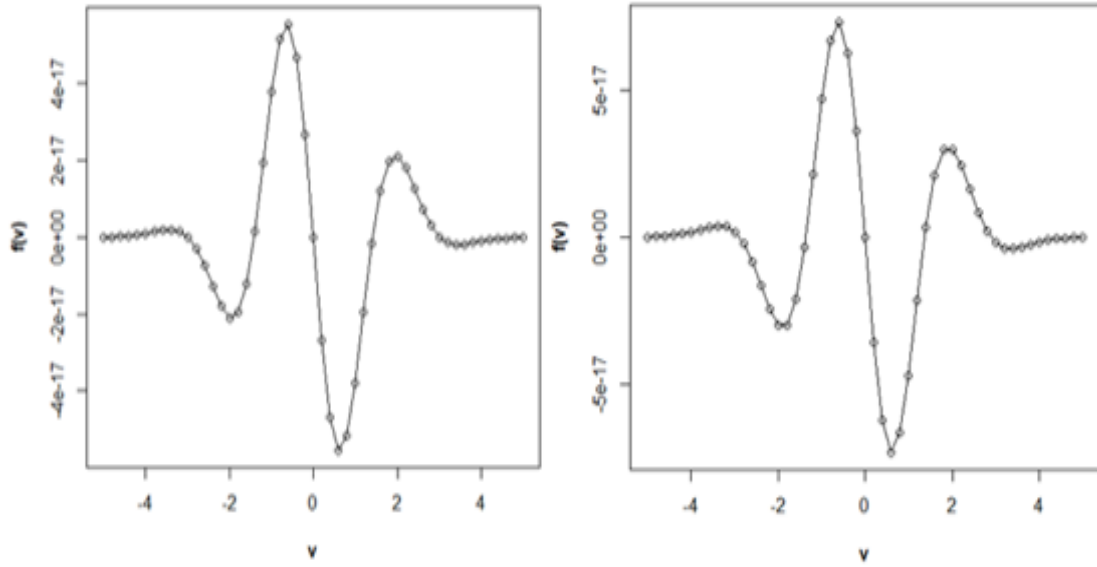


Figura 2.3.3 Sjellja asimptotike e polinomit të rendit të dytë në rastin pa bootstrap dhe rastin bootstrap (Shpërndarja Normale)

Shpërndarja Eksponenciale

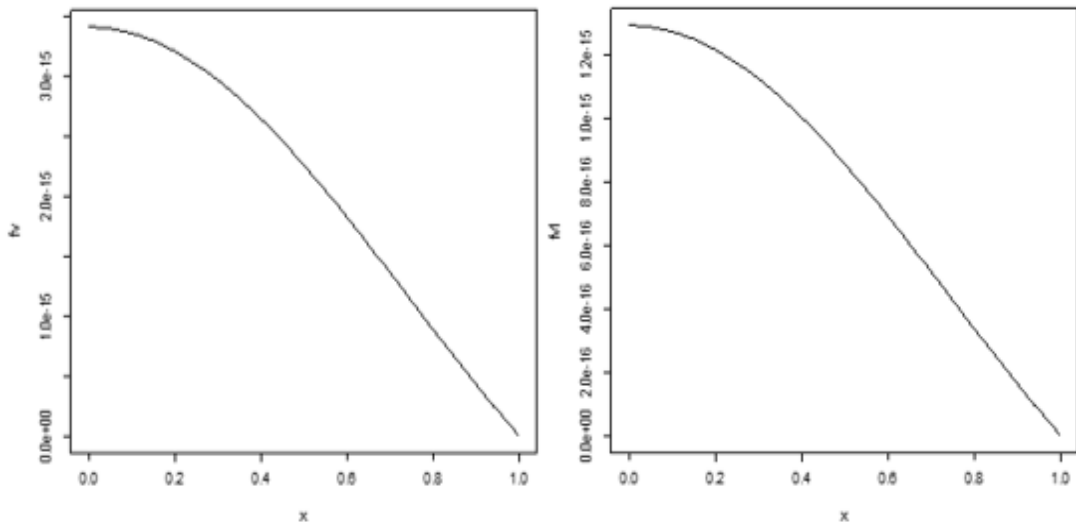


Figura 2.3.4 Sjellja polinomiale për rastin pa bootstrap dhe rastin bootstrap (Shpërndarja Eksponenciale)

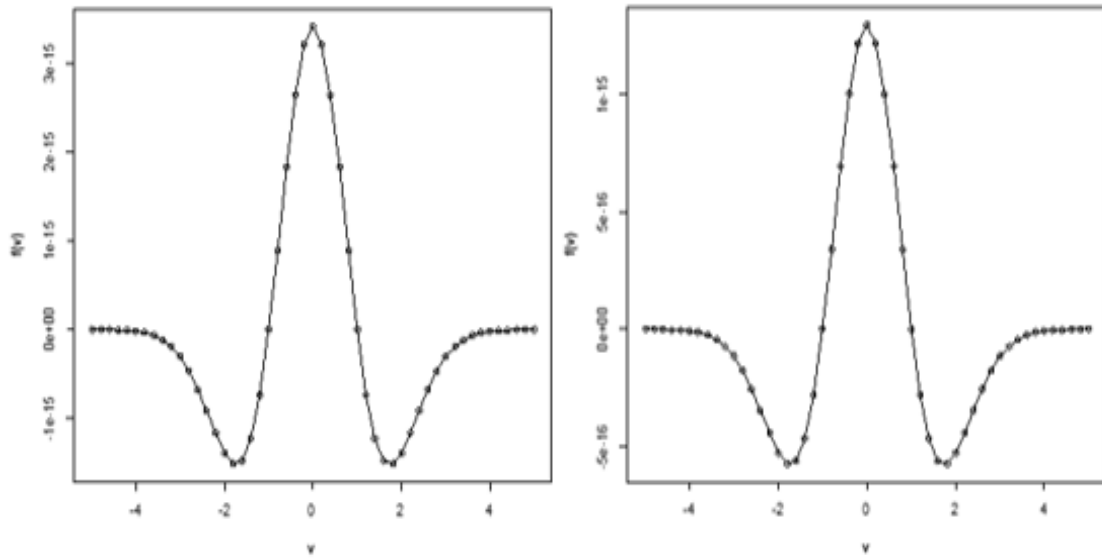


Figura 2.3.5 Sjellja asimptotike e polinomit të rendit të parë në rastin pa bootstrap dhe rastin bootstrap (Shpërndarja Eksponenciale)

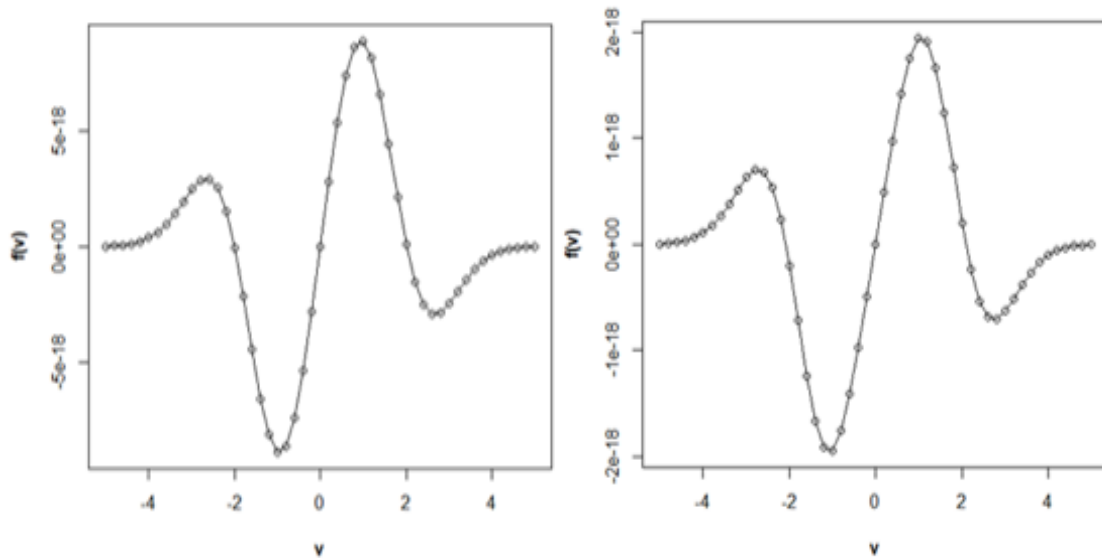


Figura 2.3.6 Sjellja asimptotike e polinomit të rendit të dytë në rastin pa bootstrap dhe rastin bootstrap (Shpërndarja Eksponenciale)

Shpërndarja Lognormale

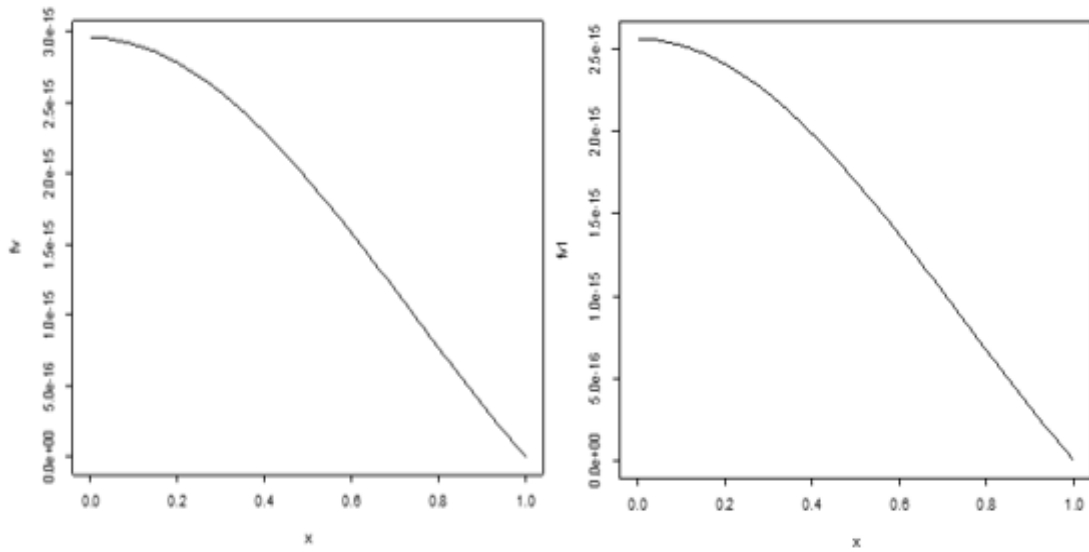


Figura 2.3.7 Sjellja polinomiale për rastin pa bootstrap dhe rastin bootstrap (Shpërndarja Lognormale)

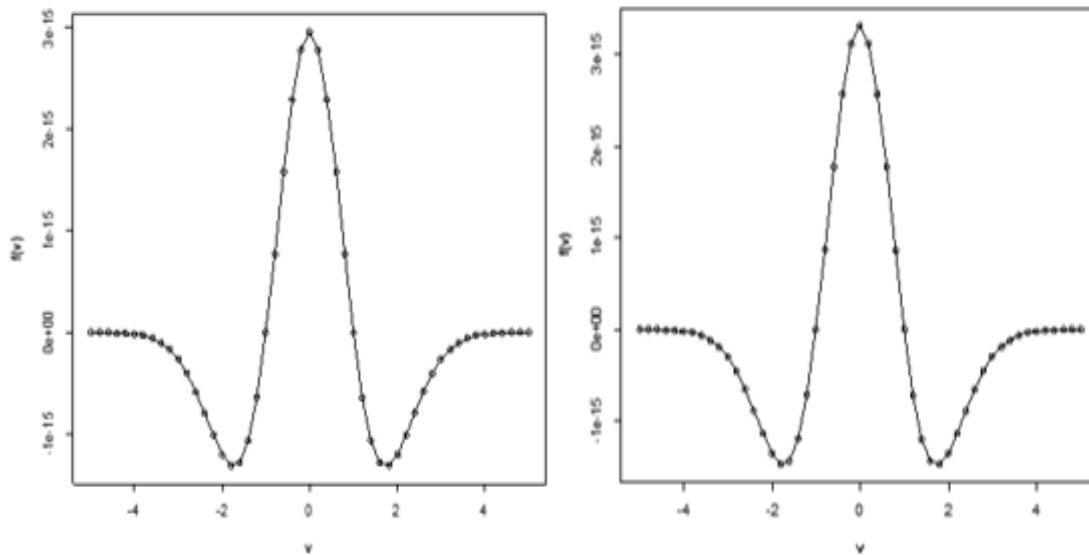


Figura 2.3.8 Sjellja asimptotike e polinomit të rendit të parë në rastin pa bootstrap dhe rastin bootstrap (Shpërndarja Lognormale)

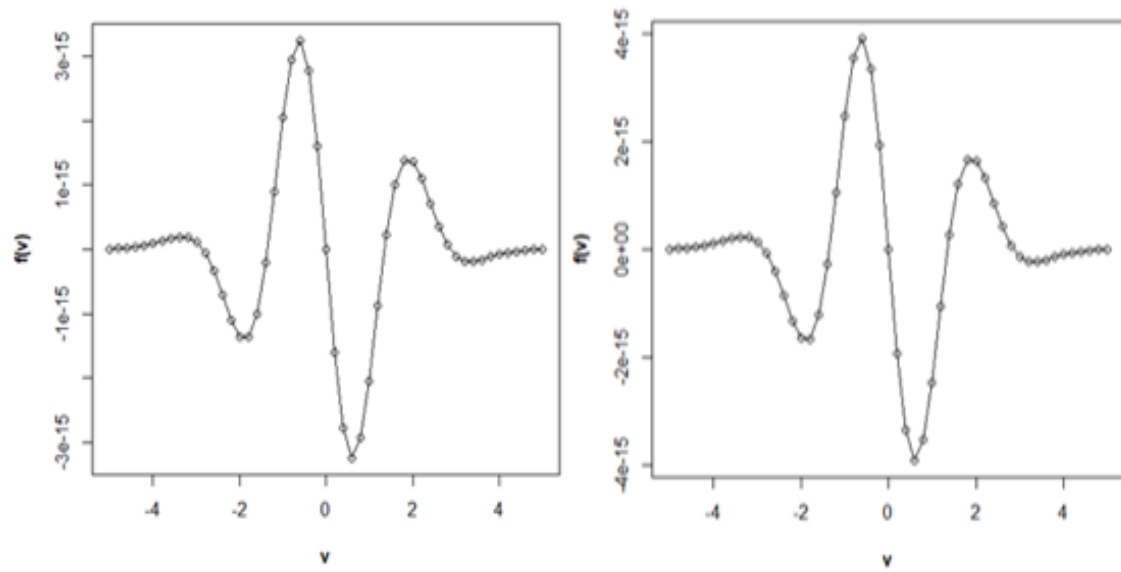


Figura 2.3.9 Sjellja asimptotike e polinomit të rendit të dytë në rastin pa bootstrap dhe rastin bootstrap (Shpërndarja Lognormale)

KAPITULLI III

INTERVALET E BESIMIT BOOTSTRAP

3.1 Intervalet e besimit bootstrap

Intervalet e besimit bootstrap u paraqitën për herë të parë nga Efroni [34]. Vend të rëndësishëm në teorinë asimptotike për intervalet e besimit bootstrap zenë zbërthimet Edgeworth. Hall [50] detajoi zbatimin e këtyre zbërthimeve tek intervalet e besimit bootstrap. Sipas tij ato janë rezultative në ndërtimin e rendit të saktësisë së intervaleve. Në shumë materiale të ndryshme janë trajtuar intervalet e besimit bootstrap, (shih [40], [14], [20], [21], [74]).

Intervalet e besimit bootstrap që do të shohim më poshtë janë: të zakonshëm, të studentizuar, kuantile dhe me korrektësi të zhvendosur (BCa).

Intervali kuantil bootstrap: Për të ndërtuar një interval besimi për një parametër θ , ne duhet të dimë sesi $\hat{\theta}$, si një vlerësues i θ , ndryshon në zgjedhjet e përsëritura nga popullata. Intervali kuantil bootstrap $(100-\alpha)\%$ përmban mesin e shpërndarjes bootstrap duke marrë kuantilet $\alpha/2$ dhe $(1-\alpha/2)$ të shpërndarjes bootstrap si skaje të intervalit: $[q_{\alpha/2}^*, q_{100-\alpha/2}^*]$. Me q_c^* është shënuar kuantili c i shpërndarjes bootstrap.

Intervali i zakonshëm bootstrap: Për të ndërtuar një interval besimi $(100-\alpha)\%$ për një parametër θ bazuar në vlerësimin e zgjedhjes $\hat{\theta}$, përcaktojmë distancën $\hat{\theta}-\theta$ për nivelin $\alpha\%$. Nëqoftëse ne i njohim vlerat kritike c_1 dhe c_2 të tilla që $P(c_1 \leq \hat{\theta}-\theta \leq c_2) = 1-\alpha/100$, atëherë mund t'i rishkruajmë inekuacionet siç vijon $P(\hat{\theta}-c_2 \leq \theta \leq \hat{\theta}-c_1) = 1-\alpha/100$ për të krijuar një interval besimi $(100-\alpha)\%$ për θ . Nëse ne nuk e njohim shpërndarjen e $\hat{\theta}$ (ose të $\hat{\theta}-\theta$) dhe kuantilet e tyre, atëherë ne marrim kuantilet korresponduese të shpërndarjes bootstrap. Nga supozimi që kuantilet e shpërndarjes së zgjedhjes dhe shpërndarjes bootstrap janë të përafërta, kemi që

$$\begin{aligned} (100-\alpha)\% &= P(q_{\alpha/2}^* - \hat{\theta} \leq \hat{\theta}^* - \hat{\theta} \leq q_{100-\alpha/2}^* - \hat{\theta}) \\ &\approx P(q_{\alpha/2}^* - \hat{\theta} \leq \hat{\theta} - \theta \leq q_{100-\alpha/2}^* - \hat{\theta}) \\ &= P(2\hat{\theta} - q_{100-\alpha/2}^* \leq \theta \leq 2\hat{\theta} - q_{\alpha/2}^*) \end{aligned}$$

duke përftuar intervalin $[2\hat{\theta} - q_{100-\alpha/2}^*, 2\hat{\theta} - q_{\alpha/2}^*]$.

Intervali i studentizuar bootstrap: Të shohim hapat për të ndërtuar këtë interval. Supozojmë se është dhënë një parametër θ dhe vlerësimi i tij $\hat{\theta}$. Le të jetë $\hat{\theta}^*$ vlerësimi bootstrap joparametrik i θ bazuar në të njëjtën zgjedhje bazë me vëllim n nga e cila u përftua $\hat{\theta}$. Gjithashtu, kemi vlerësimin e devijimit standard për $\hat{\theta}$, $\hat{\sigma}$ dhe vlerësimin bootstrap korrespondues të tij, $\hat{\sigma}^*$. Shënojmë

$$S^* = \frac{(\hat{\theta}^* - \hat{\theta})}{\hat{\sigma}^*}$$

ku S^* është versioni i studentizuar bootstrap i $S = \frac{(\hat{\theta} - \theta)}{\hat{\sigma}}$.

Nëse θ është pritja e popullatës dhe $\hat{\theta}$ pritja e zgjedhjes, atëherë S është statistika t e studentizuar dhe nëse zgjedhja është ka shpërndarje normale, S do të jetë madhësi pivote. Një madhësi pivote është një madhësi rasti, shpërndarja e të cilës është e pavarur nga parametrat e modelit. Në rastin e S së mësipërme, nga supozimi i normalitetit, ajo statistikë ka shpërndarje studentit t me $n-1$ shkallë lirie.

Në këtë rast, $P(-c < S < c)$ është vlerësuar duke përdorur shpërndarjen t me $n-1$ shkallë lirie. Pra, $P(-c < S < c) = P(-c < (\hat{\theta} - \theta) / \hat{\sigma} < c)$. Atëherë kanë vend hapat vijues:

$$P(-c < (\hat{\theta} - \theta) / \hat{\sigma} < c) = P(-c\hat{\sigma} < (\hat{\theta} - \theta) < c\hat{\sigma}) = P(\hat{\theta} - c\hat{\sigma} < \theta < \hat{\theta} + c\hat{\sigma}).$$

Në rastin e statistikës bootstrap S^* për pritjen, sipas të njëjtave supozime si më sipër, S^* është asimptotikisht pivote, d.m.th. shpërndarja e saj është e pavarur nga parametrat dhe kuantilet e saj konvergjojnë tek kuantilet e S . Në këtë rast intervali i besimit bootstrap do të jetë $[\hat{\theta} - t^* \hat{\sigma}^*, \hat{\theta} + t^* \hat{\sigma}^*]$, ku t^* është kuantili $100(1-\alpha)$ i shpërndarjes t të studentit me $n-1$ shkallë lirie. Për sa kohë që S^* dhe S janë asimptotikisht pivote, kjo mënyrë gjenerimi e intervaleve të besimit bootstrap funksionon dhe është më e saktë se metoda kuantile e Efronit.

Intervali bootstrap BCa: Metoda bootstrap BC (korrektësi e zhvendosur) është krijuar për të funksionuar kur ekziston një funksion monoton $\varphi = g(\theta)$ dhe vlerësuesi $\hat{\varphi} = g(\hat{\theta})$ ka pothuajse shpërndarje normale me pritje $\varphi - z_0 \tau$. Parametri z_0 është korrektimi i zhvendosjes dhe τ është devijimi standard i $\hat{\varphi}$. Një vlerësimi i z_0 përftohet duke përdorur bootstrap. Metoda Bca përfshin një parametër të cilin Efroni e quajti “konstante e përshpejtuar”. Ajo bazohet në momentin e rendit të tretë dhe korrekton pjerrësinë bazuar në zberthimin Edgeworth. BCa funksionon veçanërisht mirë kur $\hat{\varphi}$ është pothuajse Gaussian me pritje $\varphi - z_0 \tau_\varphi$, ku z_0 është korrektimi i zhvendosjes si tek metoda BC dhe τ_φ është një funksion linear i φ i trajtës $1+a\varphi$, ku a përfaqëson konstanten e përshpejtuar të Efronit.

Kjo metodë ka një saktësi të rendit të dytë, veti që BC nuk e ka. Kjo është e vërtetë edhe pse “ a ” mund të vlerësohet në procesin bootstrap. Në thelb, së pari përdoret përafrimi për të ndërtuar një interval besimi për φ dhe më pas duke përdorur transformimin monoton të anasjellë, gjenerohet një interval besimi për θ . Së pari, supozojmë se “ a ” dhe z_0 njihen. Le të jetë $U = \varphi + (1+a\varphi)(Z - z_0)$, ku Z ka shpërndarje normale standarde me kuantilin e vet α , të shënuar me z_α . Le të jetë $U = h(T)$, vlerësuesi i transformuar për $\varphi = h(\theta)$. Kështu, ndryshorja e rastit T merret nga përkufizimi i $h^{-1}(U)$. Funksioni h^{-1} ekziston sepse h është monoton. Dukë zëvendësuar ndryshoret e rastit U dhe Z me vlerat e tyre respektive u dhe z_α , marrim

$$\hat{\varphi}_\alpha = u + \sigma(u)(z_\alpha + z_0) / [1 - a(z_\alpha + z_0)].$$

h është transformimi monoton i panjohur që pasqyron θ tek φ . Atëherë, $\hat{\theta}_\alpha = h^{-1}(\hat{\varphi}_\alpha)$. Për vlerësimin bootstrap T^* shënojmë shpërndarjen e tij \hat{G} , atëherë

$$\begin{aligned}
\hat{G}(\hat{\theta}_\alpha) &= P^*[T^* < \hat{\theta}_\alpha | t] \\
&= P^*[U^* < \hat{\varphi}_\alpha | u] \\
&= \Phi([\hat{\varphi}_\alpha - u] / \sigma(u) + z_0) \\
&= \Phi([z_0 + z_\alpha] / [1 - a(z_0 + z_\alpha)] + z_0),
\end{aligned}$$

ku P^* është probabiliteti për zgjedhjen bootstrap, T^* dhe U^* janë vlerësimet bootstrap për T dhe U , respektivisht. Meqënëse supozuam se “ a ” dhe z_0 njihen dhe Φ është shpërndarja normale standarde kumulative, madhësia në anën e djathtë të ekuacionit njihet. Prandaj, një kufi besimi $100\alpha\%$ i poshtëm për θ do të jetë

$$\hat{\theta}_\alpha = \hat{G}^{-1}\{\Phi([z_0 + z_\alpha] / [1 - a(z_0 + z_\alpha)] + z_0)\}.$$

Mungesën e njohurive për h e rregulluam duke përdorur \hat{G} . Por kujtojmë që supozuam se “ a ” dhe z_0 njihen. Zëvendësojmë ato me vlerësimet bootstrap. Marrim $\hat{z}_0 = \Phi^{-1}\{\hat{G}(t)\}$, ku t është vlera për të cilën

$$P^*[T^* < t | t] = P^*[U^* < u | u] = P[U < \varphi | \varphi] = \Phi(z_0).$$

Kur përdoret përafrimi Monte Carlo tek bootstrapi

$$\hat{z}_0 = \Phi^{-1}(\# \text{of } \{t_i^* \leq t\} / \{B+1\})$$

ku B është numri i replikimeve bootstrap. Nga përkufizimi i U dhe $\sigma(\varphi)$, mund të përcaktojmë se

$$\hat{a} = E^* \left\{ \dot{\ell}_\theta^*(\hat{\theta})^3 \right\} / \left(6 \text{var}^* \dot{\ell}_\theta^*(\hat{\theta})^{3/2} \right),$$

ku $\dot{\ell}_\theta^*$ është derivati i parë i log të përgjasisë l për θ , ku $\hat{\theta}$ është vlerësimi i zgjedhjes së θ dhe të gjitha pritjet dhe përgjasitë janë marrë nga zgjedhja bootstrap.

3.2 Zbërthimi Edgeworth për korrektësinë e intervaleve të besimit

Supozojmë se ndërtojmë një interval besimi (CI) bazuar në përafrimin normal standard të

$$S_n = \sqrt{n}(\hat{\theta}_n - \theta) / \sigma$$

ku σ është varianca asimptotike e $\sqrt{n}\hat{\theta}_n$.

Do të themi se një interval besimi (CI) ka mbulim asimptotik α , n.q.s

$$P(\theta \in CI) \rightarrow \alpha, \text{ kur } n \rightarrow \infty.$$

α quhet nivel mbulimi i intervalit të besimit, shih [50].

Gabimi i mbulimit i intervalit të besimit është diferenca mes mbulimit të vërtetë dhe nivelit të mbulimit

$$\text{gabimi i mbulimit} = P(\theta \in CI) - \alpha.$$

Intervalet e besimit të njëanshëm me nivel mbulimi α kanë trajtën

$$I_1 = (-\infty, \hat{\theta} + n^{-1/2} \sigma z_\alpha)$$

ku z_α përcaktohet nga $\Phi(z_\alpha) = \alpha$.

$$\begin{aligned} P(\theta \in I_1) &= P(\theta < \hat{\theta} + n^{-1/2} \sigma z_\alpha) = P(S_n > -z_\alpha) \\ &= 1 - (\Phi(-z_\alpha) + n^{-1/2} p_1(-z_\alpha) \phi(-z_\alpha) + O(n^{-1})) \\ &= \alpha - n^{-1/2} p_1(z_\alpha) \phi(z_\alpha) + O(n^{-1}) = \alpha + O(n^{-1/2}). \end{aligned}$$

Intervalet e besimit të dyanshëm me nivel mbulimi α kanë trajtën

$$I_2 = (\hat{\theta} - n^{-1/2} \sigma x_\alpha, \hat{\theta} + n^{-1/2} \sigma x_\alpha)$$

ku $x_\alpha = z_{(1+\alpha)/2}$.

$$\begin{aligned} P(\theta \in I_2) &= P(S_n \leq x_\alpha) - P(S_n \leq -x_\alpha) \\ &= \Phi(x_\alpha) - \Phi(-x_\alpha) + n^{-1/2} [p_1(x_\alpha) \phi(x_\alpha) - p_1(-x_\alpha) \phi(-x_\alpha)] + \\ &\quad + n^{-1} [p_2(x_\alpha) \phi(x_\alpha) - p_2(-x_\alpha) \phi(-x_\alpha)] + n^{-3/2} [p_3(x_\alpha) \phi(x_\alpha) - p_3(-x_\alpha) \phi(-x_\alpha)] + O(n^{-2}) \\ &= \alpha + 2n^{-1} p_2(x_\alpha) \phi(x_\alpha) + O(n^{-2}) = \alpha + O(n^{-1}). \end{aligned}$$

Ta përmbledhim: Gabimi i mbulimit i intervaleve të besimit (CI) të njëanshëm është i rendit $O(n^{-1/2})$, për intervalet e besimit të dyanshëm është $O(n^{-1})$. Intervali i besimit që ka gabim mbulimi të rendit $n^{-1/2}$ quhet interval besimi me korrektësi të rendit të parë. Intervali i besimit që ka gabim mbulimi të rendit n^{-1} quhet interval besimi i rendit të dytë.

Ky fakt ka një shtrirje të rëndësishme për përsëritjet bootstrap, pasi kjo do të thotë që çdo përsëritje redukton rendin e gabimit me një faktor $n^{-1/2}$ në rastin e intervaleve të besimit (CI) me korrektësi të rendit të parë, ndërsa në rastin e intervaleve të besimit (CI) me korrektësi të rendit të dytë me një faktor n^{-1} .

Intervalet kuantile dhe të zakonshëm janë intervale besimi me korrektësi të rendit të parë, kurse intervalet e studentizuar dhe BCa janë intervale besimi me korrektësi të rendit të dytë.

3.3 Zbërthimi Edgeworth për intervalet e besimit bootstrap

Le të shqyrtojmë statistikën $T = n^{1/2}(\hat{\theta} - \theta_0) / \hat{\sigma}$ dhe $T^* = n^{1/2}(\hat{\theta}^* - \theta_0) / \hat{\sigma}^*$, dhe le të jenë w_α dhe \hat{w}_α zgjidhjet e ekuacioneve

$$P(|T| \leq w_\alpha) = P(|T^*| \leq \hat{w}_\alpha | X) = \alpha.$$

Intervali teorik $I = (\hat{\theta} - n^{-1/2} \hat{\sigma} w_\alpha, \hat{\theta} + n^{-1/2} \hat{\sigma} w_\alpha)$ ka saktësi mbulimi mjaft të mirë meqë

$$P(\theta_0 \in I) = P(|T| \leq w_\alpha) = \alpha.$$

Intervali bootstrap $\hat{I} = (\hat{\theta} - n^{-1/2} \hat{\sigma} \hat{w}_\alpha, \hat{\theta} + n^{-1/2} \hat{\sigma} \hat{w}_\alpha)$ ka mbulim α .

Të dyja w_α dhe \hat{w}_α konvergjojnë tek pika kritike e shpërndarjes Normale Standarde $z_{(1+\alpha)/2}$.

Meqë

$$\hat{w}_\alpha - w_\alpha = O_p(n^{-3/2}), \quad (3.3.1)$$

skajet e intervalit \hat{I} përputhen me ato të intervalit I sipas termave të rendit $n^{-3/2}$; pra \hat{I} ka korrektësi të rendit të tretë për I . Meqë $T \mp \hat{w}_\alpha$ dhe $T \mp w_\alpha$ ndryshojnë vetëm me terma të rendit $O_p(n^{-3/2})$, nga (3.3.1) dhe metoda delta (shih [50])

$$P(T \leq \pm \hat{w}_\alpha) = P(T \leq \pm w_\alpha) + O(n^{-3/2}).$$

Prandaj,

$$\begin{aligned} P(\theta_0 \in \hat{I}) &= P(|T| \leq \hat{w}_\alpha) \\ &= P(|T| \leq w_\alpha) + O(n^{-3/2}) = \alpha + O(n^{-3/2}). \end{aligned}$$

Gabimi i mbulimit të \hat{I} është në të vërtetë $O(n^{-2})$, jo thjesht $O(n^{-3/2})$. Fakti se intervali i besimit \hat{I} ka gabim mbulimi të rendit n^{-2} rrjedh nga vetitë e zakonshme të barazimit të polinomeve në zbërthimin Edgeworth.

Pra, mund të shkruajmë

$$\begin{aligned} P(\theta_0 \in \hat{I}) &= \\ &= \Phi(z) + n^{-3/2}r_3(z)\phi(z) + n^{-2}r_4(z)\phi(z) + \dots \\ &\quad - \left\{ \Phi(-z) + n^{-3/2}r_3(-z)\phi(-z) + n^{-2}r_4(-z)\phi(-z) + \dots \right\} \\ &= \alpha + 2n^{-2}r_4(z)\phi(z) + O(n^{-3}), \end{aligned}$$

ku r_3, r_4, \dots janë polinome, r_j janë funksione çift ose tek për j tek ose çift, respektivisht. Prandaj, $P(\theta_0 \in \hat{I}) = \alpha + O(n^{-2})$.

3.4 Krahasimi i intervaleve të besimit bootstrap

Këtë krahasim intervalësh do ta ilustruajmë me një rast konkret, shih [24]. Kemi marrë në shqyrtim çmimet e shitjes së 30 shtëpive në qytetin e Tiranës, të marra nga gazeta “Çelësi” në mënyrë rastësore. Ne kemi bërë përpunimin e nevojshëm të të dhënave, simulimet e nevojshme mbi këtë bashkësi të dhënash në gjuhën e programimit R dhe si paketë ndihmëse në ndërtimin e kodeve kemi përdorur paketën “boot”. Bashkësia jonë e të dhënave është: 63, 47, 60, 48, 63, 55, 60, 37, 37, 55, 72, 70, 52, 85, 45, 48, 55, 58, 65, 48, 60, 53, 35, 66, 63, 60, 60, 52, 80, 57 (secila në 000€). Duke përdorur Shapiro-test pamë se këto të dhëna kanë shpërndarje normale. Gjatë punës sonë kemi bërë simulime për $R=1000$ zgjedhje bootstrap. Mesatarja e të dhënave tona është afërsisht 56.96 (000 €). Kemi llogaritur intervalet e besimit (CI): të zakonshëm, kuantil, të studentizuar, BCa për nivelet 95% dhe 99%. Rezultatet janë më poshtë:

Tabela 3.4.1 Intervalet e besimit bootstrap për nivelin 95%

| Intervali i besimit (CI) | 95% CI | Gjerësia | Mesi i intervalit |
|--------------------------|----------------|----------|-------------------|
| I Zakonshëm | [52.67, 60.97] | 8.3 | 56.82 |
| I Studentizuar | [52.76, 61.29] | 8.53 | 57.025 |
| Kuantile | [52.97, 61.27] | 8.3 | 57.12 |
| BCa | [52.97, 61.30] | 8.33 | 57.135 |

Tabela 3.4.2 Intervalet e besimit bootstrap për nivelin 99%

| Intervali i besimit (CI) | 99% CI | Gjerësia | Mesi i intervalit |
|--------------------------|----------------|----------|-------------------|
| I Zakonshëm | [51.23, 62.30] | 11.07 | 56.765 |
| I Studentizuar | [51.02, 63.14] | 12.12 | 57.08 |
| Kuantile | [51.63, 62.70] | 11.07 | 57.165 |
| BCa | [51.71, 62.86] | 11.15 | 57.285 |

Nga tabelat e mësipërme shohim se intervali i besimit i Studentizuar është më i gjerë sesa të tjerët për të dy rastet e niveleve 95% dhe 99%. Gjerësinë më të vogël e kanë intervalet e zakonshëm dhe kuantil. Pikat e mesit të intervaleve të zakonshëm dhe të Studentizuar janë më afër mesatares së të dhënave tona dhe vlerat e tyre janë të ngjashme me njëra-tjetrën. Pikat e mesit të intervaleve kuantil dhe BCa janë të ngjashme për të dy rastet. Intervalet e besimit për nivelin 99% kanë gjerësi më të madhe sesa intervalet me nivel 95%. Pikat e mesit të intervaleve me nivel 95% janë më afër mesatares së të dhënave sesa ato me nivel 99%.

Gjithashtu, ne kemi llogaritur edhe probabilitetin e mbulimit për nivelet 95% dhe 99% të intervalit të besimit për pritjen, për të dhënat tona normale, probabilitetin e mbulimit për shpërndarjen e studentizuar t dhe probabilitetin e mbulimit për intervalin e besimit bootstrap për pritjen. Rezultatet janë siç vijon:

Tabela 3.4.3 Probabilitetet e mbulimit (CP) të intervaleve të besimit për nivelet 95% dhe 99%

| Probabiliteti i mbulimit (CP) | CI për pritjen | CI për shpërndarjen t | Boot. CI për pritjen |
|-------------------------------|----------------|-------------------------|----------------------|
| CP- 95% | 0.956 | 0.948 | 0.943 |
| CP- 99% | 0.993 | 0.991 | 0.99 |

Nga kjo tabelë shohim se probabiliteti i mbulimit për intervalet e besimit jo bootstrap është më afër me nivelin tonë 95%. Por probabiliteti i mbulimit i intervalit të besimit bootstrap është më afër nivelit 99% sesa intervalet e besimit jo bootstrap.

3.5 Simulime për shpërndarjen lognormale

Shpesh në statistikë shtrohet problemi i vlerësimit të intervaleve të besimit për shpërndarjet me pjerrësi pozitive. Metodatat e rizgjedhjes bootstrap kanë veti të korrektësisë dhe të saktësisë së rendit të dytë dhe shmangin ndikimet e pjerrësisë ([65], [27], [39]). Më poshtë do të shohim disa simulime të bëra në shpërndarjen lognormale.

Shpërndarja lognormale është një shpërndarje që përdoret zakonisht në fushat e ekonomisë dhe financës [64]. Dy karakteristika të saj janë: ndryshoret e rastit jo negative dhe forma e funksionit të densitetit ka pjerrësi pozitive [82]. Le të jetë ndryshorja e rastit y me shpërndarje lognormale, shënohet $\ln(y) \sim N(\mu, \sigma^2)$, atëherë kemi që $\mu = E(\ln(y))$, and $\sigma^2 = Var(\ln(y))$. Densiteti i ndryshores së rastit y lognormale është:

$$f(y) = \frac{1}{\sqrt{2\pi\sigma y}} \left[-\frac{(\ln y - a)^2}{2\sigma^2} \right], \text{ për } y > 0.$$

Ne kemi marrë në shqyrtim shpërndarjen lognormale me parametrat $\mu=0, \sigma^2=1$. Duke përdorur gjuhën e programimit R, kemi bërë simulimet për $R=10000$ zgjedhje bootstrap për rastet e vëllimeve 30, 100, 500. Kemi ndërtuar intervalet e besimit bootstrap kuantil për nivelet 95% dhe 99% për pritjen matematike në shpërndarjen tonë lognormale, si edhe grafikët e densiteteve empirike për rastet e vëllimeve të mësipërme, shih [27]. Më poshtë janë rezultatet e simulimeve.

Tabela 3.5.1 Intervalet e besimit bootstrap kuantil 95% dhe 99%, për pritjen matematike në shpërndarjen lognormale

| | 95% CI | 99% CI | Meset e intervaleve | Pritja matematike |
|-------|--------------|--------------|---------------------|-------------------|
| n=30 | [1.13, 1.94] | [1.03, 2.09] | 1.54; 1.56 | 1.52 |
| n=100 | [1.34, 2.02] | [1.26, 2.15] | 1.68; 1.71 | 1.68 |
| n=500 | [1.52, 1.99] | [1.46, 2.09] | 1.76; 1.78 | 1.73 |

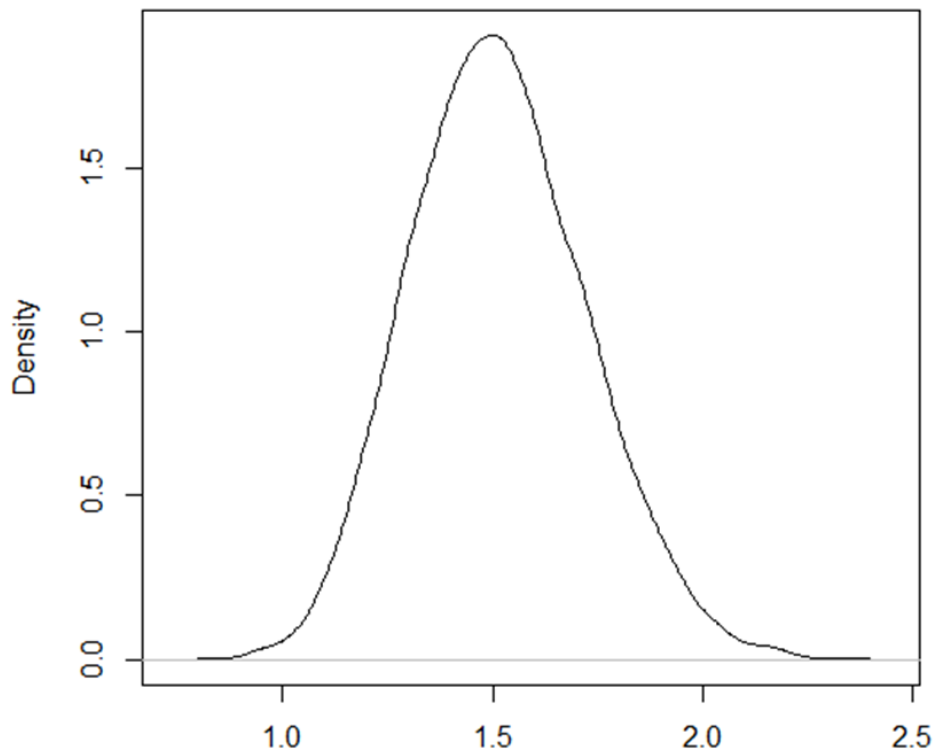


Figura 3.5.1 Vija e densitetit empirik e shpërndarjes lognormale për $R=1000$ zgjedhje bootstrap dhe $n=30$

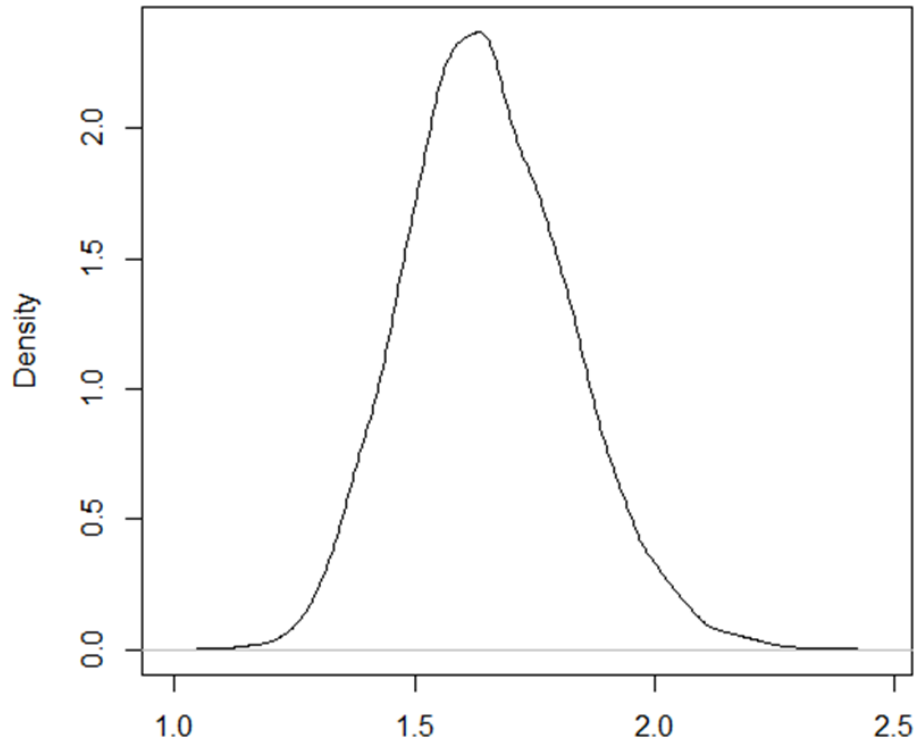


Figura 3.5.2 Vija e densitetit empirik e shpërndarjes lognormale për $R=1000$ zgjedhje bootstrap dhe $n=100$

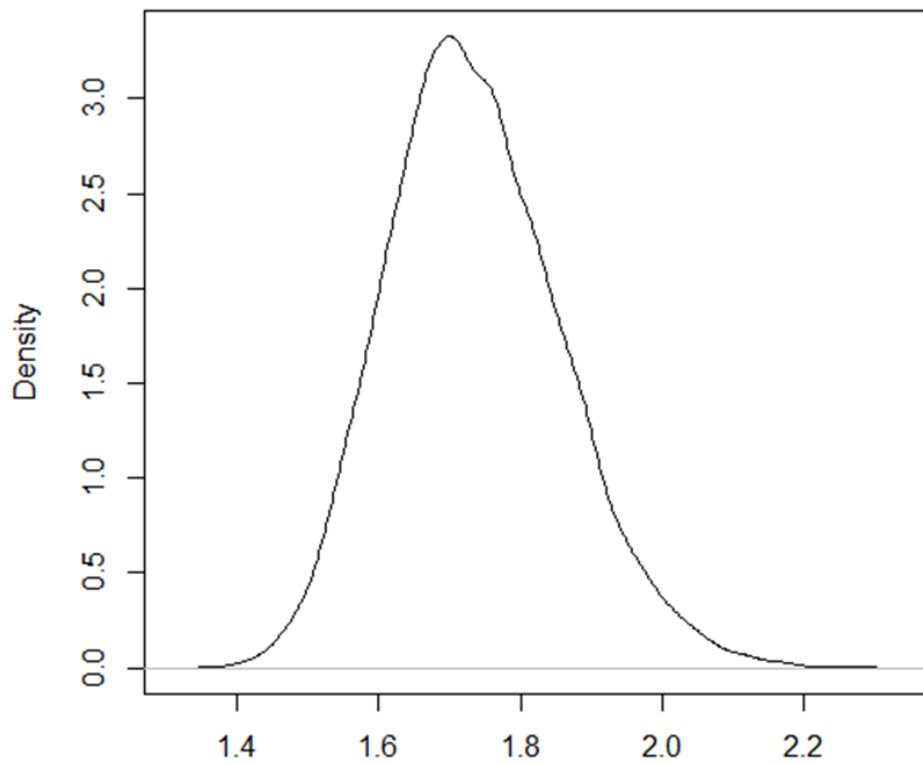


Figura 3.5.3 Vija e densitetit empirik e shpërndarjes lognormale për $R=1000$ zgjedhje bootstrap dhe $n=500$

Shohim nga figurat se me rritjen e vëllimit pjerrësia pozitive zvogëlohet, pra për $n=30$ vija e densitetit ka pjerrësi pozitive më të madhe sesa në rastin $n=500$.

Nga tabela shohim se pjerrësia është më e madhe për nivelin e sigurisë më të madhe.

Gjithsesi duke patur parasysh që futja e bootstrap sjell një korrektësi të rendit të dytë në shpërndarjet me pjerrësi pozitive, intervalet e ndërtuara nuk ndikohen shumë nga pjerrësia e shpërndarjes.

KAPITULLI IV

ZBATIMI I PËRAFRIMIT EDGEWORTH DHE METODËS BOOTSTRAP NË EKONOMETRI. APLIKIMI I PËRAFRIMIT EDGEWORTH DHE BOOTSTRAP NË ANALIZËN E REGRESIT

4.1 Këndvështrim i përgjithshëm mbi Ekonometrinë

Ekonometria e ka zanafillën e saj qysh në vitin 1930, ku fillimisht Shoqata e Ekonometristëve dhe Revista “Ekonometrika” e konsideruan atë si shkencë, e cila bën lidhjen e aspekteve teorike të Ekonomisë me fushën ekzakte të matematikës. Teoria bashkëkohore e përkufizon Ekonometrinë si studim analitik i fenomeneve me natyrë ekonomike, nëpërmjet zbatimit të metodave statistikore ([2], [3]).

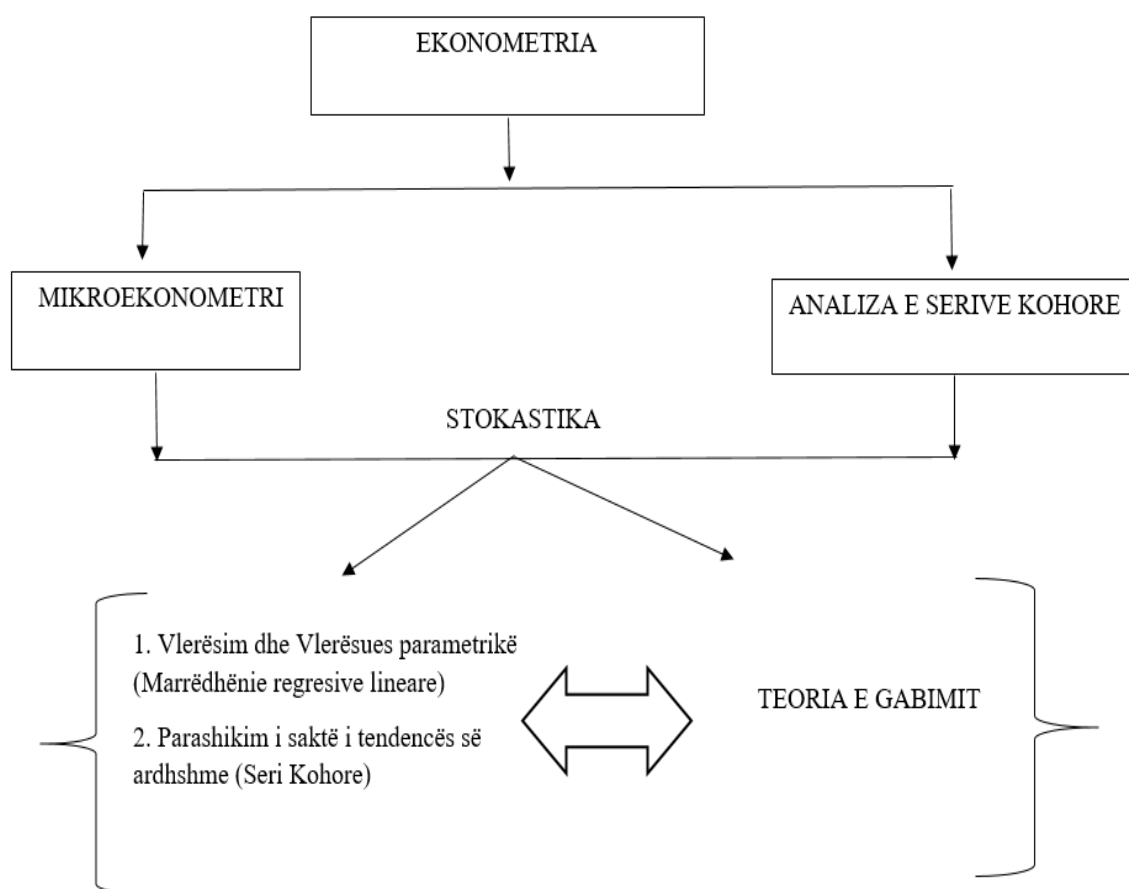


Figura 4.1.1 Këndvështrim mbi Shkencën e Ekonometrisë

Ashtu sikurse mund të evidentohet nga Figura 4.1.1, dy themelet bazë të shkencës së Ekonometrisë janë: Mikroekonometria dhe Analiza e Serive Kohore. Rezultatet që gjenerohen prej analizës së të dhënave në rastin e Mikroekonometrisë kanë tendencë heterogjeniteti dhe divergjimi, ndërkohë që rezultatet që gjenerohen nga Analiza e Serive Kohore, kanë tendencë konvergjuese dhe homogjeniteti, duke mundësuar kështu një prognozim optimal të serive. Pra, thelbi i Ekonometrisë qëndron në grumbullimin e të dhënave për një fenomen me natyrë ekonomike dhe përpunimi statistikor i këtyre të dhënave për të finalizuar me sukses një studim empirik.

Marrëdhënia mes ndryshoreve të marra në studim ka natyrë stokastike, pasi gjithë studimet ekonometrike marrin në konsideratë termin e gabimit. Pra, analiza e regresit ose analiza e serive kohore, shoqërohet me termin e gabimit, pasi mund të ndodhë që gjatë studimit të :

a) Neglizhohen një tërësi ngjarjesh rasti ose ndryshoresh, të cilat mund të kenë ndikim të konsiderueshëm tek analiza e regresit. Në këtë mënyrë shpjegueshmëria e ndryshores së varur nga ndryshoret e pavarura nuk është optimale.

b) Ekzistojnë probleme të tilla, si: homoskedasticiteti, autokorrelacioni dhe multikolineariteti.

c) Shfaqen probleme për njehsimin e drejtë të ndryshores së varur për shkak të përcaktimit jo korrekt të modelit që do të jetë subjekt i studimit

d) Realizohen gabime humane në lidhje me konceptimin jo të drejtë të modelit që do të studiohet ose mospërfshirjen e ndryshoreve me rëndësi në model.

Pra, ndonëse ekzistenca e termit të gabimit është e pashmangshme, sërish synimi i çdo studiuesi është që të gjenerojë vlerësime të sakta, duke minimizuar variancën. Ky synim do të realizohet nëpërmjet aplikimit të metodave statistikore. Nëse krijimi i një modeli ekonomik që do të jetë subjekt i një studimi statistikor është pjesë e aspektit shkencor të Ekonometrisë, zgjedhja e metodave të përshtatshme statistikore është elementi që përbën aspektin e “artit” të Ekonometrisë.

Që një metodë statistikore të jetë efektive, duhet të plotësojë disa kritere si:

1. Të ketë kosto minimale

Metoda statistikore që përdoret për të përpunuar rezultatet duhet të jetë efektive në përdorim, në mënyrë që të kursejë kosto dhe përpjekje.

2. Të vlerësojë saktë marrëdhëniet mes ndryshoreve

Metoda statistikore duhet të vlerësojë me saktësi të lartë marrëdhëniet lineare ose jolineare mes ndryshoreve të modelit të marrë në studim. Qasja e përpunimeve statistikore duhet të shfaqë tendencën drejt reduktimit të termit të gabimit.

3. Të gjenerojë një koeficient përcaktueshmërie të lartë

Koeficienti i përcaktueshmërisë përcakton se ç’pjesë e variancës totale të ndryshores së varur shpjegohet linearisht nga ndryshoret e pavarura. Sa më e lartë të jetë vlera e këtij koeficienti, aq më e lartë është shpjegueshmëria e ndryshores së varur nga ndryshoret e pavarura, dhe rrjedhimisht, saktësia e modelit rritet.

4. Të jetë objektiv

Objektiviteti i një metode statistikore konsiston në zgjedhje rastësore të përsëritura, ku vlerat e ndryshoreve të pavarura konsiderohen konstante dhe përfitohen vlera të reja të ndryshores së varur dhe termit të gabimit. Shpërndarja e vlerave të gjeneruara të ndryshores së varur quhet Shpërndarje zgjedhore. Thelbi i këtij kriteri konsiston në konvergimin e një zgjedhjeje tek një vlerë e caktuar.

5. Të plotësojë kushtin e efijencës

Kushti i efijencës për një metodë statistikore përmbushet, atëherë kur përpunimet e vazhdueshme të të dhënave sjellin minimizimin e variancës së modelit të ndërtuar. Nëse metoda që përdoret është njëkohësisht objektive, efijente dhe kosto-minimizuese, atëherë ajo konsiderohet metoda më e mirë.

6. Të reduktojë gabimin mesatar kuadratik

Që një metodë të ketë epërsi ndaj metodave të tjera, duhet të reduktojë gabimin mesatar kuadratik, i cili shpreh ndryshimin në katror të vlerave të ndryshores së varur nga vlera mesatare e kësaj ndryshoreje.

7. Të ketë veti asimptotike

Nëse nuk është e mundur gjetja e një metode objektive, atëherë synimi do të jetë zgjedhja e një metode statistikore e cila gëzon veti asimptotike. Dy cilësitë kryesore që shoqërojnë vetitë asimptotike janë:

- a) Konsistenca, e cila ka të bëjë me reduktimin e gabimit mesatar kuadratik
- b) Efijenca, e cila siguron gjenerimin e shumës së katrorëve të gabimit me tendencë konvergimi drejt zeros.

8. Të ketë përgjasi maksimale

Nëse metoda e zgjedhur gëzon këtë veti, mund të konkludohet se ekziston lehtësi në përpunimin e një vëllimi të konsiderueshëm të dhënash. Vetitë që shoqërojnë këtë kriter janë

- a) Konsistenca
- b) Efijenca
- c) Objektiviteti
- d) Shpërndarje asimptotike
- e) Reduktim i gabimit mesatar kuadratik

9. Të ndërthurret me metoda rizgjedhëse si Monte Carlo

Metoda Monte Carlo realizohet nëpërmjet simulimeve kompjuterike, për një vëllim të konsiderueshëm të dhënash dhe për një numër të caktuar rizgjedhjesh. Thelbi i këtij studimi qendron në:

- a) Organizimi i të dhënave
- b) Implementimin e një numri të caktuar rizgjedhjesh
- c) Gjenerimi i rezultateve
- d) Krahasimi i rezultateve nga simulimet Monte Carlo me rezultatet e metodës statistikore të përdorur për përpunimin e të dhënave

Në këtë disertacion, modelet ekonomike të ndërtuara prej të dhënave konkrete financiare, janë analizuar nëpërmjet aplikimit të metodës përafuese Edgeworth dhe Metodave të rizgjedhjes Bootstrap dhe Monte Carlo. Përafrimi Edgeworth plotëson kriterin e kosto-minimizimit, efijencës si dhe objektivitetit duke përmbushur kështu

tre kriteret më të rëndësishme të efektivitetit të metodës statistikore . Gjithashtu metodat e rizgjedhjes të përdorura në këtë punim përmbushin kriterin e rritjes së saktësisë së rezultateve nëpërmjet një numri të caktuar rizgjedhjesh dhe simulimesh.

Në këtë kapitull do të shohim dy studimet praktike në kuadrin Ekonometrik:

- a) Kontributi i përafrimeve Edgeworth dhe Bootstrap në marrëdhënien regressive mes çmimit dhe sasive të shitura të produkteve ushqimore në një prej bizneseve më të fuqishme të qytetit të Korçës, EuroMarket
- b) Implementimi i përafrimit Edgeworth dhe Bootstrap për studimin e trendit të fitimeve prej shitjes së produkteve kozmetike të një prej dyqaneve më të njohura kozmetike në Korçë

Ndërsa në kapitullin pasardhës, kapitulli V, do të shohim një tjetër studim praktik në kuadrin Ekonometrik:

- a) Kontributi i përafrimeve Edgeworth, Bootstrap dhe Monte Carlo në parashikimin e Serive Kohore në termat e Modelit Autoregresiv për të dhëna financiare të kompanisë me sukses global: Coca-Cola

4.2 Marrëdhënia e Regresit Linear dhe Kuantil

Regresi Linear është një marrëdhënie algjebrike lineare mes ndryshores së varur y_i dhe ndryshoreve të pavarura x_i . Matematikisht, kjo marrëdhënie shprehet:

$$y_i = b_0 + b_1 * x_i + \varepsilon .$$

Parametrat b_0 dhe b_1 janë konstante që lidhin ndryshoret në modelin e Regresit Linear. Metoda më e njohur për njehsimin e këtyre parametrave është Metoda e Katrorëve më të Vegjël.

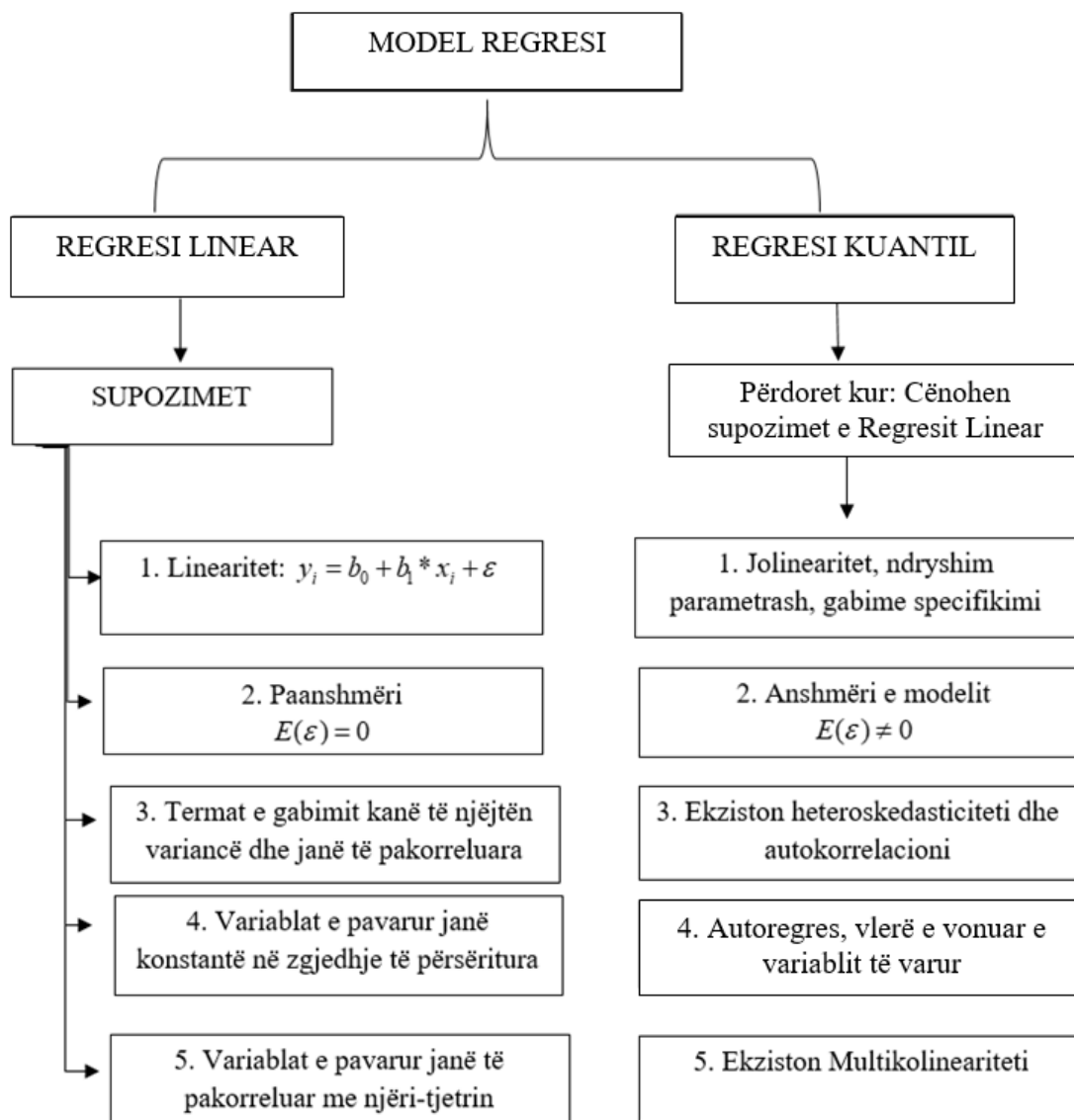


Figura 4.2.1 Regresi Linear dhe Kuantil

Pra, sikurse evidentohet nga Figura 4.2.1, kur supozimet bazë të modelit të Regresit Linear cënohen, do të zbatohet Modeli i Regresit Kuantil.

Disa nga problemet që shfaqen tek Modeli i Regresit Linear në momentin e mospërbushjes së disa prej kushteve më të rëndësishme janë: Autokorrelacioni, Heteroskedasticiteti dhe Multikolineariteti [37].

Tabela 4.2.1 Autokorrelacioni, Multikolariteti dhe Heteroskedasticiteti

| | Multikolaritet | Autokorrelacion | Heteroskedasticitet |
|----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>Kuptimi</i> | Lidhje perfekte mes disa/gjithë ndryshoreve të pavarura | Elementët jashtë matricës së kovariancës ndryshojnë nga zero. Ndodh lidhje me vonesë e një serie mbetjesh me veten e saj | Variancat tek diagonalja e matricës së kovariancës janë të ndryshme. Pra për çdo vrojtim ka shpërndarje të ndryshme gabimesh |
| <i>Shkaku</i> | Njëra ndryshore shprehet si funksion linear i ndryshores tjetër, ku koeficientët para tyre nuk mund të jenë asnjëherë zero njëkohësisht | <ol style="list-style-type: none"> 1. Përrjashtim i ndryshoreve të rëndësishme nga modeli (korrelacion serial) 2. Specifikim i gabuar i modelit 3. Gabim matje 4. Influenca të fenomeneve shkatërruese 5. Manipulim të dhënash 6. Vonesat | <ol style="list-style-type: none"> 1. Minimizim variance për shkak të përvetësimit të gabimit ose përmirësimit teknik të grupimit të të dhënave 2. Variancë e lartë për shkak të: <ul style="list-style-type: none"> - Specifikimit të gabuar -Shpërndarjes jonormale të regresorëve -Transformimeve të gabuara funksionale -Përfshirje e vlerave ekstreme -Shtim i ndryshoreve të varura |
| <i>Pasoja</i> | <ol style="list-style-type: none"> 1. Rritje e variancës dhe kovariancës mes ndryshoreve të pavarura 2. Zgjerim i Intervaleve të besimit (ul besueshmërinë) 3. Ul koeficientin e përcaktueshmërisë 4. Ndryshim thelbësor i koeficientëve dhe vlerave 5. Largim i ndryshoreve të pavarura të rëndësishme nga modeli | <ol style="list-style-type: none"> 1. Paanshmëria dhe konsistenca nuk ndikohen 2. Vlerësuesi është joefiçent 3. Rritet varianca e modelit 4. Mbivlerësim i koeficientit të përcaktueshmërisë dhe rezultatit të testit të studentit | <ol style="list-style-type: none"> 1. Vlerësuesi është i paanshëm dhe konsistent, pasi ndryshoret e pavarura nuk korrelojnë me mbetjet 2. Ul efiçencen sepse rritet varianca 3. Mbivlerësim i Testeve të Studentit dhe Fisher 4. Zgjerim i Intervaleve të Besimit |

| | | | |
|---------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><i>Evidenti mi</i></p> | <p>1. Rritje e koeficientit të korrelacionit 2. Kur rezultati i regresit ndihmës është më i lartë se rezultati kritik</p> | <p>1. Metoda grafike (tendenca ciklike evidenton autokorrelacion pozitiv) 2. Testi Darbin-Watson (përmban term të lirë, rend i parë i autokorrelacionit $u_t = \rho * u_{t-1} + \varepsilon$, nuk përmban ndryshore të varura me vonesë) 3. LM Breuch-Pagan Test (Rend të lartë autokorrelacioni $u_t = \rho_1 * u_{t-1} + \rho_2 * u_{t-2} + \dots + \rho_p * u_{t-p} + \varepsilon$, lejon ndryshore të varura me vonesë, jep përgjigje për çdo zonë) 4. Testi “h”- Vepron në prani të ndryshoreve të varura me vonesë</p> | <p>-Nevojitet intuita, eksperiencia dhe teoria për të evidentuar këtë problem 1. Metoda formale LM-identifikimi i heteroskedasticitetit nëpërmjet këtyre testeve: a) Breuch-Pagan, b) Glesjer, c) Harvey Godfrey d) Goldfeld Quard e) White tests 2. Metoda informale- jep paraqitje grafike të shpërndarjes së vlerësimit të gabimeve kundrejt ndryshoreve të pavarura dhe ndryshores së varur</p> |
| <p><i>Elimini mi</i></p> | <p>1. Referim tek vlera më e lartë e regresit bazë krahasuar me regresin ndihmës dhe tek vlera më e lartë se 2 e testit të studentit 2. Përfshirje e më tepër informacioni si: -Rritje e të dhënave, për të ulur variancën -Përcaktim i marrëdhënieve jolineare mes ndryshoreve të pavarura -Specifikim i marrëdhënies mes parametrave - Përjashtim i një ndryshoreje “të parëndësishme” nga modeli - Formim i komponentëve thelbësorë (grupim i ndryshoreve kolineare)</p> | <p>1. Kur njihet “ρ” do të kryhen transformime që të mos humbin vrojtimit 2. Kur ρ nuk njihet do të ndiqen procedurat e : a) Procedura përsëritëse të Cochrane, e cila konsiston në gjetjen e ρ optimale, e cila do të sjelle konvergjimim e vlerësimit të mbetjeve me vlerësimin e vërtetë të mbetjeve b) Procedura Kërkuese e Hildreth-Lu, e cila konsiston në kërkimin dinamik të ρ optimale dhe zgjedhjen e asaj ρ që minimizon Shumën e Katrorëve të Gabimit</p> | <p>1. Rivlerësim i Modelit me synim rritjen e efiçences 2. Korrigjim i pasaktësive në termat e kovariancës dhe testit të studentit</p> |

| | | |
|-----------------------------------------------------------------------------------|--|--|
| -Analizëfaktoriale - Kombinim linear mes ndryshoreve me kolinearitete të lartë | | |
|-----------------------------------------------------------------------------------|--|--|

Për të liruar modelin e regresit linear prej problemeve të trajtuara në Tabelën 4.2.1, zbatohet Modeli i Regresit Kuantil, i cili është një version më i thelluar i Modelit të Regresit Linear, të njehsueshëm nëpërmjet Metodës së Zakonshme të Katrorëve më të Vegjël.

4.3 Implementimi i përafrimit Edgeworth për versionin Bootstrap

Në këtë çështje do të trajtojmë rritjen e saktësisë së parashikimit të fenomeneve konkrete me natyrë ekonomike, duke përdorur Metodën e përafrimit Edgeworth për versionin Bootstrap. Studimi konsiston në identifikimin gradual të saktësisë së rezultateve, deri në arritjen e një përfundimi që siguron shpjegueshmëri të lartë dhe variancë të ulët, shih ([22], [29]).

4.3.1 Vlerësimi i Gabimit mesatar Kuadratik

Le të shënojmë me $X = \{X_1, \dots, X_n\}$ një zgjedhje rasti me vëllim n , me funksion shpërndarje F , dhe me

$$\hat{F}(x) = n^{-1} \sum_{i=1}^n I(X_i \leq x)$$

shënojmë funksionin e shpërndarjes empirike të zgjedhjes [50]. Vlerësimi bootstrap për kuantilin “ p ” të F , $\hat{\xi}_p = \hat{F}^{-1}(p)$, është

$$\hat{\xi}_p = \hat{F}^{-1}(p) = \inf \{x : \hat{F}(x) \geq p\} = X_{nr} \quad (4.3.1)$$

ku, $X_{n1} \leq \dots \leq X_{nm}$ tregon rendin statistikor të X , dhe $r = [np]$ është numri i plotë, jo më i madh se np .

Gabimi Mesatar Kuadratik i $\hat{\xi}_p$ jepet në trajtën

$$\begin{aligned} \tau^2 &= E \left\{ (\hat{\xi}_p - \xi_p)^2 \right\} \\ &= \frac{n!}{(r-1)!(n-r)!} \int_{-\infty}^{\infty} (x - \xi_p)^2 F(x)^{r-1} \{1 - F(x)\}^{n-r} dF(x) \\ &= r \binom{n}{r} \int_0^1 \{F^{-1}(u) - F^{-1}(p)\}^2 u^{r-1} (1-u)^{n-r} du, \end{aligned}$$

dhe vlerësimi bootstrap i tij është

$$\begin{aligned}\hat{\tau}^2 &= r \binom{n}{r} \int_0^1 \left\{ \hat{F}^{-1}(u) - \hat{F}^{-1}(p) \right\}^2 u^{r-1} (1-u)^{n-r} du \\ &= \sum_{j=1}^n (X_{nj} - X_{nr})^2 w_j,\end{aligned}\tag{4.3.2}$$

$$\text{ku, } w_j = r \binom{n}{r} \int_{(j-1)/n}^{j/n} u^{r-1} (1-u)^{n-r} du.$$

Vlerësimi bootstrap i variancës detajohet si në vijim. Vlerësimet bootstrap për mesataren dhe mesataren kuadratike të $\hat{\xi}_p$, janë respektivisht

$$\begin{aligned}r \binom{n}{r} \int_0^1 \hat{F}^{-1}(u) u^{r-1} (1-u)^{n-r} du &= \sum_{j=1}^n X_{nj} w_j, \\ r \binom{n}{r} \int_0^1 \hat{F}^{-1}(u)^2 u^{r-1} (1-u)^{n-r} du &= \sum_{j=1}^n X_{nj}^2 w_j.\end{aligned}$$

Prandaj, vlerësimi bootstrap i variancës është

$$\sum_{j=1}^n X_{nj}^2 w_j - \left(\sum_{j=1}^n X_{nj} w_j \right)^2.$$

4.3.2 Zbërthimi Edgeworth për Vlerësimin Bootstrap të Kuantileve

Versionet Bootstrap të ξ_p dhe τ^2 janë $\hat{\xi}_p$ dhe $\hat{\tau}^2$, të përcaktuara tek (4.3.1) dhe (4.3.2). Shpërndarja Normale Standarde përcakton limitin e $(\hat{\xi}_p - \xi_p) / \hat{\tau}$, sepse $\hat{\xi}_p$ ka shpërndarje normale (asimptotike) $N(0, \tau^2)$, dhe $\hat{\tau}^2 / \tau^2 \rightarrow 1$ konvergjon në 1 [19]. Zbërthimi Edgeworth që do të shohim tani për shkallën e konvergjencës në këtë kufi dhe vetitë e termave në këtë zbërthim nuk janë të ngjashme me ato të trajtuara në Kapitullin 2 të [50], sepse rendi i $\hat{\tau}^2 \tau^{-2} - 1$ është $n^{-1/4}$ dhe jo $n^{-1/2}$. Rendi i termit të parë të kësaj shpërndarjeje është $n^{-1/2}$, jo $n^{-1/4}$, por polinomi i koeficientëve të atij termi nuk është as çift e as tek; ndërkohë në rastet e shqyrtuara më parë, polinomi për termin me rend $n^{-1/2}$ zakonisht është çift.

Ne mund të përcaktojmë

$$\begin{aligned}q(x) &= \frac{1}{4} \left\{ \pi p(1-p) \right\}^{-1/2} x(x^2 + 1 + 2^{3/2}) \\ &\quad + \frac{1}{6} \left\{ p(1-p) \right\}^{-1/2} (1+p)(x^2 - 1) \\ &\quad - \left[\left\{ p(1-p) \right\}^{-1/2} p + \left\{ p(1-p) \right\}^{1/2} f'(\xi_p) f(\xi_p)^{-2} \right] x^2 \\ &\quad - \left\{ p(1-p) \right\}^{-1/2} \left\{ \frac{1}{2}(1-p) + r - np \right\}\end{aligned}$$

për të evidentuar më qartë rezultatin. Gjithashtu,

$$P \left\{ (\hat{\xi}_p - \xi_p) / \hat{\tau} \leq x \right\} = \Phi(x) + n^{-1/2} q(x) \phi(x) + O(n^{-3/4})$$

kur $n \rightarrow \infty$ [51]. Polinomin q është i rendit 3 dhe nuk konsiderohet as funksion tek as çift.

Zbërthimi Edgeworth për kuantilet e studentizuara krijohet nga seria “bazë” e termave që zvogëlohen me rend $n^{-1/2}$, që vjen nga numëruesi në raportin e testit të studentit, së bashku me një seri “të re” që del nga emëruesi.

Rendi i serisë “së re” zvogëlohet në $n^{-1/4}$, sepse gabimi relativ i vlerësimit të variancës është $n^{-1/4}$. Termi i j -të është çift ose tek, në përputhje me faktin nëse j është tek ose çift, përkatësisht, për të dy seritë. Nëse i referohemi serive të marra në shqyrtim, rendi $n^{-1/2}$ përfshin termin e parë tek të serisë “bazë” dhe termin e dytë çift të serisë “së re”. Hall dhe Sheather [52] diskutuan zbërthimin Edgeworth për shpërndarjen e një kuantile të studentizuar kur vlerësimi i devijimit standard bazohet në vlerësuesin e densitetit, të llojeve të përshkruara tek [8] dhe [76].

4.3.3 Modeli i Regresit Kuantil

Metodat e modelimit të shpërndarjes së kushtëzuar në aspekte të tjera paraqesin interes të veçantë, përveç modeleve të regresit që analizojnë mesataren e kushtëzuar të një ndryshoreje të varur. Regresi kuantil që modelon kuantilet e ndryshores së varur për një sërë ndryshoresh të kushtëzuara, është një nga modelet më të spikatura ([55], [56]).

Regresi kuantil vlerëson marrëdhënien lineare midis regresorëve X dhe Y , një kuantil i specifikuar i ndryshores së varur, [62]. Vlerësuesi i devijimeve më së paku absolute (LAD) është një rast i veçantë i regresit kuantil i cili rregullon mesoren e kushtëzuar të ndryshores së varur.

Përshkrimi i faktit se si mesatarja ose përqindjet e ndryshores së varur ndikohen nga ndryshimi i regresorëve ofrohet nga regresi kuantil.

4.3.4 Rasti i ndryshoreve të pavarura dhe identikisht të shpërndara (IID)

Koenker dhe Bassett [62] nxorrën rezultate asimptotikisht normale për vlerësimin e regresit kuantil në rastin e ndryshoreve të pavarura dhe identikisht të shpërndara (i.i.d), duke treguar se në kushte të rregullta dhe të qendrueshme,

$$\sqrt{n}(\hat{\beta}(\tau) - \beta(\tau)) \sim N(0, \tau(1-\tau)s(\tau)^2 J^{-1}) \quad (4.3.3)$$

ku,

$$J = \lim_{n \rightarrow \infty} \left(\sum_i X_i X_i' / n \right) = \lim_{n \rightarrow \infty} (X' X / n)$$

$$s(\tau) = F^{-1}(\tau) = 1 / f(F^{-1}(\tau))$$

dhe $s(\tau)$ është derivati i funksionit kuantil ose funksioni i anasjelltë i densitetit i vlerësuar për kuantilin τ , njihet si funksioni i densitetit kuantil ose funksioni i sparsitetit [84].

Vlerësimi direkt i përcaktorit të matricës së kovariancës, ndikohet nga vlera e densitetit për një kuantil të dhënë. Formula për kovariancën asimptotike tek (4.3.3) është e njëjtë me formulën e njehsimit të kovariancës sipas Metodës së Zakonshme të Katrorëve Më të Vegjël për rastin e ndryshoreve të pavarura dhe identikisht të shpërndara, ku $\tau(1-\tau)s(\tau)^2$ lidhet me njehsimin e variancës.

Rasti i zakonshëm i.i.d ka të bëjë me përafrimin Edgeworth dhe është shumë i përdorshëm në minimizimin e gabimit standard të koeficientit të pavarur.

4.3.5 Vlerësimi i sparsitetit

Rëndësia e funksionit të sparsitetit identifikohet nëpërmjet njehsimit të përcaktorit të matricës së kovariancës për të dhënat e pavarura dhe identikisht të shpërndara (i.i.d.). Meqë sparsiteti është funksion i një shpërndarjeje të panjohur F , është i nevojshëm një vlerësim sasior konkret, ([55], [56]).

Tre metodat që paketa Softuerike EViews 10 ofron për vlerësimin e sparsitetit skalar $s(\tau)$ janë: Siddiqui, i cili bazohet tek ndryshimi i herësive, Basset dhe Vlerësuesi i densitetit të bërthamës, ([72], [10], [59], [76], [63], [5]).

4.3.6 Metodat Bootstrap

Metodat Bootstrap nuk konsiderojnë njehsimin e sparsitetit në një pikë të dhënë ose në një zgjedhje të caktuar për të vlerësuar matricën e kovariancës, sepse shmangin vlerësimin asimptotik të kësaj matrice.

Katër metodat e ndryshme Bootstrap që EViews vë në dispozicion janë: Bootstrap i Mbetjeve (Residual), Bootstrap i çiftit XY (XY pair) dhe dy variantet e Bootstrap Margjinal i zinxhirit Markov (Markov Chain Marginal Bootstrap, MCMB), ([10], [60], [53]).

Bootstrap i Mbetjeve (Residual Bootstrap)

Bootstrap i Mbetjeve bazohet në kryerjen e një numri të caktuar rizgjedhjesh (me përsëritje), të mbetjeve $u_i(\tau)$ dhe X_i .

Le të jetë X^* një matricë me $m \times p$ përmasa, ku X janë të rizgjedhura në mënyrë të pavarur, dhe le të jetë u^* një vektor m –përmasor i mbetjeve të rizgjedhura. Pasi krijojmë ndryshoren e varur ne kryejmë vlerësimin bootstrap $\beta(\tau)$ duke përdorur Y^* dhe X^* , nga ekuacioni $Y^* = X^* \hat{\beta}(\tau) + u^*$, ([55], [56]). Nëse kjo procedurë përsëritet për M rizgjedhje bootstrap, vlerësimi asimptotik i matricës së kovariancës do të kishte formën

$$V(\hat{\beta}) = n \left(\frac{m}{n} \right) \frac{1}{B} \sum_{j=1}^B (\hat{\beta}_j(\tau) - \overline{\hat{\beta}(\tau)}) (\hat{\beta}_j(\tau) - \overline{\hat{\beta}(\tau)})' \quad (4.3.4)$$

ku $\overline{\hat{\beta}(\tau)}$ është mesatarja e elementëve bootstrap. Matrica bootstrap e kovariancës, $V(\hat{\beta})$ është një vlerësim i variancës së zgjedhur i vlerësimit bootstrap të $\beta(\tau)$. Pavarësia e u edhe X është kusht i rëndësishëm për pavarësinë e X_i dhe $u_i(\tau)$.

Bootstrap i Çiftit XY

Metoda më e përdorur për rizgjedhje është Bootstrap i Çiftit XY dhe kryesisht gjen zbatim kur u and X nuk janë të pavarura. Për të implementuar këtë metodë ne thjesht krijojmë një rizgjedhje B të ndërtuar rastësisht (me rikthim), me vëllim m nga të dhënat origjinale, më pas llogaritim vlerësimin e $\beta(\tau)$ duke përdorur (y^*, X^*) për

çdo rizgjedhje. Duke përdorur (4.3.4) mund të bëhet vlerësimi asimptotik i matricës së kovariancës nga varianca e zgjedhjes bootstrap.

4.3.7 Procesi i testimit të kuantileve

Analiza e modelit të regresit kuantil duke përfshirë më tepër se një kuantil do të përbënte një qasje interesante studimore. Për këtë arsye ne mund të krijojmë një hipotezë të përbashkët për të përdorur koeficientët e gjithë kuantileve që do të merren në konsideratë. Synimi i këtij procesi testimi është identifikimi i rëndësisë së këtyre kuantileve në të njëjtën kohë tek regresi [61].

Hapi i parë në këtë analizë është përcaktimi i teorisë së përshtatshme mbi shpërndarjen, para aplikimit të testimit të hipotezave. Forma vektoriale e koeficientit të procesit të testimit të kuantileve ka këtë formë

$$\beta = (\beta(\tau_1)', \beta(\tau_2)', \dots, \beta(\tau_K)')' .$$

Atëherë

$$\sqrt{n}(\hat{\beta} - \beta) \sim N(0, \Omega)$$

ku Ω ka blloqe të formës:

$$\Omega_{ij} = [\min(\tau_i, \tau_j) - \tau_i \tau_j] H^{-1}(\tau_i) J H^{-1}(\tau_j) . \quad (4.3.5)$$

Në rastin i.i.d. Ω thjeshtëzohet në:

$$\Omega = \Omega_0 \otimes J \quad (4.3.6)$$

ku, Ω_0 ka element përfaqësues

$$\omega_{ij} = \frac{\min(\tau_i, \tau_j) - \tau_i \tau_j}{f(F^{-1}(\tau_i))(f(F^{-1}(\tau_j)))} . \quad (4.3.7)$$

Duke përdorur formulat (4.3.5), (4.3.6) dhe (4.3.7) dhe duke aplikuar njëren nga metodat bootstrap, ne mund të njehsojmë Ω .

4.4 Kontributi i Përafrimit Edgeworth në një rast real studimor

Rasti konkret i marrë në studim bazohet në të dhënat e grumbulluara nga një prej bizneseve më të suksesshme që operojnë në qytetin e Korçës, për 129 vëzhgime të çmimeve për një diversitet produktesh dhe kërkesën përkatëse konsumatore për këto produkte përgjatë viteve 2019 dhe 2020, shih [22]. Paketa softuerike e përdorur për përpunimin e të dhënave është EViews10.

Tabela 4.4.1 Modeli i Regresit Kuantil për marrëdhënien Çmim-Kërkesë për vitin 2019

Dependent Variable: Demand
 Method: Quantile Regression (Median)
 Date: 02/04/21 Time: 22:43
 Sample: 1 129
 Included observations: 129
 Huber Sandwich Standard Errors & Covariance
 Sparsity method: Kernel (Epanechnikov) using residuals
 Bandwidth method: Hall-Sheather, bw=0.19228
 Estimation successfully identifies unique optimal solution

| Variable | Coefficien | t | Std. Error | t-Statistic | Prob. |
|------------------------|------------|--------------------|-----------------|-------------|--------|
| C | 1202.639 | | 368.8249 | 3.260731 | 0.0014 |
| Prices | -5.364852 | | 3.628374 | -1.478583 | 0.1417 |
| Pseudo R-squared | 0.034646 | Mean dependent var | | 1549.915 | |
| Adjusted R-squared | 0.027044 | S.D. dependent var | | 2535.739 | |
| S.E. of regression | 2608.706 | Objective | | 73061.39 | |
| Quantile dependent var | 624.0000 | Restr. objective | | 75683.50 | |
| Sparsity | 2134.090 | Quasi-LR statistic | | 9.829418 | |
| Prob(Quasi-LR stat) | 0.001717 | | | | |

Tabela 4.4.1 pasqyron modelin e Regresit Kuantil të marrëdhënies mes çmimeve dhe kërkesave konsumatore për një varietet produktesh përgjatë vitit 2019. Nga rezultatet e përftuara është e thjeshtë të identifikohet se marrëdhënia Çmim-Kërkesë Konsumatore është e fortë, por e zhdrejtë, për këtë arsye, nëse mediana e vlerës së çmimeve do të ketë një rritje prej 1%, kërkesa konsumatore do të ulet me një përafërsi prej 5%. Ndikimi i Çmimeve në Kërkesën Konsumatore nuk ka rëndësi të madhe statistikore, pasi probabiliteti përkatës prej 0.1417, është më i lartë se kriteri i rëndësisë statistikore prej 0.05. Gjithashtu, shpjegueshmëria e ndryshores së varur, të përfaqësuar nga Kërkesa Konsumatore, prej ndryshores së pavarur të përfaqësuar nga Çmimi i produkteve është 2.7%. Pra, 97.3% e variancës së Kërkesës shpjegohet prej ndryshoreve të tjera të pavarura, të cilat nuk janë marrë në konsideratë për këtë studim. Për këtë arsye, varianca e modelit është e lartë (1549.915).

Qëllimi për çdo studim që ndërmerret është optimalizimi i modelit në termat e reduktimit të gabimeve standard të secilit prej koeficientëve të përfshirë në model. Për të përmbushur këtë qëllim, në këtë studim u aplikuan Përafrimi Edgeworth për ndryshoret e pavarura identikisht të shpërndara (IID), versioni Bootstrap dhe Përafrimi Edgeworth për versionin Bootstrap.

Metodologjia e përdorur konsiston në përzgjedhjen e 30 kuantileve të modelit dhe identifikimin e gabimeve potenciale për secilin kuantil. Kjo metodologji u përdor për Modelin e Regresit Kuantil, dhe për modelin e regresit pas implementimit të: Përafrimit Edgeworth, Metodave Bootstrap dhe Përafrimit Edgeworth për versionin Bootstrap.

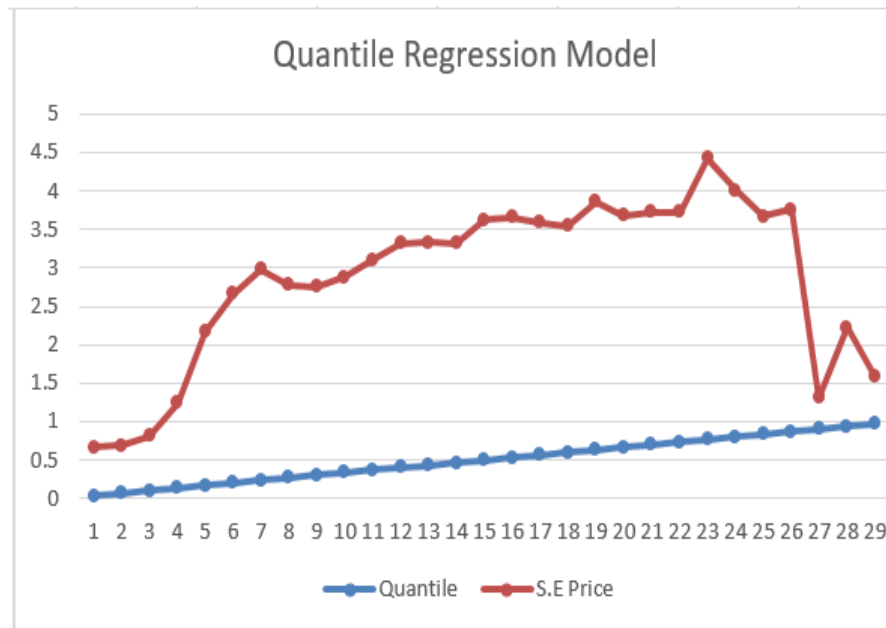


Figura 4.4.1 Shpërndarja e Devijimeve Standarde për 30 kuantile

Figura 4.4.1 pasqyron shpërndarjen e Devijimeve Standarde (Gabimeve) të modelit të marrë në shqyrtim për 30 kuantilet e përzgjedhura për zhvillimin e studimit. Identifikojmë se me rritjen e kuantileve, tendenca e shpërndarjes së gabimeve ka prirje rënëse.

Për të reduktuar devijimet standarde të modelit zbatojmë përafrimin Edgeworth.

Tabela 4.4.2 Përafrimi Edgeworth për Modelin e Regresit Kuantil

Dependent Variable: Demand
Method: Quantile Regression (Median)
Date: 02/04/21 Time: 23:05
Sample: 1 129
Included observations: 129
Ordinary (IID) Standard Errors & Covariance
Sparsity method: Siddiqui using residuals
Bandwidth method: Hall-Sheather, bw=0.19228
Estimation successfully identifies unique optimal solution

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|------------------------|-------------|--------------------|-------------|--------|
| C | 1202.639 | 184.3785 | 6.522664 | 0.0000 |
| Prices | -5.364852 | 1.676435 | -3.200156 | 0.0017 |
| Pseudo R-squared | 0.034646 | Mean dependent var | 1549.915 | |
| Adjusted R-squared | 0.027044 | S.D. dependent var | 2535.739 | |
| S.E. of regression | 2608.706 | Objective | 73061.39 | |
| Quantile dependent var | 624.0000 | Restr. objective | 75683.50 | |
| Sparsity | 2316.388 | Quasi-LR statistic | 9.055850 | |
| Prob(Quasi-LR stat) | 0.002619 | | | |

Pas zbatimit të Edgeworth-it në model (Tabela 4.4.2), devijimi standard i koeficientit të Çmimit ulet. Gjithashtu, Çmimet bëhen statistikisht të rëndësishme pasi

probabiliteti $0.0017 < 0.05$. Pra, modeli është përmirësuar ndjeshëm pas aplikimit të përafrimit Edgeworth.

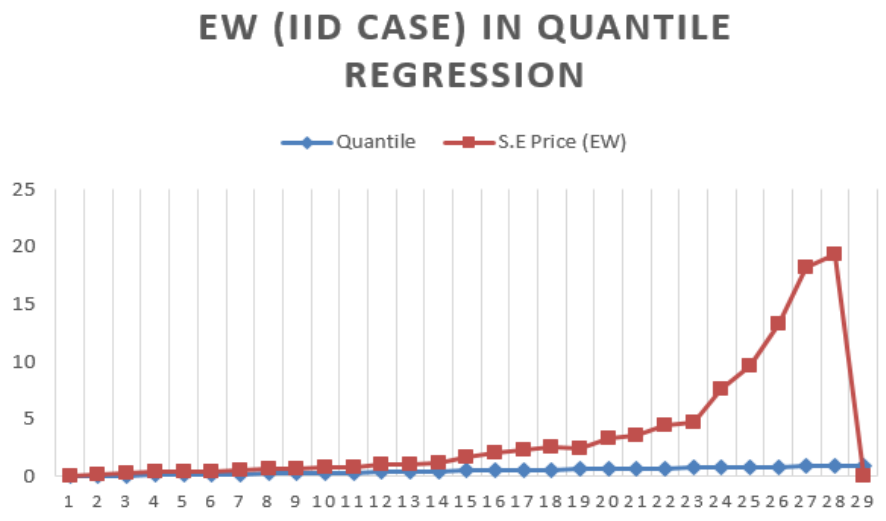


Figura 4.4.2 Shpërndarja e Devijimeve Standarde për 30 kuantile pas zbatimit të Përafrimit Edgeworth

Figura 4.4.2 pasqyron tendencën e dukshme rënëse të devijimeve standarde të modelit për secilin kuantil. Nëse do të krahasojm këtë tendencë me shpërndarjen e devijimeve standarde fillestare (Figurën 4.4.3), vërrehet se pas aplikimit të përafrimit Edgeworth, grafiku përafrohet me boshtin e abshisave duke marrë vlera të pakonsiderueshme të gabimeve.

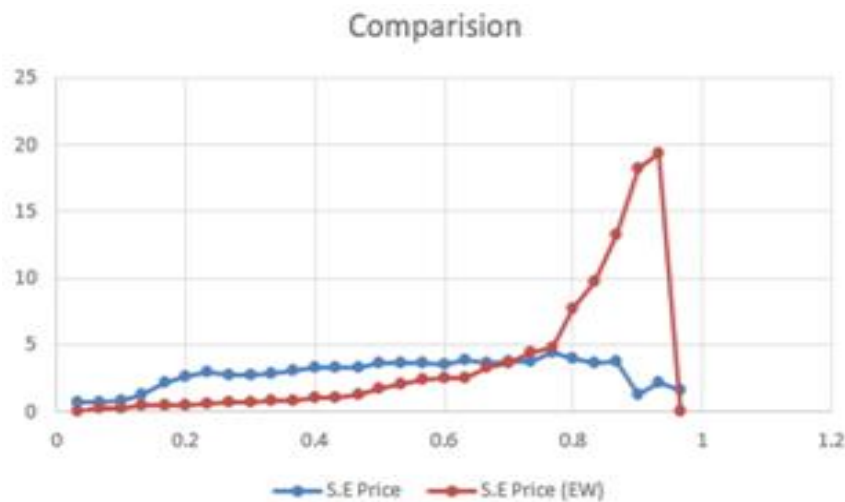


Figura 4.4.3 Këndvështrim krahasimor i tendencës së shpërndarjes së Devijimeve Standarde për çdo kuantil

Tabela 4.4.3 Zbatimi i Bootstrap tek Modeli i Regresit Kuantil

Dependent Variable: Demand
 Method: Quantile Regression (Median)
 Date: 02/04/21 Time: 23:25
 Sample: 1 129
 Included observations: 129
 Bootstrap Standard Errors & Covariance
 Bootstrap method: XY-pair, reps=100000, rng=kn,
 seed=1521391883
 Sparsity method: Siddiqui using residuals
 Bandwidth method: Hall-Sheather, bw=0.19228
 Estimation successfully identifies unique optimal solution

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|------------------------|-------------|--------------------|-------------|--------|
| C | 1202.639 | 339.4023 | 3.543402 | 0.0006 |
| Prices | -5.364852 | 2.943007 | -1.822915 | 0.0707 |
| Pseudo R-squared | 0.034646 | Mean dependent var | 1549.915 | |
| Adjusted R-squared | 0.027044 | S.D. dependent var | 2535.739 | |
| S.E. of regression | 2608.706 | Objective | 73061.39 | |
| Quantile dependent var | 624.0000 | Restr. Objective | 75683.50 | |
| Sparsity | 2316.388 | Quasi-LR statistic | 9.055850 | |
| Prob(Quasi-LR stat) | 0.002619 | | | |

Sikurse vërrehet nga Tabela 4.4.3, pas aplikimit të Versionit Bootstrap për 100,000 rizgjedhje, devijimet standarde kanë pësuar rënie krahasuar me modelin fillestar të Regresit Kuantil. Tendenca rënëse e devijimeve standard në model konkretizohet gjithashtu në Figurën 4.4.4.

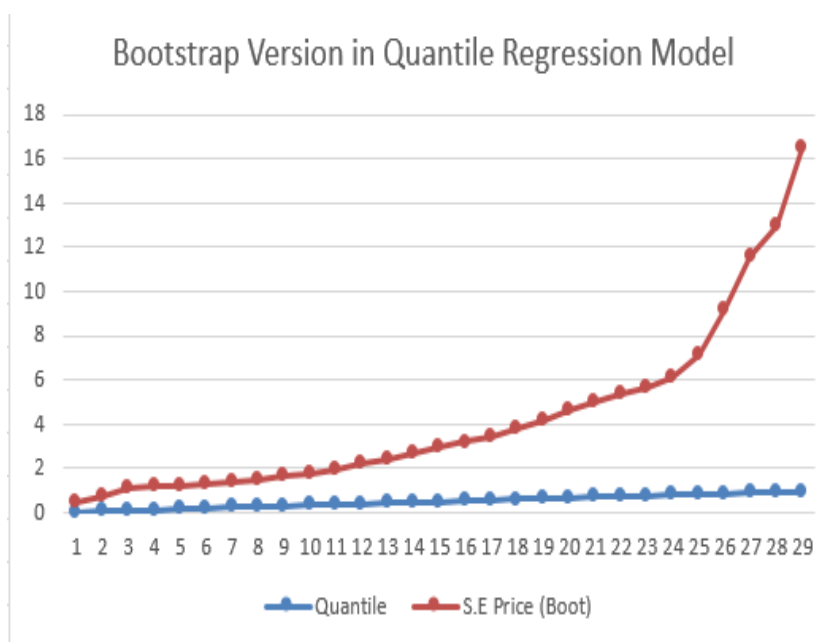


Figura 4.4.4 Shpërndarja e devijimeve standarde për secilin kuantil pas zbatimit të versionit Bootstrap

Nëse pas zbatimit të versionit Bootstrap të Modelit të Regresit Kuantil, ne implementojmë përafrimin Edgeworth, rëndësia statistikore e Çmimeve rritet ndjeshëm. Gjithashtu, rënia e devijimeve standarde është e konsiderueshme duke e bërë këtë model të optimalizuar. Këto rezultate konkretizohen në Tabela 4.4.4.

Tabela 4.4.4 Përafrimi Edgeworth i Versionit Bootstrap të Modelit

Dependent Variable: Demand
Method: Quantile Regression (Median)
Date: 02/04/21 Time: 23:40
Sample: 1 129
Included observations: 129
Ordinary (IID) Standard Errors & Covariance
Sparsity method: Kernel (Epanechnikov) using residuals
Bandwidth method: Hall-Sheather, bw=0.19228
Estimation successfully identifies unique optimal solution

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|------------------------|-------------|--------------------|-------------|--------|
| C | 1202.639 | 171.3953 | 7.016753 | 0.0000 |
| Prices | -5.364852 | 1.558387 | -3.442567 | 0.0008 |
| Pseudo R-squared | 0.034646 | Mean dependent var | 1549.915 | |
| Adjusted R-squared | 0.027044 | S.D. dependent var | 2535.739 | |
| S.E. of regression | 2608.706 | Objective | 73061.39 | |
| Quantile dependent var | 624.0000 | Restr. objective | 75683.50 | |
| Sparsity | 2153.278 | Quasi-LR statistic | 9.741827 | |
| Prob(Quasi-LR stat) | 0.001801 | | | |

Figura 4.4.4 tregon shpërndarjen e gabimeve standarde. Gjithashtu është lehtësisht e identifikueshme se tendenca e shpërndarjes së gabimeve është rënëse. Kështu, përafrimi Edgeworth i Versionit Bootstrap për Modelin e Regresit Kuantil ka përmirësuar modelin, ku rëndësia statistikore e Çmimeve në Kërkesa Konsumatore është rritur dhe devijimet standarde kanë një rënie të konsiderueshme. Në këtë mënyrë, marrëdhënia mes Çmimeve dhe Kërkesës Konsumatore është vlerësuar dhe përafruar në mënyrë të suksesshme nëpërmjet zbatimit të kësaj metode.

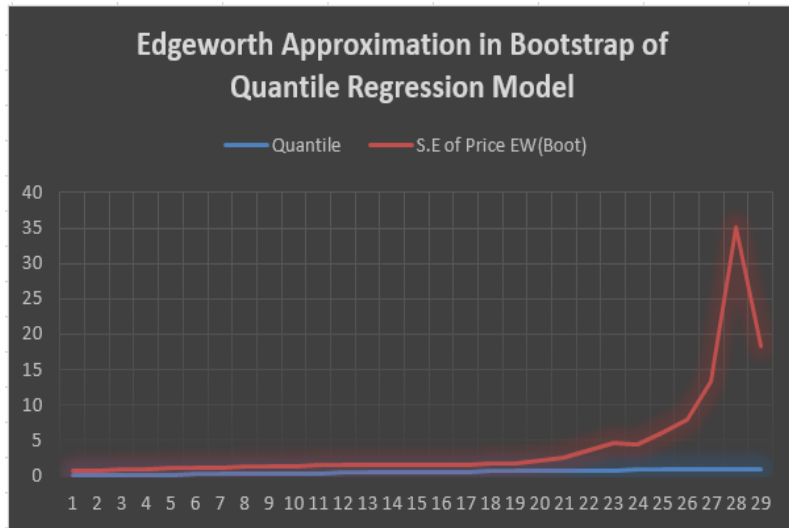


Figura 4.4.5 Shpërndarja e devijimeve standarde pas aplikimit të përafritit Edgeworth të Versionit Bootstrap

Prirja rënëse e shpërndarjes së devijimeve standarde për çdo kuantil pasqyrohet në Figurën 4.4.5.

Figura 4.4.6 përmbledh të gjitha rezultatet e arritura në një grafik të vetëm, duke mbivendosur të gjitha tendencat shpërndarëse të devijimeve standarde. Nga këto rezultate është evidente se përafriti Edgeworth i Versionit Bootstrap jep rezultatet më optimale për modelin e marrë në studim.



Figura 4.4.6 Grafik përmbledhës i rezultateve të këtij studimi

Nëse shqyrtojmë tendencën e vitit 2020, ajo është disi e ndryshme nga tendenca e gjeneruar e vitit 2019. Nëse vlera e medianës së Çmimit rritet me 1%, Kërkesa Konsumatore ulet me afro 3.5%. Duke qenë se Çmimet janë konstante, në tendencën rritëse të Kërkesës Konsumatore ka ndikuar periudha e Pandemisë e cila e ka bërë Kërkesën inelastike. Gjithashtu, shpjegueshmëria e Kërkesës Konsumatore në termat e Çmimeve është përmirësuar krahasuar me vitin 2019, duke shënuar një nivel prej

4%. Rrjedhimisht edhe varianca ka pësuar rënie të konsiderueshme në terma krahasimore.

Tabela 4.4.5 Modeli i Regresit Kuantil për marrëdhënien Çmim-Kërkesë për vitin 2020

Dependent Variable: Demand
 Method: Quantile Regression (Median)
 Date: 02/05/21 Time: 14:57
 Sample: 1 129
 Included observations: 129
 Huber Sandwich Standard Errors & Covariance
 Sparsity method: Kernel (Epanechnikov) using residuals
 Bandwidth method: Hall-Sheather, bw=0.19228
 Estimation successfully identifies unique optimal solution

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|------------------------|-------------|--------------------|-------------|--------|
| C | 1026.348 | 188.9552 | 5.431701 | 0.0000 |
| Price | -3.453567 | 1.393225 | -2.478829 | 0.0145 |
| Pseudo R-squared | 0.049662 | Mean dependent var | 1371.000 | |
| Adjusted R-squared | 0.042179 | S.D. dependent var | 1754.567 | |
| S.E. of regression | 1826.880 | Objective | 61602.81 | |
| Quantile dependent var | 639.0000 | Restr. objective | 64822.00 | |
| Sparsity | 1849.246 | Quasi-LR statistic | 13.92650 | |
| Prob(Quasi-LR stat) | 0.000190 | | | |

Përgjatë gjenerimeve vijuese është përdorur e njëjta logjikë e ndjekur edhe për vitin 2019.

Tabela 4.4.6 Përafrimi Edgeworth i Modelit të Regresit Kuantil për vitin 2020

Dependent Variable: Demand
 Method: Quantile Regression (Median)
 Date: 02/05/21 Time: 12:44
 Sample: 1 129
 Included observations: 129
 Ordinary (IID) Standard Errors & Covariance
 Sparsity method: Siddiqui using residuals
 Bandwidth method: Hall-Sheather, bw=0.19228
 Estimation successfully identifies unique optimal solution

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|------------------------|-------------|--------------------|-------------|--------|
| C | 1026.348 | 176.4725 | 5.815912 | 0.0000 |
| Price | -3.453567 | 1.597728 | -2.161548 | 0.0325 |
| Pseudo R-squared | 0.049662 | Mean dependent var | 1371.000 | |
| Adjusted R-squared | 0.042179 | S.D. dependent var | 1754.567 | |
| S.E. of regression | 1826.880 | Objective | 61602.81 | |
| Quantile dependent var | 639.0000 | Restr. objective | 64822.00 | |
| Sparsity | 2087.466 | Quasi-LR statistic | 12.33722 | |
| Prob(Quasi-LR stat) | 0.000444 | | | |

Tabela 4.4.7 Zbatimi Bootstrap i Modelit të Regresit Kuantil për vitin 2020

Dependent Variable: Demand
 Method: Quantile Regression (Median)
 Date: 02/05/21 Time: 12:46
 Sample: 1 129
 Included observations: 129
 Bootstrap Standard Errors & Covariance
 Bootstrap method: XY-pair, reps=100000, rng=kn,
 seed=1509069309
 Sparsity method: Siddiqui using residuals
 Bandwidth method: Hall-Sheather, bw=0.19228
 Estimation successfully identifies unique optimal solution

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|------------------------|-------------|--------------------|-------------|--------|
| C | 1026.348 | 259.6332 | 3.953070 | 0.0001 |
| Price | -3.453567 | 1.975663 | -1.748055 | 0.0829 |
| Pseudo R-squared | 0.049662 | Mean dependent var | 1371.000 | |
| Adjusted R-squared | 0.042179 | S.D. dependent var | 1754.567 | |
| S.E. of regression | 1826.880 | Objective | 61602.81 | |
| Quantile dependent var | 639.0000 | Restr. objective | 64822.00 | |
| Sparsity | 2087.466 | Quasi-LR statistic | 12.33722 | |
| Prob(Quasi-LR stat) | 0.000444 | | | |

Tabela 4.4.8 Përafrimi Edgeworth i versionit Bootstrap i Modelit për vitin 2020

Dependent Variable: Demand
 Method: Quantile Regression (Median)
 Date: 02/05/21 Time: 12:48
 Sample: 1 129
 Included observations: 129
 Ordinary (IID) Standard Errors & Covariance
 Sparsity method: Kernel (Epanechnikov) using residuals
 Bandwidth method: Hall-Sheather, bw=0.19228
 Initial Values: C(1)=0.00000, C(2)=0.00000
 Estimation successfully identifies unique optimal solution

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|------------------------|-------------|--------------------|-------------|--------|
| C | 1026.348 | 157.8247 | 6.503089 | 0.0000 |
| Price | -3.453567 | 1.428898 | -2.416945 | 0.0171 |
| Pseudo R-squared | 0.049662 | Mean dependent var | 1371.000 | |
| Adjusted R-squared | 0.042179 | S.D. dependent var | 1754.567 | |
| S.E. of regression | 1826.880 | Objective | 61602.81 | |
| Quantile dependent var | 639.0000 | Restr. objective | 64822.00 | |
| Sparsity | 1866.885 | Quasi-LR statistic | 13.79492 | |
| Prob(Quasi-LR stat) | 0.000204 | | | |



Figura 4.4.7 Rezultate përmblendhëse për Modelin e vitit 2020

Edhe në këtë rast, është e evidentueshme se përafrimi Edgeworth për Versionin Bootstrap të Modelit të Regresit Kuantil, është zgjidhja optimale, pasi redukton devijimin standard të Çmimit kundrejt Kërkesës Konsumatore.

4.5 Kontributi i Përafrimit Edgeworth për versionin Bootstrap në analizën e trendit të shitjeve për produktet kozmetike

Në këtë rast studimor janë konsideruar 130 vëzhgime për shitjet e produkteve kozmetike dhe fitimet e realizuara për vitet 2019 dhe 2020 për një biznes që operon në qytetin e Korçës. Objektivi i këtij studimi është identifikimi i trendit të fitimeve për produktet kozmetike në termat e një marrëdhënie potenciale mes shitjeve dhe fitimeve. Arsyeja pse nuk është përdorur Modeli i Regresit Linear lidhet me faktin e jolinearitetit, heteroskedasticitetit dhe autokorrelacionit që shoqërojnë këtë model, shih [29]. Këto probleme që shoqërojnë Modelin e Regresit identifikohen kur Katrori i Vëzhgimeve është më i vogël se 0.05, dhe në këtë rast ky faktor ka vlerë 0.0000. Në këtë mënyrë, hipotezat bazë të cilat verifikojnë inekzistencën e autokorrelacionit dhe heteroskedasticitetit nuk mund të pranohen. Meqë probabiliteti i Hi-Katrorit në këto teste është $0.0000 < 0.05$ provon ekzistencën e këtyre problemeve. Këto janë të pasqyruara në tabelat 4.5.1, 4.5.2, 4.5.3 dhe 4.5.4.

Rezultatet lidhur me jolinaritetin e modelit të regresit linear konkretizohen tek Figura 4.5.1 dhe Figura 4.5.2.

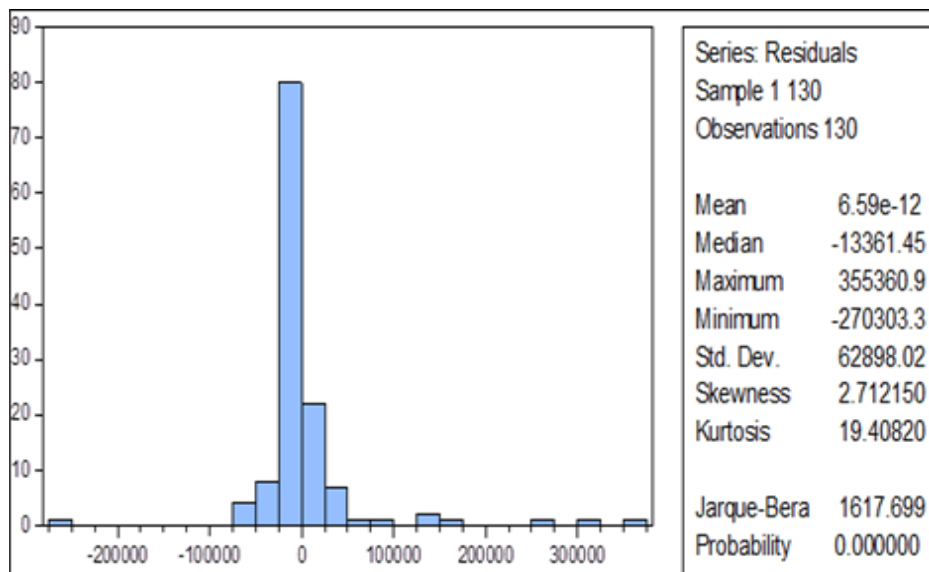


Figura 4.5.1 Testi i Normalitetit për Regresin Linear, viti 2019

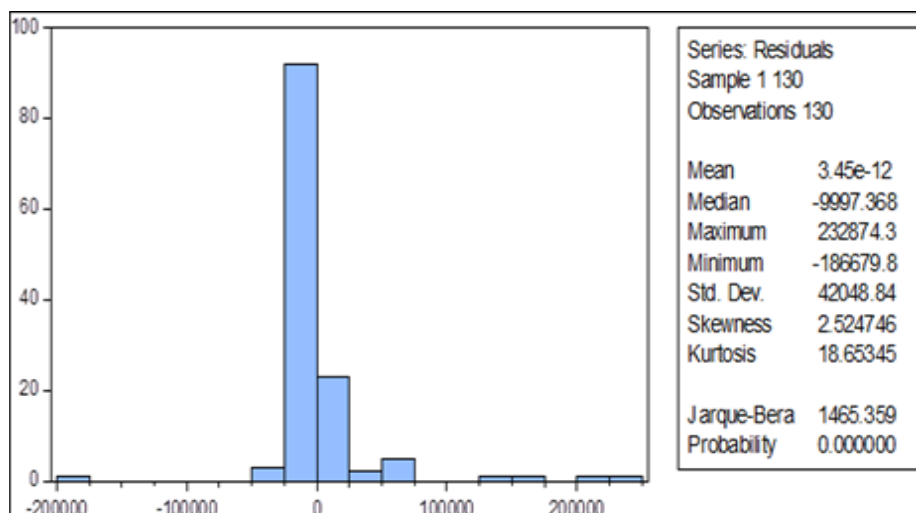


Figura 4.5.2 Testi i Normalitetit për Regresin Linear, viti 2020

Sikurse evidentohet nga Fig.1 dhe 2, regreset lineare mes fitimeve dhe sasive të shitura përkatësisht të viteve 2019 dhe 2020 kanë problemin e jolinearitetit. Ky problem evidentohet nga paraqitja grafike si edhe nga vlerat e probabilitetit faktik për çdo vit i cili merr probabilitet 0.000, duke qenë më i vogël se probabiliteti kritik prej 5%. Në këtë mënyrë ne nuk pranojmë hipotezën bazë e cila shpreh inekzistencën e këtij problemi.

Tabela 4.5.1 Testi i korrelacionit serial (LM), viti 2019

Breusch-Godfrey Serial Correlation LM Test (2019)

| | | | |
|---------------|----------|---------------------|--------|
| F-statistic | 11.02827 | Prob. F(2,126) | 0.0000 |
| Obs*R-squared | 19.36659 | Prob. Chi-Square(2) | 0.0001 |

Test Equation:
 Dependent Variable:RESID
 Method: Least Squares
 Date: 02/10/21 Time: 14:59
 Sample: 1 130
 Included observations: 130
 Presample missing value lagged residuals set to zero.

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------------|-------------|------------|-------------|--------|
| C | 1438.918 | 5599.924 | 0.256953 | 0.7976 |
| SALES_UNIT2020 | -6.320355 | 9.617993 | -0.657139 | 0.5123 |
| RESID(-1) | 0.415617 | 0.089324 | 4.652913 | 0.0000 |
| RESID(-2) | -0.099451 | 0.088556 | -1.123037 | 0.2636 |

| | | | |
|--------------------|----------|-----------------------|----------|
| R-squared | 0.148974 | Mean dependent var | 6.59E-12 |
| Adjusted R-squared | 0.128711 | S.D. dependent var | 62898.02 |
| S.E. of regression | 58710.8 | Akaike info criterion | 24.82892 |
| Sum squared resid | 4.34E+11 | Schwarz criterion | 24.91715 |
| Log likelihood | -1609.88 | Hannan-Quinn criter. | 24.86477 |
| F-statistic | 7.35218 | Durbin-Watson stat | 1.970731 |
| Prob(F-statistic) | 0.00014 | | |

Tabela 4.5.2 Testi i korrelacionit serial (LM), viti 2020

Breusch-Godfrey Serial Correlation LM Test (2020)

| | | | |
|---------------|----------|---------------------|--------|
| F-statistic | 8.597132 | Prob. F(2,126) | 0.0003 |
| Obs*R-squared | 15.60994 | Prob. Chi-Square(2) | 0.0004 |

Test Equation:
 Dependent Variable:RESID
 Method: Least Squares
 Date: 02/10/21 Time: 15:01
 Sample: 1 130
 Included observations: 130
 Presample missing value lagged residuals set to zero.

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------------|-------------|------------|-------------|--------|
| C | 729.0175 | 3757.082 | 0.194038 | 0.8465 |
| SALES_UNIT2020 | -4.596872 | 8.537285 | -0.538447 | 0.5912 |
| RESID(-1) | 0.364978 | 0.089529 | 4.076639 | 0.0001 |
| RESID(-2) | -0.058543 | 0.088891 | -0.658592 | 0.5114 |

| | | | |
|--------------------|-----------|-----------------------|----------|
| R-squared | 0.120076 | Mean dependent var | 3.45E-12 |
| Adjusted R-squared | 0.099126 | S.D. dependent var | 42048.84 |
| S.E. of regression | 39910.4 | Akaike info criterion | 24.05695 |
| Sum squared resid | 2.01E+11 | Schwarz criterion | 24.14518 |
| Log likelihood | -1559.702 | Hannan-Quinn criter. | 24.09280 |
| F-statistic | 5.731422 | Durbin-Watson stat | 1.975683 |
| Prob(F-statistic) | 0.001041 | | |

Tabela 4.5.1 dhe 4.5.2 pasqyrojnë rezultatet e testit LM të korrelacionit serial, që përdoret për identifikimin e ekzistencës ose jo të problemit të autokorrelacionit, i cili në thelb shpreh lidhje me vonesë të një serie mbetjesh me veten e saj. Duke u bazuar tek shpjegueshmëria e modelit për çdo vrojtim (obs*R square) shohim se probabiliteti faktik është 0.000, pra më i vogël se 5%. Në këtë mënyrë pranohet hipoteza alternative e cila vërteton ekzistencën e autokorrelacionit.

Tabela 4.5.3 Testi i heteroskedasticitetit, viti 2019

Heteroskedasticity Test: Breusch-Pagan-Godfrey (2019)

| | | | |
|---------------------|----------|---------------------|--------|
| F-statistic | 25.22631 | Prob. F(1,128) | 0.0000 |
| Obs*R-squared | 21.40246 | Prob. Chi-Square(1) | 0.0000 |
| Scaled explained SS | 190.9758 | Prob. Chi-Square(1) | 0.0000 |

Test Equation:
 Dependent Variable:RESID^2
 Method: Least Squares
 Date: 02/10/21 Time: 15:00
 Sample: 1 130
 Included observations: 130

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------------|-------------|------------|-------------|--------|
| C | 1.03E+09 | 1.48E+09 | 0.699362 | 0.4856 |
| SALES_UNIT2020 | 12636064 | 2515851 | 5.02258 | 0.0000 |

| | | | |
|--------------------|-----------|-----------------------|----------|
| R-squared | 0.164634 | Mean dependent var | 3.93E+09 |
| Adjusted R-squared | 0.158108 | S.D. dependent var | 1.69E+10 |
| S.E. of regression | 1.55E+10 | Akaike info criterion | 49.78319 |
| Sum squared resid | 3.08E+22 | Schwarz criterion | 49.82731 |
| Log likelihood | -3233.908 | Hannan-Quinn criter. | 49.80112 |
| F-statistic | 25.22631 | Durbin-Watson stat | 1.50046 |
| Prob(F-statistic) | 0.000002 | | |

Tabela 4.5.4 Testi i heteroskedasticitetit, viti 2020

Heteroskedasticity Test: Breusch-Pagan-Godfrey (2020)

| | | | |
|---------------------|----------|---------------------|--------|
| F-statistic | 24.91218 | Prob. F(1,128) | 0.0000 |
| Obs*R-squared | 21.17937 | Prob. Chi-Square(1) | 0.0000 |
| Scaled explained SS | 181.2366 | Prob. Chi-Square(1) | 0.0000 |

Test Equation:
 Dependent Variable:RESID^2
 Method: Least Squares
 Date: 02/10/21 Time: 15:01
 Sample: 1 130
 Included observations: 130

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------------|-------------|------------|-------------|--------|
| C | 6.04E+08 | 6.39E+08 | 0.944908 | 0.3465 |
| SALES_UNIT2020 | 7194458 | 1441426 | 4.99121 | 0.0000 |

| | | | |
|--------------------|----------|-----------------------|----------|
| R-squared | 0.162918 | Mean dependent var | 1.75E+09 |
| Adjusted R-squared | 0.156379 | S.D. dependent var | 7.40E+09 |
| S.E. of regression | 6.80E+09 | Akaike info criterion | 48.13265 |
| Sum squared resid | 5.91E+21 | Schwarz criterion | 48.17676 |
| Log likelihood | -3126.62 | Hannan-Quinn criter. | 48.15057 |
| F-statistic | 24.91218 | Durbin-Watson stat | 1.657908 |
| Prob(F-statistic) | 0.000002 | | |

Tabela 4.5.3 dhe 4.5.4 paraqet testin Breuch-Pagan-Godfrey që përdoret për identifikimin e problemit të heteroskedasticitetit, i cili në thelb shpreh gjenerimin e gabimeve të ndryshme për vrojtme të ndryshme. Sipas këtij testi probabiliteti i statistikës hi-katror që bazohet në shpjegueshmërinë e modelit për çdo vrojtje është 0. Meqë ky probabiliteti është më i vogël se probabiliteti 5% vërtetohet ekzistenca e heteroskedasticiteti përkatësisht për të dyja modelet e krijuara për 2019 dhe 2020.

Tabela 4.5.5 Modeli i Regresit Kuantil për marrëdhënien Shitje-Fitime, viti 2019

Dependent Variable: PROFIT2019
 Method: Quantile Regression (Median)
 Date: 02/09/21 Time: 21:01
 Sample: 1 130
 Included observations: 130
 Huber Sandwich Standard Errors & Covariance
 Sparsity method: Kernel (Epanechnikov) using residuals
 Bandwidth method: Hall-Sheather, bw=0.19179
 Estimation successfully identifies unique optimal solution

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|------------------------|-------------|--------------------|-------------|----------|
| C | 12909.71 | 1769.685 | 7.294921 | 0.0000 |
| SALES_UNIT_2019 | 143.1429 | 6.387266 | 22.41067 | 0.0000 |
| Pseudo R-squared | 0.390078 | Mean dependent var | | 55550.89 |
| Adjusted R-squared | 0.385313 | S.D. dependent var | | 91839.30 |
| S.E. of regression | 64835.38 | Objective | | 1768740. |
| Quantile dependent var | 22070.00 | Restr. objective | | 2899943. |
| Sparsity | 40509.30 | Quasi-LR statistic | | 223.3963 |
| Prob(Quasi-LR stat) | 0.000000 | | | |

Tabela 4.5.6 Modeli i Regresit Kuantil për marrëdhënien Shitje-Fitime, viti 2020

Dependent Variable: PROFIT_2020
 Method: Quantile Regression (Median)
 Date: 02/09/21 Time: 21:14
 Sample: 1 130
 Included observations: 130
 Huber Sandwich Standard Errors & Covariance
 Sparsity method: Kernel (Epanechnikov) using residuals
 Bandwidth method: Hall-Sheather, bw=0.19179
 Estimation successfully identifies unique optimal solution

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|------------------------|-------------|--------------------|-------------|----------|
| C | 7075.527 | 1074.773 | 6.583273 | 0.0000 |
| SALES_UNIT2020 | 141.0085 | 9.263846 | 15.22138 | 0.0000 |
| Pseudo R-squared | 0.406264 | Mean dependent var | | 36395.49 |
| Adjusted R-squared | 0.401625 | S.D. dependent var | | 63516.50 |
| S.E. of regression | 44147.96 | Objective | | 1179896. |
| Quantile dependent var | 12044.10 | Restr. objective | | 1987239. |
| Sparsity | 24475.46 | Quasi-LR statistic | | 263.8865 |
| Prob(Quasi-LR stat) | 0.000000 | | | |

Pra, meqë modeli i regresit ka problemet e autokorrelacionit, heteroskedasticitetit dhe jolinearititetit, ne ndërtuam modelin e regresit kuantil, përkatësisht për dy vitet 2019 dhe 2020. Rezultatet e gjenerimit të regresit kuantil për të dy vitet pasqyrohen tek Tabela 4.5.5 dhe 4.5.6. Ndryshorja e varur në të dy modelet e krijuara është fitimi,

kurse ndryshorja e pavarur është sasia e shitur për çdo produkt kozmetik. Nëse i referohemi vitit 2019 shohim se kur njësitë e shitura pësojnë një rritje prej 1% në vlerën e medianës, fitimi rritet me 143,14%. Pra, ekziston një marrëdhënie e fortë pozitive mes ndryshores së varur (fitimit) dhe ndryshores së pavarur (sasi të shitura).

Ndryshorja e pavarur ka një rëndësi të lartë statistikore tek fitimi sepse probabiliteti faktik është 0.000, pra më i vogël se probabiliteti kritik prej 0.005. Shpjegueshmëria e modelit (R-squared) është 39%, duke përbërë kështu një nivel të konsiderueshëm por jo optimal shpjegueshmërie të variancës së fitimit prej variancës së sasive të shitura. Kjo ndodh për shkak të ekzistencës së faktorëve të tjerë shpjegues të fitimit sikurse janë: preferenca e konsumatorit, ndryshimet e tregut, ndryshimet teknologjike etj.

Gabimi standard i koeficientit që tregon rëndësinë e njësive të shitura tek fitimi është 6%. Varianca e gjithë modelit të krijuar është e konsiderueshme sepse edhe shpjegueshmëria e gjithë modelit të krijuar nuk është më optimale.

E njëjta logjikë ndiqet për të shpjeguar edhe modelin e krijuar për 2020. I vetmi ndryshim që haset tek 2020 është fakti se kur vlera e medianës së njësive të shitura në 2020 rritet me 1% , fitimet e përfuara do të pësojnë një ulje duke arritur nivelin 141%. Kjo tendencë e ndryshme ndodhi për shkak të situatës pandemike e cila ndikoi në uljen e kërkesës konsumatore të vajzave për blerjen e produkteve kozmetike për shkak të mbajtjes së maskave .

Synimi i këtij studimi është që të reduktojë gabimin standard (devijimin) e koeficientit që përcakton ndikimin e njësive të shitura tek fitimet përkatësisht për të dyja vitet. Për të realizuar këtë objektiv është zbatuar procesi i testimit të kuantileve. Për këtë model ne morëm në konsideratë 10 kuantile dhe pamë shpërndarjen e gabimeve për çdo kuantil përgjatë 2 viteve të marra në studim.

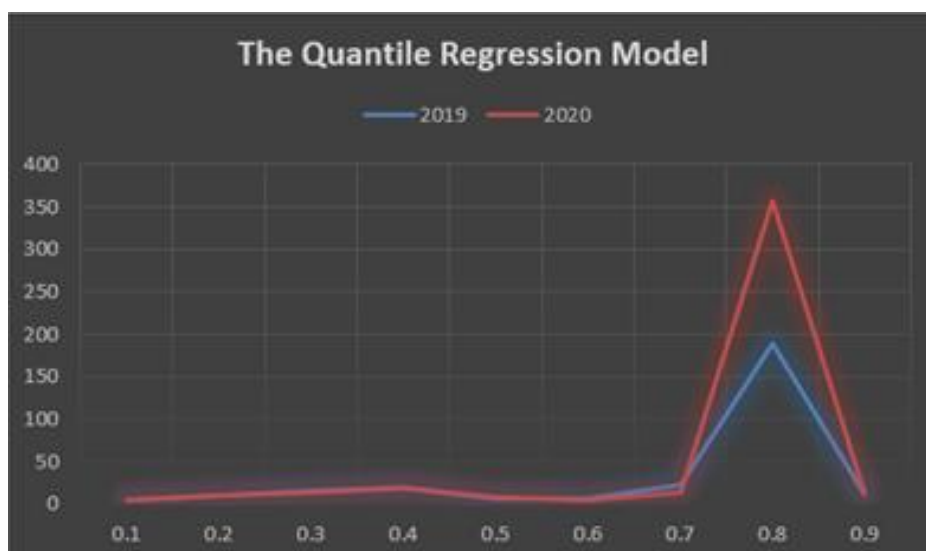


Figura 4.5.3 Shpërndarja e devijimeve standarde të Modeleve Kuantile për 2019 dhe 2020

Figura 4.5.3 paraqet shpërndarjen e devijimit standard (gabimit) për kuantilet e marra në shqyrtim. Për 10 kuantilet e para shpërndarja e gabimeve ka të njëjtën tendencë për të dyja vitet. Megjithatë vërrehet se piku më i lartë i gabimit për vitin 2020 është më i

lartë krahasuar me pikun e vitit 2019, për shkak të variancës më të madhe që ekziston në modelin e 2020.

Për të optimizuar modelin si fillim kemi implementuar përafrimin Edgeworth për modelin e regresit kuantil përkatësisht për të dy vitet. Tabela 4.5.7 dhe 4.5.8 pasqyrojnë përafrimin Edgeworth për rastin (IID) të modelit të regresit kuantil për të dy vitet e marra në studim. Rezultati rrjedhës i tyre është ndjeshëm i përmirësuar për shkak të reduktimit të konsiderueshëm të gabimit standard (devijimit). Konkretisht, gabimi pas implementimit të përafrimit Edgeworth në modelin e 2019 është 3.3%, duke pësuar kështu një rënie prej 3%. Gabimi i modelit të 2020 pas zbatimit të përafrimit Edgeworth është 2.4%, duke siguruar një rënie prej 6.9%, në nivelin e gabimit standard.

Tabela 4.5.7 Përafrimi Edgeworth për Modelin e Regresit Kuantil për vitin 2019

Dependent Variable: PROFIT2019
 Method: Quantile Regression (Median)
 Date: 02/09/21 Time: 21:03
 Sample: 1 130
 Included observations: 130
 Ordinary (IID) Standard Errors & Covariance
 Sparsity method: Kernel (Epanechnikov) using residuals
 Bandwidth method: Hall-Sheather, bw=0.19179
 Estimation successfully identifies unique optimal solution

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|------------------------|-------------|--------------------|-------------|----------|
| C | 12909.71 | 1950.852 | 6.617473 | 0.0000 |
| SALES_UNIT_2019 | 143.1429 | 3.321782 | 43.09220 | 0.0000 |
| Pseudo R-squared | 0.390078 | Mean dependent var | | 55550.89 |
| Adjusted R-squared | 0.385313 | S.D. dependent var | | 91839.30 |
| S.E. of regression | 64835.38 | Objective | | 1768740. |
| Quantile dependent var | 22070.00 | Restr. objective | | 2899943. |
| Sparsity | 40968.23 | Quasi-LR statistic | | 220.8938 |
| Prob(Quasi-LR stat) | 0.000000 | | | |

Tabela 4.5.8 Përafrimi Edgeworth për Modelin e Regresit Kuantil për vitin 2020

Dependent Variable: PROFIT_2020
 Method: Quantile Regression (Median)
 Date: 02/09/21 Time: 21:18
 Sample: 1 130
 Included observations: 130
 Ordinary (IID) Standard Errors & Covariance
 Sparsity method: Siddiqui using residuals
 Bandwidth method: Hall-Sheather, bw=0.19179
 Estimation successfully identifies unique optimal solution

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|------------------------|-------------|--------------------|-------------|----------|
| C | 7075.527 | 1042.195 | 6.789060 | 0.0000 |
| SALES_UNIT2020 | 141.0085 | 2.350360 | 59.99445 | 0.0000 |
| Pseudo R-squared | 0.406264 | Mean dependent var | | 36395.49 |
| Adjusted R-squared | 0.401625 | S.D. dependent var | | 63516.50 |
| S.E. of regression | 44147.96 | Objective | | 1179896. |
| Quantile dependent var | 12044.10 | Restr. objective | | 1987239. |
| Sparsity | 22166.23 | Quasi-LR statistic | | 291.3776 |
| Prob(Quasi-LR stat) | 0.000000 | | | |

Nëse marrim në konsideratë Fig. 4.5.4 shohim se rezultatet e përafrimit Edgeworth të shpërndarjes së gabimeve për çdo kuantil të konsideruar vizualizon një tendencë përafruese pozitive drejt sheshtësimit të pikut të gabimit.

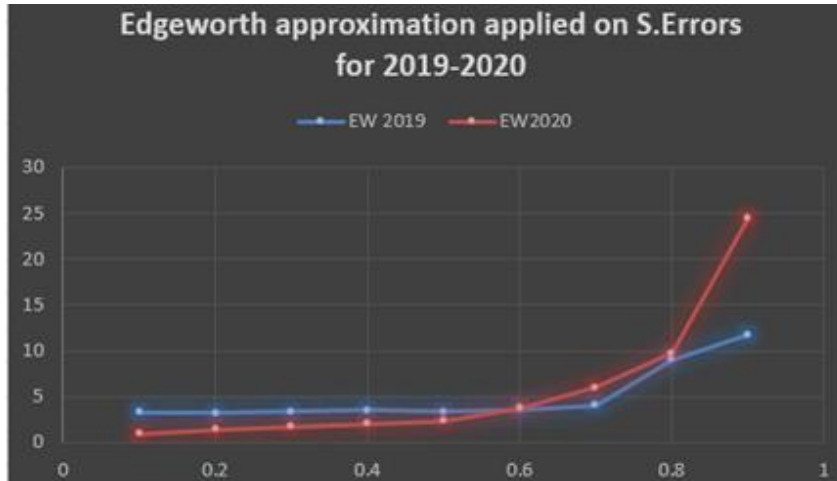


Figura 4.5.4 Përafrimi Edgeworth i shpërndarjes së gabimeve (devijimi standard)

Për të parë rezultatet e një metode tjetër në termat e reduktimit të gabimeve standarde, kemi zbatuar metodën bootstrap për mbetjet. Rezultatet e kësaj metode paraqiten tek Tabela 4.5.9 dhe 4.5.10.

Në 2019 gabimi standard ka arritur nivelin prej 3.33% duke siguruar kështu një rënie prej 3.05%, kurse në vitin 2020 gabimi standard është 2.91% duke siguruar një rënie prej 6.35%. Pra, pas zbatimit të versionit bootstrap të mbetjeve për 10000 rizgjedhje tek të dy modelet e krijuara, evidentohet rritja e saktësisë së marrëdhënies së marrë në studim mes njërive të shitura dhe fitimeve të përfutuara për shkak të reduktimit të gabimit.

Tabela 4.5.9 Versioni Bootstrap i Modelit të Regresit Kuantil për vitin 2019

Dependent Variable: PROFIT2019
Method: Quantile Regression (Median)
Date: 02/09/21 Time: 21:05
Sample: 1 130
Included observations: 130
Bootstrap Standard Errors & Covariance
Bootstrap method: Residual, reps=10000, rng=kn, seed=633957423
Sparsity method: Siddiqui using residuals
Bandwidth method: Hall-Sheather, bw=0.19179
Initial Values: C(1)=0.00000, C(2)=0.00000
Estimation successfully identifies unique optimal solution

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|------------------------|-------------|--------------------|-------------|----------|
| C | 12909.71 | 1619.185 | 7.972970 | 0.0000 |
| SALES_UNIT_2019 | 143.1429 | 3.329511 | 42.99217 | 0.0000 |
| Pseudo R-squared | 0.390078 | Mean dependent var | | 55550.89 |
| Adjusted R-squared | 0.385313 | S.D. dependent var | | 91839.30 |
| S.E. of regression | 64835.38 | Objective | | 1768740. |
| Quantile dependent var | 22070.00 | Restr. objective | | 2899943. |
| Sparsity | 34588.92 | Quasi-LR statistic | | 261.6337 |
| Prob(Quasi-LR stat) | 0.000000 | | | |

Tabela 4.5.10 Versioni Bootstrap i Modelit të Regresit Kuantil për vitin 2020

Dependent Variable: PROFIT_2020
 Method: Quantile Regression (Median)
 Date: 02/09/21 Time: 21:36
 Sample: 1 130
 Included observations: 130
 Bootstrap Standard Errors & Covariance
 Bootstrap method: Residual, reps=10000, rng=kn, seed=911414654
 Sparsity method: Kernel (Epanechnikov) using residuals
 Bandwidth method: Hall-Sheather, bw=0.19179
 Initial Values: C(1)=0.00000, C(2)=0.00000
 Estimation successfully identifies unique optimal solution

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|------------------------|-------------|--------------------|-------------|----------|
| C | 7075.527 | 1054.578 | 6.709346 | 0.0000 |
| SALES_UNIT2020 | 141.0085 | 2.914924 | 48.37469 | 0.0000 |
| Pseudo R-squared | 0.406264 | Mean dependent var | | 36395.49 |
| Adjusted R-squared | 0.401625 | S.D. dependent var | | 63516.50 |
| S.E. of regression | 44147.96 | Objective | | 1179896. |
| Quantile dependent var | 12044.10 | Restr. objective | | 1987239. |
| Sparsity | 24498.41 | Quasi-LR statistic | | 263.6393 |
| Prob(Quasi-LR stat) | 0.000000 | | | |

Në të dyja modelet e Regresit Kuantil të krijuara, përkatësisht për vitin 2019 dhe 2020, u aplikua versioni Bootstrap i mbetjeve (Residual Bootstrap) për 10,000 rizgjedhje. Rezultati pas implementimit të këtij versioni ishte reduktimi i devijimeve standarde (gabimit) duke rritur kështu saktësinë e marrëdhënies që po studiojmë mes njësive të shitura dhe fitimeve të përfuara. Ky rezultat pozitiv detajohet dhe në tendencën rënëse të paraqitjes në figurën 4.5.5.

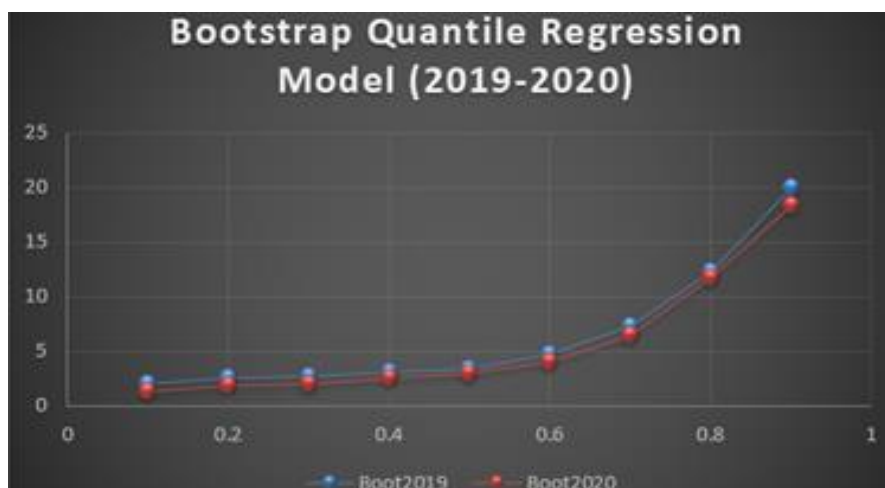


Figura 4.5.5 Versioni Bootstrap i shpërndarjes së gabimeve për çdo kuantil

Për të finalizuar studimin e optimizimit të modeleve të krijuara kemi aplikuar përafrimin Edgeworth për versionin Bootstrap të regresit kuantil për të dy vitet. Nëse metoda e rizgjedhjes bootstrap reduktoi termin e gabimit duke rritur saktësinë e parashikimit të trendit të fitimeve të gjeneruara, implementimi i përafrimit Edgeworth

për versionin bootstrap do të optimizojë edhe më tepër modelin, duke rritur nivelin e saktësisë. Rezultatet e kësaj metode konkretizohet tek Tabela 4.5.11 dhe 4.5.12.

Tabela 4.5.11 Përafrimi Edgeworth i versionit Bootstrap të Regresit Kuantil, 2019

Dependent Variable: PROFIT2019
 Method: Quantile Regression (Median)
 Date: 02/09/21 Time: 21:06
 Sample: 1 130
 Included observations: 130
 Ordinary (IID) Standard Errors & Covariance
 Sparsity method: Siddiqui using residuals
 Bandwidth method: Hall-Sheather, bw=0.19179
 Initial Values: C(1)=0.00000, C(2)=0.00000
 Estimation successfully identifies unique optimal solution

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|------------------------|-------------|--------------------|-------------|----------|
| C | 12909.71 | 1647.078 | 7.837947 | 0.0000 |
| SALES_UNIT_2019 | 143.1429 | 2.804536 | 51.03979 | 0.0000 |
| Pseudo R-squared | 0.390078 | Mean dependent var | | 55550.89 |
| Adjusted R-squared | 0.385313 | S.D. dependent var | | 91839.30 |
| S.E. of regression | 64835.38 | Objective | | 1768740. |
| Quantile dependent var | 22070.00 | Restr. objective | | 2899943. |
| Sparsity | 34588.92 | Quasi-LR statistic | | 261.6337 |
| Prob(Quasi-LR stat) | 0.000000 | | | |

Tabela 4.5.12 Përafrimi Edgeworth i versionit Bootstrap të Regresit Kuantil, 2020

Dependent Variable: PROFIT_2020
 Method: Quantile Regression (Median)
 Date: 02/09/21 Time: 21:34
 Sample: 1 130
 Included observations: 130
 Ordinary (IID) Standard Errors & Covariance
 Sparsity method: Siddiqui using residuals
 Bandwidth method: Hall-Sheather, bw=0.19179
 Initial Values: C(1)=0.00000, C(2)=0.00000
 Estimation successfully identifies unique optimal solution

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|------------------------|-------------|--------------------|-------------|----------|
| C | 7075.527 | 1042.195 | 6.789060 | 0.0000 |
| SALES_UNIT2020 | 141.0085 | 2.350360 | 59.99445 | 0.0000 |
| Pseudo R-squared | 0.406264 | Mean dependent var | | 36395.49 |
| Adjusted R-squared | 0.401625 | S.D. dependent var | | 63516.50 |
| S.E. of regression | 44147.96 | Objective | | 1179896. |
| Quantile dependent var | 12044.10 | Restr. objective | | 1987239. |
| Sparsity | 22166.23 | Quasi-LR statistic | | 291.3776 |
| Prob(Quasi-LR stat) | 0.000000 | | | |

Kjo metodë ka dhënë rezultate mjaft pozitive në të dy vitet e marra në studim, ku për 2019 niveli i gabimit standard ka njour rënie prej 3.6%, kurse për 2020 rënia e gabimit standard ka njohur masën 6.91%.

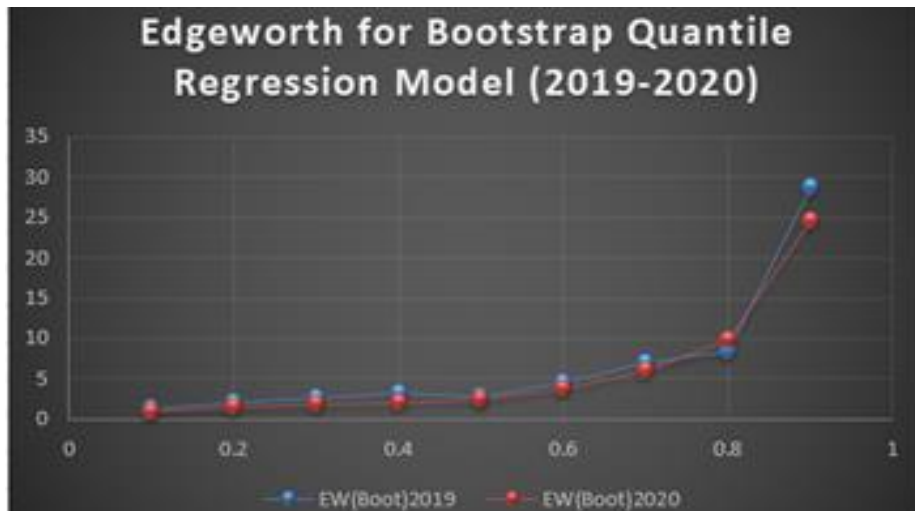


Figura 4.5.6 Përafrimi Edgeworth i versionit Bootstrap të Regresit Kuantil (2019-2020)

Prirja e shpërndarjes së gabimeve për secilin kuantil të zgjedhur pas aplikimit të përafrimit Edgeworth për versionin bootstrap ka tendencë të dukshme rënëse. Ky rezultat pozitiv është i evidentueshëm në të dy vitet e marra në studim që konkretizohen në Fig. 4.5.6.

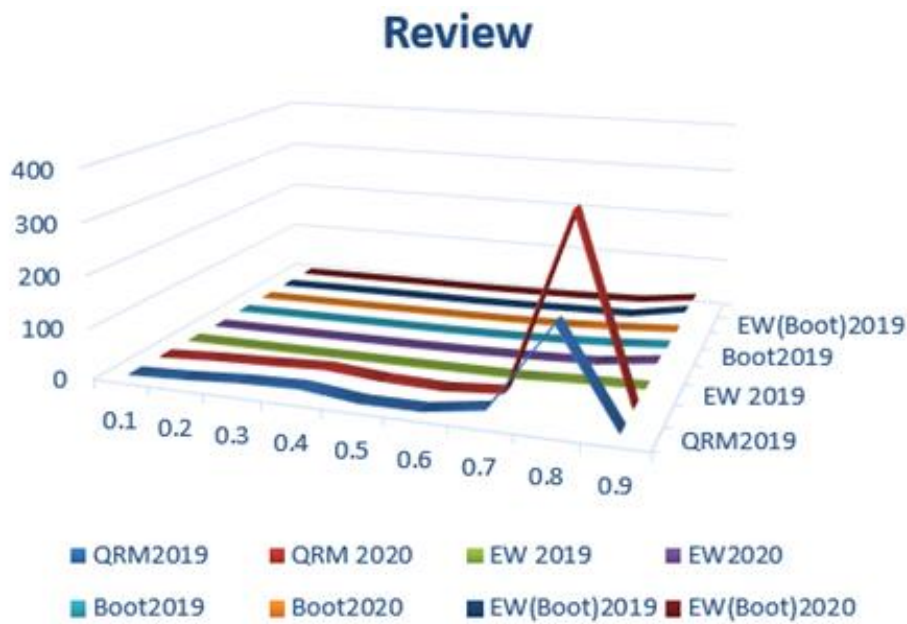


Figura 4.5.7 Përmbledhje rezultatesh

Prirja e shpërndarjes së gabimeve për secilin kuantil të zgjedhur pas aplikimit të përafrimit Edgeworth për versionin bootstrap ka tendencë të dukshme rënëse. Ky rezultat pozitiv është i evidentueshëm në të dy vitet e marra në studim që konkretizohen në Fig. 4.5.7.

KAPITULLI V

ZBATIMI I ZBËRTHIMIT EDGEWORTH DHE VERSIONIT BOOTSTRAP NË SERITË KOHORE

5.1 Kontributi i përafrimit Edgeworth dhe Metodave të rizgjedhjes Bootstrap në prognozimin e serive kohore

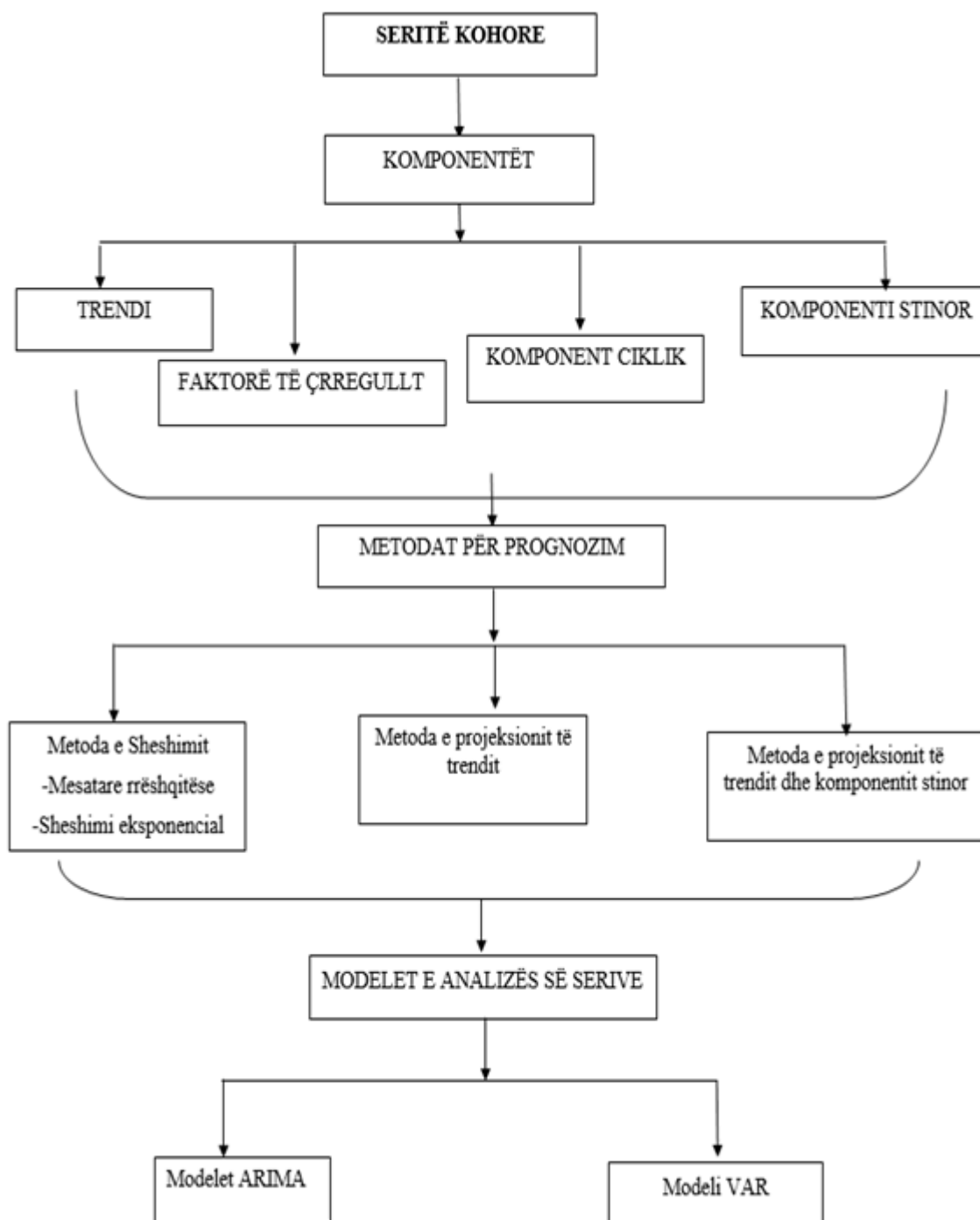


Figura 5.1.1 Analizë e Serive Kohore

Seritë ndihmojnë në prognozimin e ardhshëm sasior ose cilësor të të dhënave të marra në studim. Metodatat kryesore prognozuese në termat e serive kohore (shih [2], [3]), janë:

- a) Metoda e sheshimit, e cila detajohet në Mesatarja rrëshqitëse dhe Metoda e sheshimit eksponencial
- b) Metoda e projeksionit të trendit
- c) Metoda e projeksionit të trendit dhe komponentit stinor

Komponentët e një serie kohore janë:

- 1) Trendi, i cili zhvendos serinë kohore si rezultat i faktorëve që ndikojnë në të përgjatë një periudhe të gjatë kohore
- 2) Komponenti Ciklik, i cili nënkupton vargun e pikave jashtë vijës së trendit
- 3) Komponenti stinor, i cili ndikon serinë në periudha më të shkurtra se një vit
- 4) Faktorët e çrregullt, të cilët ndikojnë në mënyrë rastësore serinë

Dy metodatat e analizës së Serive Kohore janë:

- a) Modeli ARIMA (Mesatare Rrëshqitëse Autoregresive e Integruar), i cili analizon strukturat dinamike të të dhënave të serive kohore të marra në studim
- b) Modeli VAR (Vektori Autoregresiv), i cili është një algoritëm parashikues që zbatohet kur dy ose më tepër seri influencojnë njëra-tjetrën.

Në këtë studim konkret, fokusi do të jetë Modeli VAR.

5.1.1 Vektori i Autoregresit (VAR)

Modeli i Vektorit të Autoregresit (VAR) zbatohet gjerësisht në parashikimin dhe analizën e sistemeve të serive kohore të ndërlidhura të cilat mund të jenë subjekt i ndikimit dinamik të influencave të rastësishme. Forma e reduktuar e Vektorit të Autoregresit nuk kryen modelim strukturor të sistemit të serive kohore duke trajtuar çdo ndryshore endogjene të sistemit si një funksion i vlerave të vonuara të të gjitha ndryshoreve endogjene të përfshira në model, ([43], [54]).

Trajtimet e bazave të analizës VAR dhe tiparet e rëndësishme të këtij modeli, përshkruhen në [67].

Ne mund të shprehim modelin e serive stacionare k-dimensionale, VAR (p) si:

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + Cx_t + \varepsilon_t \quad (5.1.1)$$

ku $y_t = (y_{1t}, y_{2t}, \dots, y_{kt})'$ është vektori $k \times 1$ -përmasor i ndryshoreve endogjene, $x_t = (x_{1t}, x_{2t}, \dots, x_{dt})'$ është vektori $d \times 1$ -përmasor i ndryshoreve ekzogjene, A_1, \dots, A_p janë matricat me përmasa $k \times k$ që kontribuojnë në vlerësimin e koeficientëve të vonesës, C është matrica me përmasa $k \times d$ e cila përdoret për vlerësimin e koeficientëve në rastin e ndryshoreve ekzogjene, $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{kt})'$ është zhurma e bardhë e procesit inovativ me përmasë $k \times 1$, e cila mund të shprehet:

$$E(\varepsilon_t) = 0, E(\varepsilon_t \varepsilon_t') = \Sigma_\varepsilon, \text{ dhe } E(\varepsilon_t \varepsilon_s') = 0, \text{ for } t \neq s.$$

Ky ekuacion vërteton se vektori i inovacionit korrelohet me rangun e plotë të matricës, por nuk shfaq asnjë tendencë lidhshmërie me drejtimin dhe vonesat e inovacionit dhe ndryshoreve të anës së djathtë.

Supozojmë se vektori $(pk+d) \times 1$ -përmasor ,

$$Z_t = (y_{t-1}', \dots, y_{t-p}', x_t')$$

pasqyron gjithë regresorët e periudhës kohore t në modelin e Vektorit të Autoregresit. Gjithashtu vëzhgimet për këtë model shprehen si:

$$Y = BZ + E$$

ku $Y = (y_1, y_2, \dots, y_T)$ dhe $E = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T)$ janë matricat me përmasa $k \times T$ të ndryshoreve endogjenë dhe inovacioneve, dhe,

$$B = (A_1, A_2, \dots, A_p, C)$$

$$Z = (Z_1, Z_2, \dots, Z_T)$$

janë përkatësisht, matrica me përmasë $k(pk+d)$ e sistemit të koeficientëve dhe matrica me përmasë $(pk+d) \times T$ e të dhënave të regresorëve.

Në formë të përmbledhur do të kishim:

$$\tilde{y} = (Z' \otimes I_k) \beta + \tilde{\varepsilon}$$

ku $\tilde{y} = \text{vec}(Y)$, $\beta = \text{vec}(B)$, dhe

$$\tilde{\varepsilon} = \text{vec}(E), \text{ ku } E(\tilde{\varepsilon}\tilde{\varepsilon}') = (I_k \otimes \Sigma_\varepsilon).$$

Për aq kohë sa vlerat e vonuara të inovacioneve dhe vlerave endogjene supozohen të jenë të pakorreluara me inovacionet dhe ndryshoret ekzogjene të vonuara, kushtet standarde janë të plotësuara dhe paraqitja grafike e Metodës së Zakonshme të Katrorëve më të Vegjël jep një vlerësim të qëndrueshëm.

Efiçenca e Metodës së Zakonshme të Katrorëve më të Vegjël e cila përgjithësisht është ekuivalente me Metodën e Përgjithshme të Katrorëve më të Vegjël, konsiston në faktin se të gjitha ekuacionet e sistemit kanë regresorë identikë.

Nëpërmjet EViews-it, Metoda e Zakonshme e Katrorëve më të vegjël aplikohet tek çdo ekuacion për të vlerësuar Vektorin Standard të Autoregresit në modelin e marrë në studim.

Implementimi i Metodës së Zakonshme të Katrorëve më të Vegjël do të gjeneronte vlerësimin:

$$\hat{\beta} = ((ZZ')^{-1} Z \otimes I_k) \tilde{y}$$

i cili ka një matricë kovariance:

$$V(\hat{\beta}) = (ZZ')^{-1} \otimes \Sigma_\varepsilon.$$

Për të përfutur matricën e kovariancës nevojitet vlerësimi, i cili realizohet nëpërmjet diferencës së vlerësuesit të korigjuar të mbetjeve për momente të caktuara:

$$\hat{\Sigma}_\varepsilon = \frac{\hat{E}\hat{E}'}{T - (pk + d)}$$

ku $\hat{E} = Y - \hat{B}Z$ për $\hat{\beta} = \text{vec}(\hat{B})$.

5.1.2 Testet e Vonesës

Testi i përjashtimit të vonesës

Ky test kryen një vlerësim të detajuar të çdo vonesë të përjashtuar të modelit të Vektorit të Autoregresit. Statistika (Wald) jep rezultate të pavarura dhe të përbashkëta për rëndësinë e përbashkët të të gjitha ndryshoreve endogjene të secilit ekuacion tek çdo vonesë e përfshirë në model.

Kriteret e gjatësisë së vonesës

Gjeneron kritere të ndryshme të zgjedhjes së rendit të vonesës për një VAR të pakufizuar. Është i nevojshëm specifikimi i rendit maksimal të vonesave për realizimin e testit. Rezultatet e gjeneruara tregojnë kritere të ndryshme informacioni për të gjitha vonesat për maksimumin e përcaktuar. (Vonesa fillon me 1 nëse modeli VAR është i lirë nga ndryshoret ekzogjene dhe fillon me 0 për alternativën e anasjelltë) Kriteret shpjegohen më hollësisht në [66]. Testi LR fillon me vonesën maksimale dhe përdor statistikën për të testuar hipotezën që shprehet për vlerën zakonisht zero të koeficientëve në vonesën l:

$$LR = (T - m) \{ \log |\Sigma_{\varepsilon, l-1}| - \log |\Sigma_{\varepsilon, l}| \} \sim \chi^2(k^2)$$

ku m është numri i parametrave. Gjithashtu mund të konsiderohet shembulli i modifikuar i Sims' [77], i cili përdor $(T-m)$ në vend të T . Rezultatet statistikore të LR së modifikuar duhet të krahasohen me vlerën kritike prej 5%, e cila fillon me maksimumin e vonesës dhe redukton vonesat për çdo periudhë kohore derisa arrin në një ndërprerje të caktuar. Nëse testi nuk identifikon ndërprerje, minimumi i vonesave do të shënohet me një yll. Një përfundim i rëndësishëm lidhet me faktin se ndonëse testet individuale kanë madhësi 0,05, e gjithë madhësia e testit është e pamundur të jetë 5%, ([66], [43], [54]).

5.1.3 Testet e mbetjeve

Testi i Autokorrelacionit (LM)

Testi statistikor LM për identifikimin e korrelacionit serial të mbetjeve në një rend të caktuar, përdoret si Testi i Autokorrelacionit LM në rastin e një modeli shumëpërmasor. Mënyra e llogaritjes së autokorrelacionit të rendit h nëpërmjet testit Breusch-Godfrey konsiston në përdorimin e një regresi ndihmës për mbetjet u_t në regresorët e anës së djathtë dhe vlerat fillestare të humbura të h tek u_{t-h} përafrohen drejt zeros. Alternativa e këtij testi LM është formula e LR [57]. Korrigjimi i rezultateve të testit LM mundësohet nga zberthimi korrektues Edgeworth për një vonesë të caktuar [30]. Testi statistikor LM ka shpërndarje asimptotike χ^2 me k^2 shkallë lirie, në rastin kur nuk ekziston korrelacion serial i rendit, si hipoteza bazë.

Një alternative tjetër e testit statistikor LM është Rao F -test, i cili llogaritet lehtësisht nëpërmjet paketës EViews [30]. Simulimet e Edgerton dhe Shukur sugjerojnë këtë

alternativë, si opsionin që performon më mirë se variantet e tjera që mund të jenë potencialisht të aplikueshme.

Testi White i Heteroskedasticitetit

Testi i regresit realizohet nëpërmjet kontrollit të domethënies së përbashkët të ndërlidhjeve përkatëse të mbetjeve me atë të regresorëve tek modeli i regresit të krijuar. Për modelin me terma jo të ndërlidhura, konsiderohet niveli dhe katrori i regresorëve origjinalë, ndërsa për modelin me terma të ndërlidhura, përfshihen të gjithë termat e tepërta të gjenerimeve të ndërlidhura si rezultat i regresorëve origjinalë të ekuacionit të testuar. Termi constant mund të përfshihet si regresor në një test regresioni.

Rezultatet fillestare përfshijnë rëndësinë e përbashkët të regresorëve me përjashtim të termit constant. Këto teste verifikojnë qendrueshmërinë e secilit element në mënyrë të pavarur në matricën e kovariancës. Hipoteza bazë paraqet supozimin se regresorët janë të ndryshueshëm dhe nuk kanë një domethënie të përbashkët, [37].

5.1.4 Përgjigjet e impulsit

Një ndryshim rastësor i ndryshores së i -të ndikon ndryshoren e i -të duke influencuar kështu të gjithë ndryshoret endogjene të përfshira në modelin e Vektorit të Autoregresit. Rezultati i impulsit përkon me ndikimin gjatë një periudhe të inovacioneve të vlerave aktuale dhe të ardhme të ndryshoreve të marra në studim, ([43], [54]).

Rezultati i përgjigjeve të impulsit është i qartë nëse inovacionet nuk janë të korreluara së bashku. Inovacioni i i -të është një ndryshim rastësor i ndryshores së i -të endogjene.

Inovacionet janë përgjithësisht të korreluara dhe nuk mund të shoqërohen me një ndryshore specifike. Për një interpretim të saktë të impulseve, është i nevojshëm konvertimi i inovacioneve për t 'i kthyer ato në të pakorreluara.

$$u_t = P\varepsilon_t \sim (0, D)$$

Ku D është matrica diagonale e kovariancës. EViews siguron opsione të ndryshme për të zgjedhur se cila P prodhon një matricë diagonale.

Në përgjigjet e impulsit, një ndër komandat më të rëndësishme e cila jep një paraqitje më të mirë të vlerësimit të modelit VAR është Analiza Asimptotike, e cila i referohet përafritimit Edgeworth.

5.2 Zbërthimi Edgeworth për seritë autoregressive stacionare

Teknikat e përdorura për përfitim të zbërthimeve Edgeworth për shpërndarjen e statistikave një-përmasore, të cilat i pamë të zgjeruara tek kapitulli I, mund të përgjithësohen në rastin shumë përmasor. Megjithatë, në rastin shumë përmasor, shënimet, simbolet dhe gjithë problemi bëhet më kompleks. Një rezultat themelor në zbërthimet Edgeworth shumë përmasore është ai i Bhattacharya dhe Ghosh, [6].

Tani, le të shohim rendin e zbërthimit Edgeworth për një seri autoregressive stacionare.

Le të jetë Y_t një proces autoregresiv stacionar,

$$Y_t = \sum_{i=1}^p \theta_i Y_{t-i} + \varepsilon_t,$$

ku ne supozojmë se:

(A.1) (ε_t) janë të pavarura dhe identikisht të shpërndara (i.i.d.) $\sim F_0, E\varepsilon_t = 0, E\varepsilon_t^2 = 1, E\varepsilon_t^{2(s+1)} < \infty$, për $s \geq 3$.

(A.2) $(\varepsilon_1, \varepsilon_1^2)$ përmbush kriterin e Cramer-it, ku për çdo $d > 0$, ekziston $\delta > 0$ sikurse: $\sup_{\|t\|>d} |E \exp(it'(\varepsilon_1, \varepsilon_1^2))| \leq \exp(-\delta)$.

(A.3) Rrënjët e $\sum_{j=0}^p \theta_j z^{p-j} = 0$ shtrihen brenda rrethit njësi. Pra, $\theta_0 = 1$.

Argumentat ([9], [47]) tregojnë që nëse përmbushen kushtet (A.1) dhe (A.3) dhe ε_1 kënaq kushtin e Cramer-it, shpërndarja e $n^{-1/2} \tilde{S}_n(\theta_n - \theta)$ lejon një zbërthim Edgeworth të rendit $o(n^{-(s-2)/2})$.

Metoda Monte Carlo

Monte Karlo është një nga metodat më të mira që mundëson përmirësimin e rezultateve të gjeneruara pas aplikimit të një sërë rizgjedhjesh. Monte Carlo ngjan me Metodën Bootstrap kur aplikohet në përpunimin e të dhënave reale. Të dhënat e përdorura në një studim konkret janë të kufizuara, dhe një metodë rizgjedhje si Monte Carlo do të rriste saktësinë e rezultateve duke reduktuar gabimet.

Përgjigjet e Impulsit lidhur me Gabimet Standarde dhe Intervalet e Besimit

Përafrimi asimptotik jep gjenerim më të saktë të intervaleve të besimit për një model VAR dhe kjo lloj shpërndarje ka një performancë optimale në zgjedhjet e kufizuara [66].

Gabimet standarde të përgjigjes së impulsit në vlerësimin e Intervaleve të Besimit lidhen me zbatimin e metodave Monte Carlo dhe Bootstrap në të dhëna të kufizuara të marra në studim. Synimi i implementimit të tyre është rritja e optimalizimit të modelit.

5.3 Modeli Bootstrap për modelin e Vektorit të Autoregresit (VAR)

Nëse sistemi i serive kohore mund të modifikohet si një Vektor Autoregresiv i rendit p , procesi i gjenerimit të të dhënave për këtë sistem shprehet si:

$$y_t = v + \sum_{j=1}^p A_j y_{t-j} + u_t \quad (5.3.1)$$

ku y_t dhe v janë vektorë me K elementë, A_1, \dots, A_p janë matrica me përmasë $K \times K$; dhe Σ_u është matrica e kovariancës me përmasë $K \times K$, e cila merret prej $u_t \sim N(0, \Sigma_u)$.

Supozimi i mëposhtëm ka të bëjë me funksionalitetin e objektit të synuar si: vlerësuesi, testet statistikore, etj.

$$q_T = q(y_1, \dots, y_T, v, A_1, \dots, A_p) \quad (5.3.2)$$

Shpërndarja $F_{q,T}$ e q_T mund të përafrohet nëpërmjet simulimeve të një numri të konsiderueshëm të dhënash, p.sh. B , të rezultateve të ndryshme si $(q_T^{*b}$ për $b=1, \dots, B$), ku “*” përdoret për të treguar derivatin e objektit pas simulimeve. Gjithashtu, rezultati i B , q_T^{*b} , nxirret nëpërmjet procesit të përpunimit të të dhënave të përfuara nga (5.3.1), për t’i dhënë B rezultatet e pavarura si y_t^{*b} të y_t dhe për çdo rezultat gjejmë q_T^{*b} përkatese duke përdorur (5.3.2) për të dhënat e gjeneruara. Shpërndarja e q_T është e njëjtë me shpërndarjen empirike të rezultateve të simuluar $\{y_t^{*b}\}_{b=1}^B$.

5.4 Përpunimi i të dhënave financiare nëpërmjet zbatimit të Metodave Edgeworth, Bootstrap dhe Monte Carlo

Në këtë çështje ne do të trajtojmë marrëdhënien mes Parave (Ekuivalentëve të Parave) dhe tre aktiviteve kryesore të Pasqyrës së Rrjedhjes së Parave si: Aktiviteti i Shfrytëzimit (operacional), Aktiviteti Financiar dhe Aktiviteti Investues, shih [28]. Të dhënat për katër seritë që do të analizojmë janë marrë në bazë tre mujore për 30 vitet e konsideruara, nga viti 1991 deri në vitin 2020. Informacioni u siguroi nga raportimi financiar në faqen zyrtare të një prej bizneseve me sukses global: Coca-Cola. Ky studim konsiston në identifikimin e metodave më të përshtatshme dhe të sakta për të parashikuar seritë kohore, por hapi fillestar i këtij studimi është studimi i stacionaritetit të serive [38].

Studimi i stacionaritetit të serive është një çështje thelbësore, e cila duhet të realizohet para analizës së mëtejshme të serive, sepse në të kundërt rezultatet e analizës si edhe parashikimit të serive do të ishin të pasakta. Në rastin tonë evidentojmë se pas zbatimit të Testit Unit Root, i cili identifikon stacionaritetin e serive, katër seritë e marra në studim nuk janë të qendrueshme (stacionare).

Cash&Cash-Equivalents

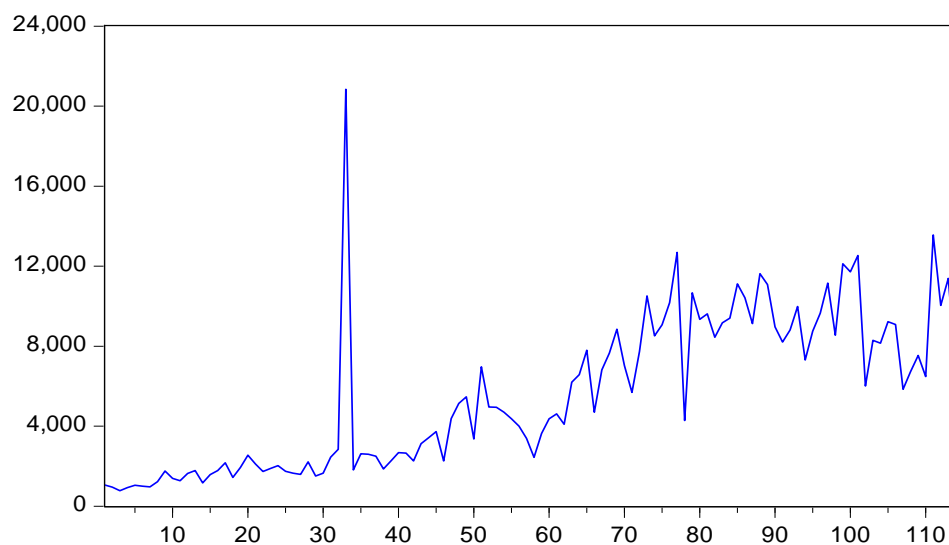


Figura 5.4.1 Jostacionariteti i serisë së Parave dhe Ekuivalentëve të Parave

Tabela 5.4.1 Testi Unit Root për serinë e Parave dhe Ekuivalentëve të Parave

Null Hypothesis: CASH_CASH_EQUIVALENTS has a unit root
 Exogenous: Constant
 Lag Length: 3 (Automatic - based on SIC, maxlag=12)

| | t-Statistic | Prob.* |
|----------------------------------------|-------------|--------|
| Augmented Dickey-Fuller test statistic | -1.694481 | 0.4312 |
| Test critical values: | | |
| 1% level | -3.490772 | |
| 5% level | -2.887909 | |
| 10% level | -2.580908 | |

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(CASH_CASH_EQUIVALENTS)
 Method: Least Squares
 Date: 06/26/21 Time: 13:58
 Sample (adjusted): 5 114
 Included observations: 110 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|------------------------------|-------------|-----------------------|-------------|----------|
| CASH_CASH_EQUIVALENTS(-1) | -0.123865 | 0.073099 | -1.694481 | 0.0931 |
| D(CASH_CASH_EQUIVALENTS(-1)) | -0.637346 | 0.108687 | -5.864047 | 0.0000 |
| D(CASH_CASH_EQUIVALENTS(-2)) | -0.392307 | 0.115073 | -3.409202 | 0.0009 |
| D(CASH_CASH_EQUIVALENTS(-3)) | -0.259047 | 0.096588 | -2.681984 | 0.0085 |
| C | 884.5453 | 473.5551 | 1.867882 | 0.0646 |
| R-squared | 0.383562 | Mean dependent var | | 53.29091 |
| Adjusted R-squared | 0.360079 | S.D. dependent var | | 3160.892 |
| S.E. of regression | 2528.558 | Akaike info criterion | | 18.55307 |
| Sum squared resid | 6.71E+08 | Schwarz criterion | | 18.67582 |
| Log likelihood | -1015.419 | Hannan-Quinn criter. | | 18.60286 |
| F-statistic | 16.33339 | Durbin-Watson stat | | 2.042067 |
| Prob(F-statistic) | 0.000000 | | | |

Financial Activities

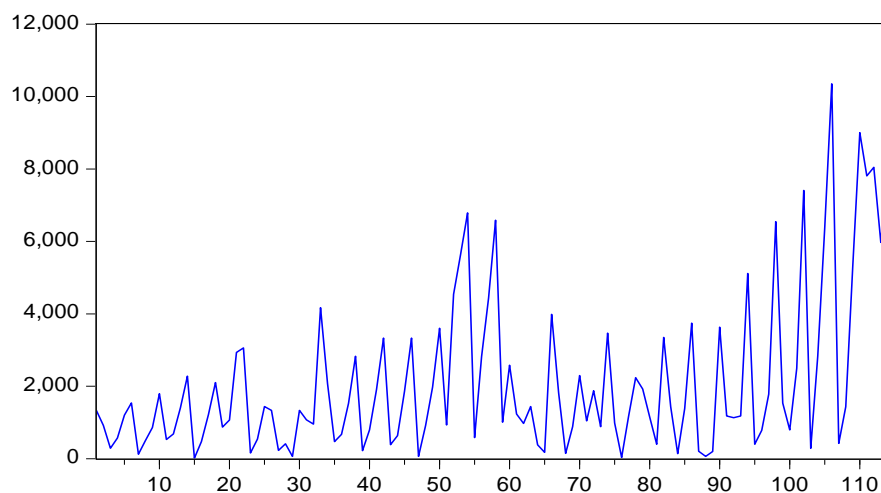


Figura 5.4.2 Jostacionariteti i serisë së Aktivitetit Financiar

Tabela 5.4.2 Testi Unit Root për serinë Aktivitetit Financiar

Null Hypothesis: FINANCIAL_ACTIVITIES has a unit root
 Exogenous: Constant
 Lag Length: 4 (Automatic - based on SIC, maxlag=12)

| | t-Statistic | Prob.* |
|----------------------------------------|-------------|--------|
| Augmented Dickey-Fuller test statistic | -0.716074 | 0.8374 |
| Test critical values: | | |
| 1% level | -3.491345 | |
| 5% level | -2.888157 | |
| 10% level | -2.581041 | |

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(FINANCIAL_ACTIVITIES)
 Method: Least Squares
 Date: 06/26/21 Time: 13:57
 Sample (adjusted): 6 114
 Included observations: 109 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|-----------------------------|-------------|-----------------------|-------------|----------|
| FINANCIAL_ACTIVITIES(-1) | -0.084124 | 0.117479 | -0.716074 | 0.4756 |
| D(FINANCIAL_ACTIVITIES(-1)) | -0.500249 | 0.154780 | -3.232004 | 0.0017 |
| D(FINANCIAL_ACTIVITIES(-2)) | -0.498657 | 0.139767 | -3.567777 | 0.0005 |
| D(FINANCIAL_ACTIVITIES(-3)) | -0.481677 | 0.121898 | -3.951470 | 0.0001 |
| D(FINANCIAL_ACTIVITIES(-4)) | 0.254210 | 0.106313 | 2.391141 | 0.0186 |
| C | 311.5795 | 267.5443 | 1.164590 | 0.2469 |
| R-squared | 0.635211 | Mean dependent var | | 62.99083 |
| Adjusted R-squared | 0.617502 | S.D. dependent var | | 2420.508 |
| S.E. of regression | 1496.996 | Akaike info criterion | | 17.51378 |
| Sum squared resid | 2.31E+08 | Schwarz criterion | | 17.66193 |
| Log likelihood | -948.5011 | Hannan-Quinn criter. | | 17.57386 |
| F-statistic | 35.87096 | Durbin-Watson stat | | 1.903158 |
| Prob(F-statistic) | 0.000000 | | | |

Investing Activities

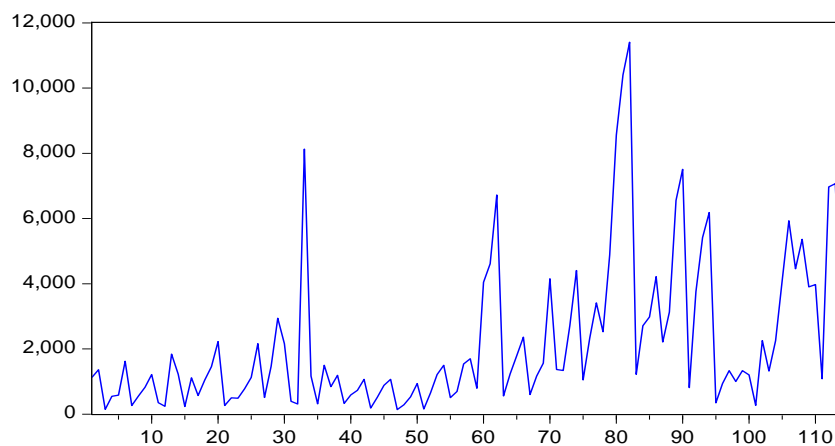


Figura 5.4.3 Jostacionariteti i serisë së Aktivitetit të Investimit

Tabela 5.4.3 Testi Unit Root për serinë Aktivitetit të Investimit

Null Hypothesis: INVESTING_ACTIVITIES has a unit root
 Exogenous: Constant
 Lag Length: 0 (Automatic - based on SIC, maxlag=12)

| | t-Statistic | Prob.* |
|----------------------------------------|-------------|--------|
| Augmented Dickey-Fuller test statistic | -5.971390 | 0.3000 |
| Test critical values: | | |
| 1% level | -3.489117 | |
| 5% level | -2.887190 | |
| 10% level | -2.580525 | |

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(INVESTING_ACTIVITIES)
 Method: Least Squares
 Date: 06/26/21 Time: 13:56
 Sample (adjusted): 2 114
 Included observations: 113 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------------|-------------|-----------------------|-------------|----------|
| INVESTING_ACTIVITIES(-1) | -0.485743 | 0.081345 | -5.971390 | 0.0000 |
| C | 1059.579 | 256.3708 | 4.132996 | 0.0001 |
| R-squared | 0.243135 | Mean dependent var | | 3.123894 |
| Adjusted R-squared | 0.236316 | S.D. dependent var | | 2256.952 |
| S.E. of regression | 1972.328 | Akaike info criterion | | 18.02936 |
| Sum squared resid | 4.32E+08 | Schwarz criterion | | 18.07763 |
| Log likelihood | -1016.659 | Hannan-Quinn criter. | | 18.04895 |
| F-statistic | 35.65750 | Durbin-Watson stat | | 2.005795 |
| Prob(F-statistic) | 0.000000 | | | |

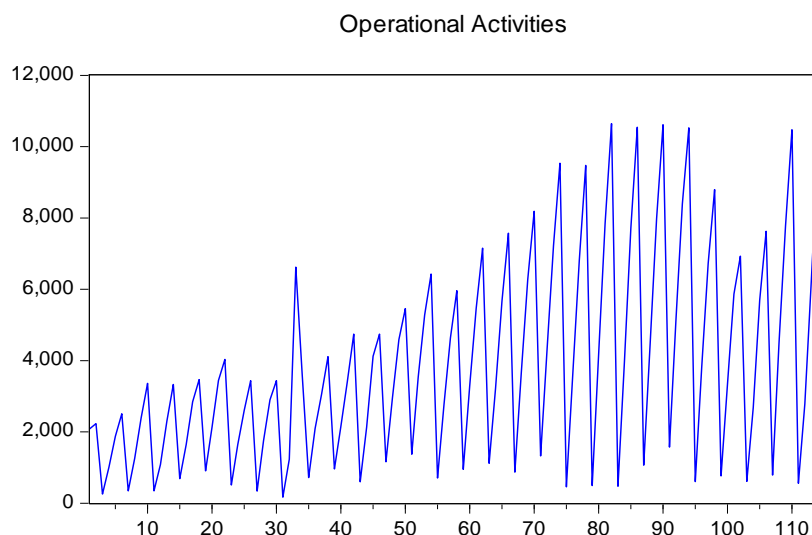


Figura 5.4.4 Jostacionariteti i serisë së Aktivitetit të Shfrytëzimit

Tabela 5.4.4 Testi Unit Root për serinë Aktivitetit të Shfrytëzimit

Null Hypothesis: OPERATIONAL_ACTIVITIES has a unit root
 Exogenous: Constant
 Lag Length: 4 (Automatic - based on SIC, maxlag=12)

| | t-Statistic | Prob.* |
|----------------------------------------|-------------|--------|
| Augmented Dickey-Fuller test statistic | -1.668773 | 0.4442 |
| Test critical values: | | |
| 1% level | -3.491345 | |
| 5% level | -2.888157 | |
| 10% level | -2.581041 | |

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(OPERATIONAL_ACTIVITIES)
 Method: Least Squares
 Date: 06/26/21 Time: 13:55
 Sample (adjusted): 6 114
 Included observations: 109 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|-------------------------------|-------------|-----------------------|-------------|--------|
| OPERATIONAL_ACTIVITIES(-1) | -0.092612 | 0.055497 | -1.668773 | 0.0982 |
| D(OPERATIONAL_ACTIVITIES(-1)) | -0.598940 | 0.102139 | -5.863994 | 0.0000 |
| D(OPERATIONAL_ACTIVITIES(-2)) | -0.629917 | 0.098155 | -6.417584 | 0.0000 |
| D(OPERATIONAL_ACTIVITIES(-3)) | -0.655022 | 0.095713 | -6.843628 | 0.0000 |
| D(OPERATIONAL_ACTIVITIES(-4)) | 0.323277 | 0.095898 | 3.371056 | 0.0011 |
| C | 440.0482 | 227.7070 | 1.932519 | 0.0560 |
| R-squared | 0.951588 | Mean dependent var | 73.08257 | |
| Adjusted R-squared | 0.949238 | S.D. dependent var | 3769.201 | |
| S.E. of regression | 849.2195 | Akaike info criterion | 16.37999 | |
| Sum squared resid | 74280891 | Schwarz criterion | 16.52813 | |
| Log likelihood | -886.7092 | Hannan-Quinn criter. | 16.44006 | |
| F-statistic | 404.9126 | Durbin-Watson stat | 2.031690 | |
| Prob(F-statistic) | 0.000000 | | | |

Sikurse evidentohet nga Figura 5.4.1, 5.4.2, 5.4.3 dhe 5.4.4, seritë që kemi marrë në studim janë jostacionare. Ky fakt konkretizohet edhe nëpërmjet rezultateve të testit Unit Root, të zbatuar për secilën seri tek Tabela 5.4.1, 5.4.2, 5.4.3 dhe 5.4.4.

Meqenëse probabilitetet përkatëse të testeve për secilën janë: 0.4312, 0.300, 0.8374 dhe 0.4442, pra janë më të mëdha se vlera kritike prej 5%, konkludojmë se hipoteza bazë që shpreh jostacionaritetin e serive pranohet. Për të eliminuar problemin e jostacionaritetit ne kemi kryer diferencimin e secilës prej serive, për të rregulluar ndryshimet në nivelet e serive, dhe për të zbutur trendin e tyre.

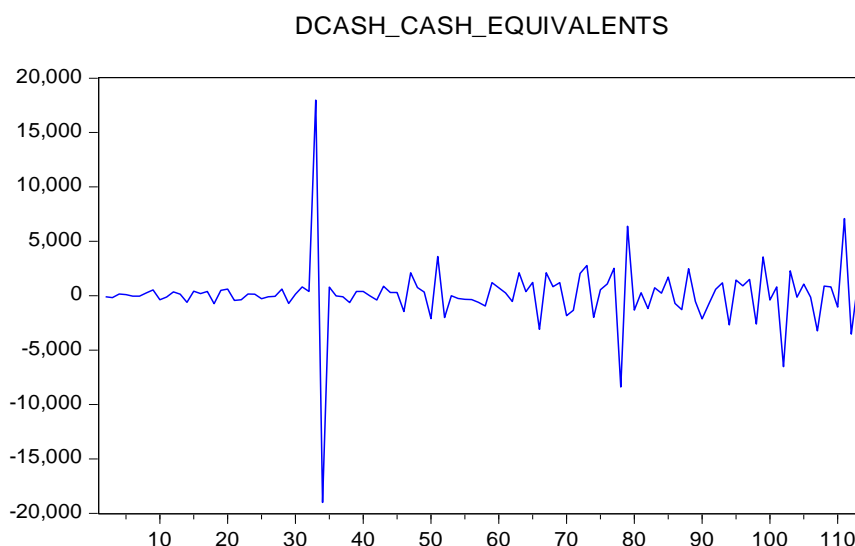


Figura 5.4.5 Stacionariteti i serisë së Parave dhe Ekuivalentëve të Parave

Tabela 5.4.5 Stacionariteti i serisë së Parave dhe Ekuivalentëve të Parave

Null Hypothesis: DCASH_CASH_EQUIVALENTS has a unit root

Exogenous: Constant

Lag Length: 2 (Automatic - based on SIC, maxlag=12)

| | t-Statistic | Prob.* |
|----------------------------------------|-------------|--------|
| Augmented Dickey-Fuller test statistic | -10.27355 | 0.0000 |
| Test critical values: | | |
| 1% level | -3.490772 | |
| 5% level | -2.887909 | |
| 10% level | -2.580908 | |

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation

Dependent Variable: D(DCASH_CASH_EQUIVALENTS)

Method: Least Squares

Date: 06/26/21 Time: 13:58

Sample (adjusted): 5 114

Included observations: 110 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|-------------------------------|-------------|-----------------------|-------------|--------|
| DCASH_CASH_EQUIVALENTS(-1) | -2.482254 | 0.241616 | -10.27355 | 0.0000 |
| D(DCASH_CASH_EQUIVALENTS(-1)) | 0.751484 | 0.180939 | 4.153233 | 0.0001 |
| D(DCASH_CASH_EQUIVALENTS(-2)) | 0.294790 | 0.095085 | 3.100280 | 0.0025 |
| C | 194.9927 | 244.3143 | 0.798122 | 0.4266 |
| R-squared | 0.787872 | Mean dependent var | -43.17273 | |
| Adjusted R-squared | 0.781868 | S.D. dependent var | 5461.514 | |
| S.E. of regression | 2550.779 | Akaike info criterion | 18.56187 | |
| Sum squared resid | 6.90E+08 | Schwarz criterion | 18.66007 | |
| Log likelihood | -1016.903 | Hannan-Quinn criter. | 18.60170 | |
| F-statistic | 131.2324 | Durbin-Watson stat | 2.056134 | |
| Prob(F-statistic) | 0.000000 | | | |

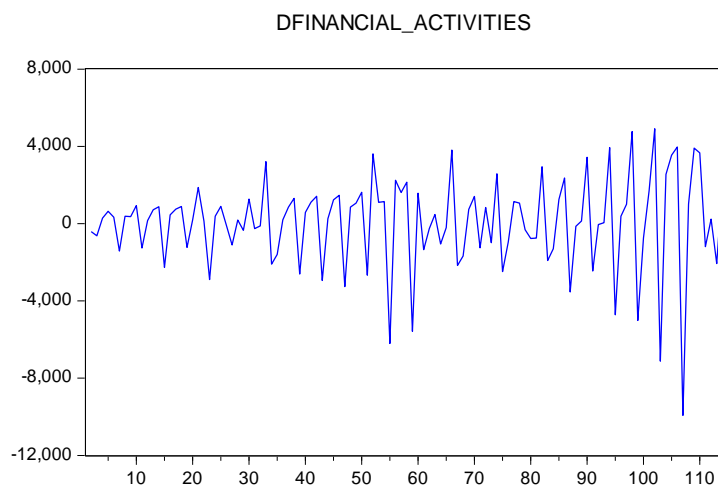


Figura 5.4.6 Stacionariteti i serisë së Aktivitetit Financiar

Tabela 5.4.6 Stacionariteti i serisë së Aktivitetit Financiar

Null Hypothesis: DFINANCIAL_ACTIVITIES has a unit root
 Exogenous: Constant
 Lag Length: 4 (Automatic - based on SIC, maxlag=12)

| | t-Statistic | Prob.* |
|----------------------------------------|-------------|--------|
| Augmented Dickey-Fuller test statistic | -7.231456 | 0.0000 |
| Test critical values: | | |
| 1% level | -3.491928 | |
| 5% level | -2.888411 | |
| 10% level | -2.581176 | |

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(DFINANCIAL_ACTIVITIES)
 Method: Least Squares
 Date: 06/26/21 Time: 13:57
 Sample (adjusted): 7 114
 Included observations: 108 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|------------------------------|-------------|-----------------------|-------------|----------|
| DFINANCIAL_ACTIVITIES(-1) | -3.002844 | 0.415248 | -7.231456 | 0.0000 |
| D(DFINANCIAL_ACTIVITIES(-1)) | 1.464688 | 0.368460 | 3.975162 | 0.0001 |
| D(DFINANCIAL_ACTIVITIES(-2)) | 0.782237 | 0.274892 | 2.845614 | 0.0054 |
| D(DFINANCIAL_ACTIVITIES(-3)) | 0.125306 | 0.180782 | 0.693135 | 0.4898 |
| D(DFINANCIAL_ACTIVITIES(-4)) | 0.221545 | 0.097757 | 2.266295 | 0.0255 |
| C | 178.2512 | 143.9301 | 1.238456 | 0.2184 |
| R-squared | 0.866993 | Mean dependent var | | 16.30556 |
| Adjusted R-squared | 0.860473 | S.D. dependent var | | 3937.374 |
| S.E. of regression | 1470.739 | Akaike info criterion | | 17.47887 |
| Sum squared resid | 2.21E+08 | Schwarz criterion | | 17.62788 |
| Log likelihood | -937.8590 | Hannan-Quinn criter. | | 17.53929 |
| F-statistic | 132.9754 | Durbin-Watson stat | | 1.990605 |
| Prob(F-statistic) | 0.000000 | | | |

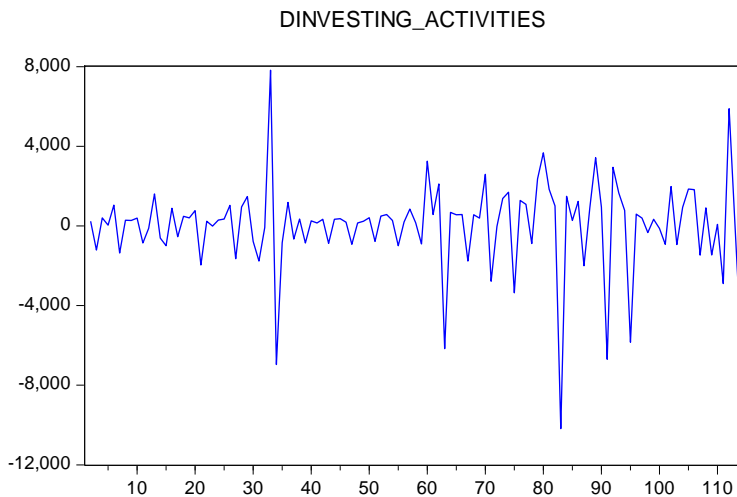


Figura 5.4.7 Stacionariteti i serisë së Aktivitetit të Investimit

Tabela 5.4.7 Stacionariteti i serisë së Aktivitetit të Investimit

Null Hypothesis: DINVESTING_ACTIVITIES has a unit root
 Exogenous: Constant
 Lag Length: 2 (Automatic - based on SIC, maxlag=12)

| | t-Statistic | Prob.* |
|----------------------------------------|-------------|--------|
| Augmented Dickey-Fuller test statistic | -10.57837 | 0.0000 |
| Test critical values: | | |
| 1% level | -3.490772 | |
| 5% level | -2.887909 | |
| 10% level | -2.580908 | |

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(DINVESTING_ACTIVITIES)
 Method: Least Squares
 Date: 06/26/21 Time: 13:56
 Sample (adjusted): 5 114
 Included observations: 110 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|------------------------------|-------------|-----------------------|-------------|-----------|
| DINVESTING_ACTIVITIES(-1) | -2.253784 | 0.213056 | -10.57837 | 0.0000 |
| D(DINVESTING_ACTIVITIES(-1)) | 0.791364 | 0.158643 | 4.988326 | 0.0000 |
| D(DINVESTING_ACTIVITIES(-2)) | 0.331317 | 0.096843 | 3.421173 | 0.0009 |
| C | 60.91701 | 190.5989 | 0.319608 | 0.7499 |
| R-squared | 0.698574 | Mean dependent var | | -54.50909 |
| Adjusted R-squared | 0.690043 | S.D. dependent var | | 3586.713 |
| S.E. of regression | 1996.859 | Akaike info criterion | | 18.07222 |
| Sum squared resid | 4.23E+08 | Schwarz criterion | | 18.17042 |
| Log likelihood | -989.9724 | Hannan-Quinn criter. | | 18.11205 |
| F-statistic | 81.88719 | Durbin-Watson stat | | 1.929694 |
| Prob(F-statistic) | 0.000000 | | | |

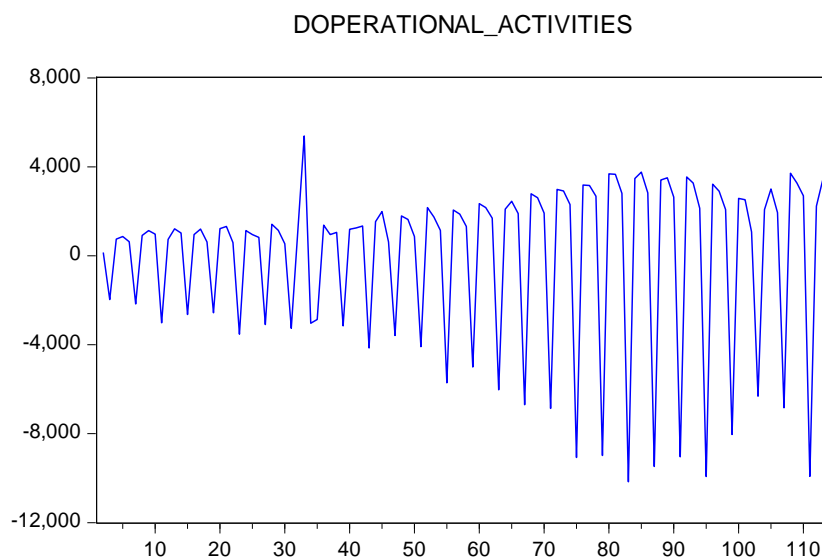


Figura 5.4.8 Stacionariteti i serisë së Aktivitetit të Shfrytëzimit

Tabela 5.4.8 Stacionariteti i serisë së Aktivitetit të Shfrytëzimit

Null Hypothesis: DOPERATIONAL_ACTIVITIES has a unit root
 Exogenous: Constant
 Lag Length: 3 (Automatic - based on SIC, maxlag=12)

| | t-Statistic | Prob.* |
|----------------------------------------|-------------|--------|
| Augmented Dickey-Fuller test statistic | -7.240307 | 0.0000 |
| Test critical values: | | |
| 1% level | -3.491345 | |
| 5% level | -2.888157 | |
| 10% level | -2.581041 | |

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(DOPERATIONAL_ACTIVITIES)
 Method: Least Squares
 Date: 06/26/21 Time: 13:55
 Sample (adjusted): 6 114
 Included observations: 109 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------------------|-------------|-----------------------|-------------|----------|
| DOPERATIONAL_ACTIVITIES(-1) | -2.708803 | 0.374128 | -7.240307 | 0.0000 |
| D(DOPERATIONAL_ACTIVITIES(-1)) | 1.038439 | 0.282588 | 3.674743 | 0.0004 |
| D(DOPERATIONAL_ACTIVITIES(-2)) | 0.360060 | 0.189931 | 1.895744 | 0.0608 |
| D(DOPERATIONAL_ACTIVITIES(-3)) | -0.320446 | 0.096702 | -3.313747 | 0.0013 |
| C | 85.73161 | 82.98324 | 1.033120 | 0.3039 |
| R-squared | 0.980442 | Mean dependent var | | 25.28440 |
| Adjusted R-squared | 0.979690 | S.D. dependent var | | 6009.791 |
| S.E. of regression | 856.4754 | Akaike info criterion | | 16.38831 |
| Sum squared resid | 76289219 | Schwarz criterion | | 16.51177 |
| Log likelihood | -888.1631 | Hannan-Quinn criter. | | 16.43838 |
| F-statistic | 1303.391 | Durbin-Watson stat | | 2.019719 |
| Prob(F-statistic) | 0.000000 | | | |

Tek Fig. 5.4.5, 5.4.6, 5.4.7 dhe 5.4.8 evidentohet vizualisht se pas kryerjes së diferencimeve, seritë janë të qendrueshme (stacionare) dhe ne mund të vazhdojmë me analizën e mëtejshme.

Gjithashtu rezultatet e Testit Unit Root, përkatësisht tek Tabela 5.4.5, 5.4.6, 5.4.7 dhe 5.4.8, vërtetojnë ekzistencën e stacionaritetit, pasi probabilitetet faktike janë 0.0000, pra më të vogla se probabiliteti kritik prej 0.05, gjë e cila çon në pranimin e hipotezës alternative e cila konsiston tek stacionariteti i serive.

Tabela 5.4.9 Modeli i Vektorit të Autoregresit (VAR)

Vector Autoregression Estimates

Sample (adjusted): 8 114

Included observations: 107 after adjustments

Standard errors in () & t-statistics in []

| | D(CASH&CASH EQUIVALENTS) | D(FINANCIAL ACTIVITIES) | D(INVESTING ACTIVITIES) | D(OPERATIONAL ACTIVITIES) |
|---------------------------------|---------------------------------------------|---------------------------------------------|---------------------------------------------|---------------------------------------------|
| D(CASH&CASH EQUIVALENTS(-1)) | -0.668866 (0.13349) [-5.01057] | -0.091659 (0.07610) [-1.20448] | -0.053719 (0.04381) [-1.22632] | -0.332616 (0.09126) [-3.64485] |
| D(CASH&CASH EQUIVALENTS(-2)) | -0.525977 (0.15436) [-3.40756] | -0.084526 (0.08799) [-0.96060] | -0.062972 (0.05065) [-1.24323] | -0.240144 (0.10552) [-2.27581] |
| D(CASH&CASH EQUIVALENTS(-3)) | -0.356864 (0.15492) [-2.30357] | -0.040438 (0.08831) [-0.45789] | 0.054818 (0.05084) [1.07832] | -0.097919 (0.10590) [-0.92460] |
| D(CASH&CASH EQUIVALENTS(-4)) | -0.382452 (0.14903) [-2.56632] | -0.179911 (0.08496) [-2.11770] | -0.123457 (0.04890) [-2.52449] | -0.185442 (0.10188) [-1.82024] |
| D(CASH&CASH EQUIVALENTS(-5)) | -0.189886 (0.15430) [-1.23064] | 0.052661 (0.08796) [0.59869] | 0.019503 (0.05063) [0.38518] | -0.159995 (0.10548) [-1.51682] |
| D(CASH&CASH EQUIVALENTS(-6)) | -0.167242 (0.13840) [-1.20838] | -0.010659 (0.07890) [-0.13510] | 0.032401 (0.04542) [0.71340] | 0.131453 (0.09461) [1.38936] |
| D(FINANCIAL ACTIVITIES(-1)) | -0.160960 (0.20676) [-0.77850] | -0.406667 (0.11787) [-3.45027] | -0.053313 (0.06785) [-0.78578] | 0.232119 (0.14134) [1.64225] |
| D(FINANCIAL ACTIVITIES(-2)) | 0.031356 (0.22243) [0.14097] | -0.541706 (0.12680) [-4.27222] | 0.027651 (0.07299) [0.37884] | 0.260827 (0.15205) [1.71537] |
| D(FINANCIAL ACTIVITIES(-3)) | -0.319897 (0.25503) [-1.25437] | -0.423035 (0.14538) [-2.90981] | -0.122929 (0.08369) [-1.46891] | -0.098035 (0.17434) [-0.56232] |

| | | | | |
|----------------------------------|---------------------------------------------|---------------------------------------------|---------------------------------------------|---------------------------------------------|
| D(FINANCIAL ACTIVITIES(-4)) | -0.541117 (0.26642) [-2.03104] | 0.007726 (0.15188) [0.05087] | -0.044188 (0.08743) [-0.50543] | -0.191411 (0.18213) [-1.05095] |
| D(FINANCIAL ACTIVITIES(-5)) | -0.324150 (0.23857) [-1.35870] | -0.276057 (0.13600) [-2.02979] | -0.036015 (0.07829) [-0.46003] | -0.350070 (0.16309) [-2.14644] |
| D(FINANCIAL ACTIVITIES(-6)) | -0.327675 (0.23020) [-1.42342] | 0.033372 (0.13123) [0.25430] | -0.075560 (0.07554) [-1.00025] | -0.299464 (0.15737) [-1.90293] |
| D(OPERATIONAL ACTIVITIES(-1)) | 0.186690 (0.38321) [0.48717] | 0.314605 (0.21846) [1.44014] | -0.471118 (0.12575) [-3.74645] | -0.301135 (0.26197) [-1.14951] |
| D(OPERATIONAL ACTIVITIES(-2)) | -0.032288 (0.42176) [-0.07656] | 0.328493 (0.24043) [1.36625] | -0.512308 (0.13840) [-3.70158] | 0.220912 (0.28832) [0.76619] |
| D(OPERATIONAL ACTIVITIES(-3)) | 0.649021 (0.46551) [1.39422] | 0.258118 (0.26537) [0.97267] | -0.625787 (0.15276) [-4.09661] | -0.207844 (0.31823) [-0.65313] |
| D(OPERATIONAL ACTIVITIES(-4)) | 0.384515 (0.48351) [0.79526] | 0.474269 (0.27563) [1.72065] | 0.293443 (0.15866) [1.84945] | 0.376037 (0.33054) [1.13766] |
| D(OPERATIONAL ACTIVITIES(-5)) | 0.530273 (0.47324) [1.12053] | 0.176837 (0.26978) [0.65550] | -0.179361 (0.15529) [-1.15499] | -0.008315 (0.32351) [-0.02570] |
| D(OPERATIONAL ACTIVITIES(-6)) | 0.226339 (0.42818) [0.52861] | 0.022701 (0.24409) [0.09300] | -0.290575 (0.14051) [-2.06804] | -0.289459 (0.29271) [-0.98889] |
| D(INVESTING ACTIVITIES(-1)) | -0.299757 (0.18942) [-1.58250] | -0.077593 (0.10798) [-0.71857] | -0.000185 (0.06216) [-0.00298] | -0.051807 (0.12949) [-0.40009] |
| D(INVESTING ACTIVITIES(-2)) | -0.108244 (0.19415) [-0.55751] | -0.099940 (0.11068) [-0.90295] | 0.091103 (0.06371) [1.42991] | -0.332775 (0.13273) [-2.50720] |
| D(INVESTING ACTIVITIES(-3)) | -0.312472 (0.19793) [-1.57871] | -0.103640 (0.11283) [-0.91853] | 0.010922 (0.06495) [0.16816] | -0.202393 (0.13531) [-1.49581] |
| D(INVESTING ACTIVITIES(-4)) | 0.046118 (0.20178) [0.22856] | -0.015811 (0.11503) [-0.13746] | 0.072284 (0.06621) [1.09168] | -0.306442 (0.13794) [-2.22158] |
| D(INVESTING ACTIVITIES(-5)) | -0.328145 | -0.100942 | -0.010035 | 0.069585 |

| | | | | |
|-----------------------------------------|-----------------|-----------------|-----------------|------------------|
| | (0.20484) | (0.11677) | (0.06722) | (0.14003) |
| | [-1.60196] | [-0.86443] | [-0.14928] | [0.49692] |
| D(INVESTING ACTIVITIES(-6)) | 0.042531 | 0.047708 | 0.071153 | -0.291272 |
| | (0.20354) | (0.11603) | (0.06679) | (0.13915) |
| | [0.20895] | [0.41115] | [1.06528] | [-2.09329] |
| C | 285.8963 | 136.8409 | 105.6496 | 121.9427 |
| | (243.890) | (139.034) | (80.0329) | (166.727) |
| | [1.17223] | [0.98423] | [1.32008] | [0.73139] |
| R-squared | 0.568153 | 0.757586 | 0.966879 | 0.611606 |
| Adj. R-squared | 0.441759 | 0.686636 | 0.957185 | 0.497930 |
| Sum sq. resids | 4.70E+08 | 1.53E+08 | 50642496 | 2.20E+08 |
| S.E. equation | 2394.841 | 1365.220 | 785.8698 | 1637.152 |
| F-statistic | 4.495095 | 10.67769 | 99.73923 | 5.380245 |
| Log likelihood | -970.1643 | -910.0301 | -850.9362 | -929.4659 |
| Akaike AIC | 18.60120 | 17.47720 | 16.37264 | 17.84048 |
| Schwarz SC | 19.22569 | 18.10169 | 16.99713 | 18.46498 |
| Mean dependent | 54.54206 | 74.30841 | 88.79439 | 11.32710 |
| S.D. dependent | 3205.280 | 2438.810 | 3797.960 | 2310.504 |
| Determinant resid covariance (dof adj.) | | 8.79E+24 | | |
| Determinant resid covariance | | 3.03E+24 | | |
| Log likelihood | | -3623.183 | | |
| Akaike information criterion | | 69.59221 | | |
| Schwarz criterion | | 72.09018 | | |
| Number of coefficients | | 100 | | |

Bazuar në Tabela 5.4.9 vërrejmë se rëndësia dhe ndikimi aktual dhe potencial i çdo serie në seritë e tjera shpjegohet nëpërmjet tre rreshtave ku përkatësisht, rreshti i parë shpreh koeficientin me të cilin një seri ndikon seritë e tjera, rreshti i dytë përcakton gabimin standard të relacionit midis dy serive dhe rreshti i tretë është rezultati i testit të studentit që fitohet përmes raportit të koeficientit me gabimet (devijimet standarde). Meqë Vektori i Autoregresit bazohet tek trajtimi i çdo serie të sistemit si një funksion i vlerave të vonuara të të gjitha serive të përfshira në model, ne mund të interpretojmë tendenën e serive, të cilat ndikohen prej njëra-tjetrës në funksion të vlerave të vonuara të përfshira tek Tabela 5.4.9. Pra, nëse ne synojmë të kemi rritje të gjendjes së parave në fund të tremujorit të parë të 2021 me 1%, duhet që paratë e angazhuara gjatë tremujorit të katërt të vitit 2020 tek Aktiviteti i Investimit të ulen me 5.4%, paratë e angazhuara tek Aktiviteti i Shfrytëzimit të reduktohen me 33.3% dhe paratë e Aktivitetit të Financimit në formën e huamarrjeve të reduktohen me 9.2%. Gjithashtu, ky parashikim sigurisht që ka një nivel të caktuar gabimi, i cili paraqitet në rreshtin e dytë të çdo gjenerimi. Konkretisht, nevoja e reduktimit të parave të angazhuara në Aktiviteti i Financimit, Investimit dhe Shfrytëzimit ka nivel pasaktësie përkatësisht: 7.6%, 4.4% dhe 9.13%. E njëjta logjikë ndiqet për interpretimin e influencave të secilës prej vlerave të vonuara të katër serive të konsideruara tek çdo seri e marrë në studim.

Meqë algoritmi parashikues VAR analizon dhe parashikon seritë kohore që influencojnë njëra-tjetrën, në fund të Tabelës 5.4.9 jepen gjenerimet mbi efektet e përbashkëta të serive tek modeli. Shpjegueshmëria e modelit është e lartë, por shpjegueshmëria më e lartë sigurohet prej serisë së Aktivitetit të Investimit në masën 97%. Varianca (mbetjet) janë minimale dhe rëndësia e modelit është e konsiderueshme.

Tabela 5.4.10 Testimi për korrelacion serial (LM Test)

Sample: 1 114
Included observations: 107

Null hypothesis: No serial correlation at lag h

| Lag | LRE* stat | df | Prob. | Rao F-stat | df | Prob. |
|-----|-----------|----|--------|------------|-------------|--------|
| 1 | 19.71434 | 16 | 0.2334 | 1.246910 | (16, 229.8) | 0.2337 |
| 2 | 17.07720 | 16 | 0.3806 | 1.074038 | (16, 229.8) | 0.3810 |
| 3 | 22.78647 | 16 | 0.1196 | 1.450740 | (16, 229.8) | 0.1198 |
| 4 | 32.42915 | 16 | 0.0088 | 2.107965 | (16, 229.8) | 0.0888 |
| 5 | 21.64000 | 16 | 0.1552 | 1.374364 | (16, 229.8) | 0.1555 |
| 6 | 19.52886 | 16 | 0.2422 | 1.234688 | (16, 229.8) | 0.2426 |
| 7 | 13.28754 | 16 | 0.6516 | 0.828962 | (16, 229.8) | 0.6519 |

Null hypothesis: No serial correlation at lags 1 to h

| Lag | LRE* stat | df | Prob. | Rao F-stat | df | Prob. |
|-----|-----------|-----|--------|------------|--------------|--------|
| 1 | 19.71434 | 16 | 0.2334 | 1.246910 | (16, 229.8) | 0.2337 |
| 2 | 38.33890 | 32 | 0.2040 | 1.215297 | (32, 263.4) | 0.2054 |
| 3 | 52.56696 | 48 | 0.3016 | 1.105646 | (48, 260.1) | 0.3060 |
| 4 | 75.42465 | 64 | 0.1554 | 1.202771 | (64, 248.9) | 0.1622 |
| 5 | 100.0980 | 80 | 0.0637 | 1.295347 | (80, 235.2) | 0.0707 |
| 6 | 109.2063 | 96 | 0.1685 | 1.158040 | (96, 220.4) | 0.1901 |
| 7 | 112.1851 | 112 | 0.4773 | 0.987564 | (112, 205.1) | 0.5235 |

*Edgeworth expansion corrected likelihood ratio statistic.

Për të studiuar më tej modelin, është e nevojshme të identifikohet nëse modeli është i lirë nga heteroskedastikiteti dhe autokorrelacioni. Pra, në Tabela 5.4.10 paraqitet Testi LM, i cili studion ekzistencën ose jo të korrelacionit serial për vonesën h në modelin e krijuar VAR. Meqënëse të gjitha probabilitetet për secilën nga shtatë vonesat e marra në shqyrtim kanë vlera më të mëdha se 0.05, rrjedhimisht hipoteza bazë nuk do të pranohet. Edhe modeli Edgeworth i cili jep një rezultat më të saktë rreth ekzistencës së korrelacionit serial, sërish vërteton se modeli VAR i krijuar për katër seritë kohore të marra në studim është i lirë nga korrelacioni serial.

Tabela 5.4.11 Testi i Heteroskedasticitetit për modelin VAR

VAR Residual Heteroskedasticity Tests (Levels and Squares)
 Date: 02/14/21 Time: 13:54
 Sample: 1 114
 Included observations: 107

| Joint test: | | | | | |
|-------------|-----|--------|--|--|--|
| Chi-sq | Df | Prob. | | | |
| 573.0495 | 480 | 0.2022 | | | |

| Individual components: | | | | | |
|------------------------|-----------|----------|--------|------------|--------|
| Dependent | R-squared | F(48,58) | Prob. | Chi-sq(48) | Prob. |
| res1*res1 | 0.226286 | 0.353398 | 0.9998 | 24.21260 | 0.9984 |
| res2*res2 | 0.374411 | 0.723181 | 0.8753 | 40.06201 | 0.7854 |
| res3*res3 | 0.317155 | 0.561224 | 0.9794 | 33.93559 | 0.9377 |
| res4*res4 | 0.342811 | 0.630305 | 0.9492 | 36.68075 | 0.8833 |
| res2*res1 | 0.285315 | 0.482389 | 0.9948 | 30.52874 | 0.9768 |
| res3*res1 | 0.219804 | 0.340423 | 0.9999 | 23.51903 | 0.9989 |
| res3*res2 | 0.354825 | 0.664543 | 0.9266 | 37.96625 | 0.8500 |
| res4*res1 | 0.226309 | 0.353445 | 0.9998 | 24.21508 | 0.9983 |
| res4*res2 | 0.365132 | 0.694950 | 0.9020 | 39.06913 | 0.8175 |
| res4*res3 | 0.252436 | 0.408028 | 0.9991 | 27.01067 | 0.9938 |

Në Tabela 5.4.11, është e lehtë të identifkohet se probabilitetet e studimit të varësisë së çdo mbetje nga njëra-tjetra, janë më të mëdha se 0.05, dhe kjo tregon se hipoteza zero që shpreh ekzistencën e heteroskedasticitetit nuk mund të pranohet.

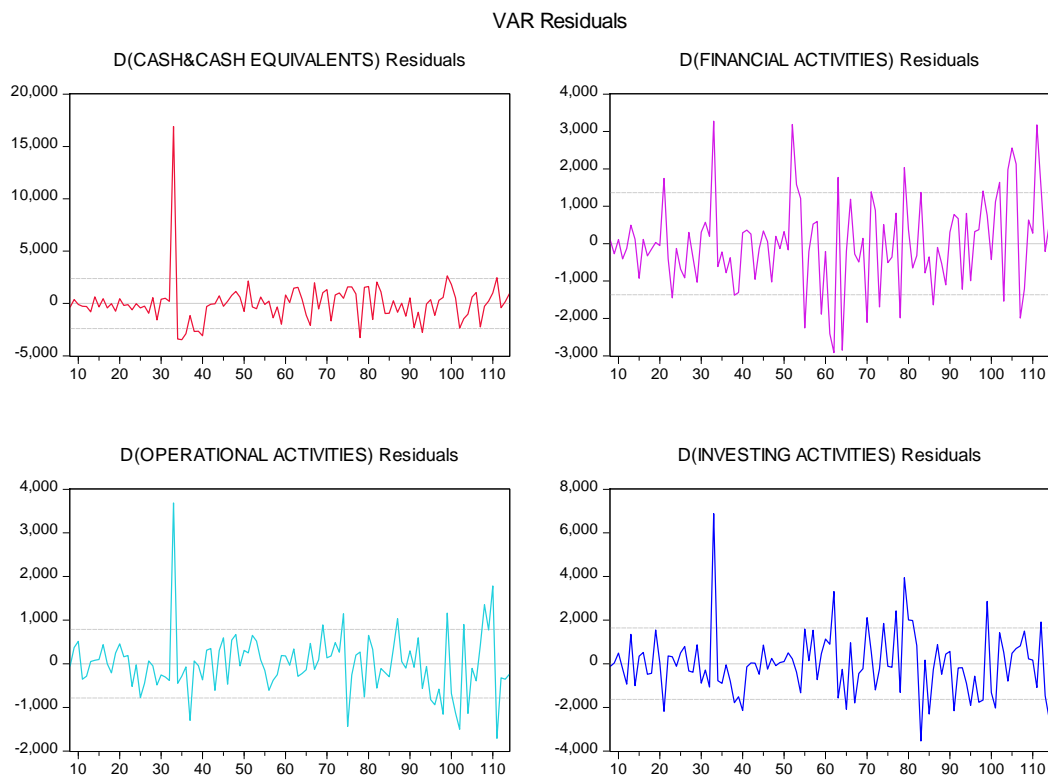


Figura 5.4.9 Mbetjet e Modelit VAR

Meqenëse Modeli VAR mbart termin e gabimit (mbetjeve), Fig. 5.4.9 jep një paraqitje grafike të mbetjeve për secilën prej serive, nga ku evidentohet qendrueshmëri e prirjes së mbetjeve dhe lehtësi në parashikimin e trendit të gabimit në vitet e ardhshme.

Për të rritur kuptueshmërinë e modelit VAR në termat e parashikimit të tendencave të ardhshme, ne kemi paraqitur Përgjigjet e Impulsit, ku një çështje e rëndësishme e tyre janë: Analiza Asimptotike që nënkupton Modelin e Përafrimit Edgeworth, Përafrimin Monte Carlo për 1000 rizgjedhje dhe Versionin Bootstrap për 1000 rizgjedhje dhe për nivel besueshmërie 95% .

Tabela 5.4.12 Analiza Asimptotike (Përafrimi Edgeworth) për identifikimin e përgjigjes përkatëse të impulsit për çdo seri

Response of D(CASH&CASH EQUIVALENTS):

| Period | D(CASH&CASH EQUIVALENTS) | D(FINANCIAL ACTIVITIES) | D(INVESTING ACTIVITIES) | D(OPERATIONAL ACTIVITIES) |
|--------|--------------------------|-------------------------|-------------------------|---------------------------|
| 1 | 2394.841 (163.708) | 0.000000 (0.000000) | 0.000000 (0.000000) | 0.000000 (0.000000) |
| 2 | -1826.573 (289.309) | -157.7709 (265.879) | -10.38674 (256.358) | -411.8117 (261.747) |
| 3 | 134.3653 (328.880) | 153.8764 (326.177) | -87.20782 (327.976) | 165.1840 (315.895) |
| 4 | -29.87395 (329.661) | -466.9623 (321.799) | 330.6794 (305.294) | -179.7698 (322.856) |
| 5 | -233.3102 (335.568) | -314.7734 (303.122) | -76.14772 (184.436) | 399.0172 (331.251) |
| 6 | 297.3177 (339.679) | 554.6333 (307.600) | -183.4638 (197.308) | -191.1481 (337.646) |
| 7 | 49.41530 (338.443) | 74.94726 (309.923) | -28.38113 (195.673) | 161.8234 (346.212) |
| 8 | 135.6295 (336.174) | 82.87116 (296.047) | 144.2722 (181.838) | 141.3084 (275.995) |
| 9 | -284.6496 (237.246) | -78.07673 (215.872) | -41.52852 (146.351) | -349.2871 (245.257) |
| 10 | -62.83569 (183.436) | -166.1535 (227.294) | 60.96567 (145.817) | 197.3186 (221.519) |

Response of D(FINANCIAL ACTIVITIES):

| Period | D(CASH&CASH EQUIVALENTS) | D(FINANCIAL ACTIVITIES) | D(INVESTING ACTIVITIES) | D(OPERATIONAL ACTIVITIES) |
|--------|--------------------------|-------------------------|-------------------------|---------------------------|
| 1 | 452.1291 (128.311) | 1288.179 (88.0581) | 0.000000 (0.000000) | 0.000000 (0.000000) |
| 2 | -351.0695 (158.861) | -537.1552 (155.530) | 183.6154 (145.278) | -106.5981 (148.527) |
| 3 | -130.6953 (170.712) | -507.9599 (165.972) | 23.61292 (164.415) | -50.76067 (160.764) |
| 4 | 92.96241 (172.659) | -79.66856 (169.289) | -118.5471 (152.975) | 37.34091 (185.526) |
| 5 | -4.760887 (181.465) | 511.8143 (172.489) | 74.25473 (96.4010) | 153.5016 (185.797) |
| 6 | 217.8045 (183.054) | -195.4912 (174.402) | 4.729393 (111.332) | -54.24950 (207.385) |
| 7 | -62.22006 (182.373) | 51.91789 (178.470) | -37.03803 (109.356) | 32.10449 (209.134) |
| 8 | -97.76470 (179.007) | 41.81660 (177.297) | -182.4105 (89.1186) | 38.38268 (163.269) |
| 9 | -102.6911 (113.010) | 37.31912 (157.644) | 282.0539 (87.5493) | -121.5019 (159.113) |
| 10 | 74.33718 (108.801) | -339.0421 (160.595) | 17.39194 (88.5959) | -1.790863 (121.752) |

Response of D(OPERATIONAL ACTIVITIES):

| Period | D(CASH&CASH EQUIVALENTS) | D(FINANCIAL ACTIVITIES) | D(INVESTING ACTIVITIES) | D(OPERATIONAL ACTIVITIES) |
|--------|--------------------------|-------------------------|-------------------------|---------------------------|
| 1 | 344.1629 | -98.12310 | 699.6536 | 0.000000 |

| | | | | |
|----|-----------|-----------|-----------|-----------|
| | (72.2384) | (67.9698) | (47.8273) | (0.00000) |
| 2 | -315.0287 | -22.40695 | -329.7068 | -0.254375 |
| | (92.8589) | (91.9093) | (86.1053) | (85.3943) |
| 3 | 16.49977 | 112.8611 | -169.4401 | 153.0970 |
| | (100.606) | (99.6693) | (99.7639) | (95.8599) |
| 4 | 51.02785 | -96.27422 | -196.2446 | -46.59434 |
| | (102.746) | (100.761) | (92.3634) | (104.871) |
| 5 | -75.44071 | -5.808033 | 608.0812 | -2.434657 |
| | (119.201) | (113.973) | (75.4234) | (107.397) |
| 6 | 87.91146 | 61.75002 | -370.4441 | -65.27948 |
| | (127.496) | (122.946) | (94.3426) | (132.626) |
| 7 | -48.70167 | -6.001829 | -102.4900 | 117.3536 |
| | (129.149) | (126.140) | (97.9135) | (136.154) |
| 8 | 32.71189 | -96.41442 | -177.7790 | -35.30156 |
| | (128.583) | (125.259) | (75.3909) | (116.996) |
| 9 | -49.14106 | 114.8115 | 637.3146 | -146.5617 |
| | (112.800) | (130.570) | (82.4227) | (112.711) |
| 10 | 61.76571 | -57.35303 | -323.5077 | 41.83503 |
| | (114.588) | (138.660) | (101.816) | (132.626) |

Response of D(INVESTING_ACTIVITIES):

| Period | D(CASH&CASH EQUIVALENTS) | D(FINANCIAL ACTIVITIES) | D(INVESTING ACTIVITIES) | D(OPERATIONAL ACTIVITIES) |
|--------|--------------------------|-------------------------|-------------------------|---------------------------|
| 1 | 721.3286 | -226.4929 | 470.3973 | 1373.819 |
| | (150.392) | (141.233) | (136.649) | (93.9123) |
| 2 | -832.6242 | 340.2930 | -235.0604 | -71.17374 |
| | (189.381) | (182.608) | (174.004) | (177.963) |
| 3 | 42.87032 | 306.5980 | 155.5651 | -341.1764 |
| | (206.734) | (205.562) | (205.748) | (202.845) |
| 4 | -24.01338 | -499.4195 | -107.3987 | -278.4853 |
| | (212.788) | (209.133) | (192.826) | (217.033) |
| 5 | -234.5903 | -265.0160 | 83.12071 | -164.3941 |
| | (218.630) | (211.758) | (145.528) | (226.157) |
| 6 | 198.1058 | -22.04405 | -311.1000 | 287.5668 |
| | (221.606) | (211.201) | (148.224) | (234.811) |
| 7 | 241.8573 | 186.2475 | -139.6171 | -26.57571 |
| | (221.090) | (209.659) | (138.017) | (243.331) |
| 8 | 236.7618 | 291.3367 | 8.846908 | -120.1115 |
| | (220.739) | (208.306) | (139.841) | (207.571) |
| 9 | -220.7173 | 43.25884 | 312.2315 | -2.796193 |
| | (159.639) | (179.017) | (118.895) | (203.625) |
| 10 | -210.8424 | -225.1515 | -54.15937 | -33.54819 |
| | (145.215) | (175.564) | (119.570) | (184.544) |

Cholesky Ordering: D(CASH&CASH EQUIVALENTS)
D(FINANCIAL ACTIVITIES) D(OPERATIONAL ACTIVITIES)
D(INVESTING ACTIVITIES)

Standard Errors: Asymptotic Analytic

Në Tabela 5.4.12 paraqitet Analiza Asimptotike (Përafrimi Edgeworth) për identifikimin e përgjigjes së një serie të caktuar, pas një ndryshimi të papritur tek seritë influencuese. Për të realizuar parashikimin, janë marrë në konsideratë 10 periudha të ardhme, nga tremujori i parë i vitit 2021 deri tek tremujori i dytë i vitit 2023.

Pra, nëse seria e Aktivitetit Financues pëson një ndryshim të papritur në të ardhmen, seria e Parave dhe Ekuivalentëve të parave nuk do të ketë ndryshim përgjatë tremujorit të parë dhe të dytë të vitit 2021, ndërkohë që në tremujorin e tretë do të ketë rënie të sasisë së gjendjes së parave në fund, dhe piku më i lartë arrihet tek tremujori i tretë i vitit 2022, në vlerën 555\$ më tepër. Nga ana tjetër, nëse Aktiviteti i Shfrytëzimit pëson ndryshim të papritur lidhur me paratë që do të nevojiten për t'u angazhuar në këtë lloj aktiviteti, gjendja e parave në fund do të jetë statike përgjatë tre tremujorëve të parë të vitit 2021, dhe vetëm në tremujorin e katërt do të ketë rritje prej

331 \$. Ndërkohë që rënia më e madhe e gjendjes së parave në fund do të evidentohet gjatë tremujorit të dytë të vitit 2022 në masën 183\$, dhe në fund të 2022, gjendja e parave në fund do të njohë një rritje prej 144\$. Përgjatë 2023, trendi i gjendjes së parave ka tendencën për të qenë i qendrueshëm duke ruajtur përafërsisht gjendjen si në fund të vitit 2022. Nëse seria e Aktivitetit të Investimit pëson ndryshim të papritur, gjendja në fund e parave do të pësojë rritje përgjatë: tremujorit të tretë të vitit 2021 dhe 2022, tremujorit të parë të 2022 dhe 2023 dhe tremujorit të katërt të 2022.

Nga ana tjetër nëse gjendja e parave në fund të vitit 2020 do të pësojë një ndryshim të papritur, paratë e angazhuara tek Aktiviteti Financiar do të njohin një rënie të konsiderueshme gjatë tremujorit të dytë në masën 351\$ dhe rritja më e lartë arrihet përgjatë tremujorit të dytë të 2022. Ndërkohë tendenca e parave të angazhuara pas qershorit të 2022 do të njohë një tendencë rënëse të qendrueshme me afërsisht 100\$. Ndërkohë një ndryshim i papritur në nevojën e angazhimit të parave tek Aktiviteti i Shfrytëzimit do të shkaktojë rritje të parave të Aktivitetit Financues përgjatë tremujorit të dytë të 2021 dhe tremujorit të parë të 2022 dhe 2023. Piku arrihet tek tremujori i parë i 2023 në masën 282\$. Rënia më e lartë arrihet në fund të 2022, në masën 182\$. Nëse paratë e Investuara do të njohin një ndryshim të papritur me tendencë rritëse, paratë e angazhuara tek Aktiviteti i Financimit do të ketë përgjithësisht tendencë rënëse të qendrueshme dhe vetëm gjatë tremujorit të parë të 2022 do të njohë rritje prej 154\$.

Nëse do të ndodhë një ndryshim i papritur i gjendjes së parave në fund, Aktiviteti i Shfrytëzimit do të njohë rënie të sasisë së parave në qarkullim për aktivitete operacionale përgjatë: tremujorit të dytë të 2021 në masën 315\$, tremujorit të parë të 2022 me 75\$ dhe tremujorit të tretë të 2022 dhe tremujorit të parë të 2023 me 49\$. Tremujori i dytë i 2022 njeh rritje prej 88\$. Nëse Aktiviteti i Financimit pëson ndryshim të papritur, sasia e parave të Aktivitetit të shfrytëzimit do të njohë një rënie të menjëhershme në tremujorin e parë të 2021 në masën 98\$, rënia reduktohet në nivelin 22 \$ në tremujorin e dytë të 2021, duke arritur pikun në tremujorin e tretë të 2021 me 113\$. Pas kësaj pëson rënie në fund të 2021 në vlerën 96\$, Përgjatë 2022 do të ketë rritje vetëm gjatë tremujorit të dytë në vlerën 62\$ dhe më pas një rënie prej 96\$. Përgjatë 2023 do të njohë rritje prej 115\$ gjatë tremujorit të parë dhe një rënie prej 57\$ në tremujorin e dytë. Pas ndryshimit të papritur në gjendjen e parave të angazhuara në Aktivitetin e Investimit, paratë e Aktivitetit të Shfrytëzimit do rriten përgjatë: tremujorit të tretë të 2021 me 153\$, tremujorit të parë dhe të tretë të 2022 me 2\$ dhe 117\$ dhe përgjatë tremujorit të dytë të 2023 me 42\$. Rënia e parave të angazhuara në Aktivitetin e Shfrytëzimit do të vërrehet përgjatë: fundit të 2021 me 47\$, tremujorit të dytë të 2022 me 65\$, fundit të 2022 me 35\$ si dhe gjatë tremujorit të parë të 2023 me 147\$.

Kur gjendja e parave do të ndryshojë rastësisht, gjendja e parave të destinuara për t'u investuar do të rritet përgjatë tremujorit të parë të 2021 me 721\$, tremujorit të tretë të 2021 me 43\$, tremujorit të dytë të 2022 me 198\$ dhe tremujorit të katërt të 2022 me 237\$. Rënia më e lartë parashikohet të ndodhë gjatë tremujorit të dytë të 2021 me 833\$. Nëse gjendja e parave të ardhura nga financimi do të ndryshojë në mënyrë rastësore, piku më i lartë i rritjes së parave për t'u investuar do të arrihet përgjatë tremujorit të dytë të 2021 me 340\$, tremujorit të tretë të 2021 me 307\$ dhe fundit të 2022 me 291\$, ndërkohë që rënia më e lartë parashikohet në fund të 2021 me 499\$. Nëse gjendja e parave për Aktivitetet e Shfrytëzimit ndryshon në mënyrë rastësore, gjendja e parave për investim do njohë rritje përgjatë: tremujoreve të para të 2021,

2022 dhe 2023, përkatësisht 470\$, 83\$ dhe 312\$, gjatë tremujorit të tretë të 2021 me 156\$ dhe në fund të 2022 me 9\$. Nga ana tjetër gjendja e parave për investim do të ketë rënie përgjatë: temujoreve të dyta të 2021, 2022 dhe 2023, përkatësisht 235\$, 311\$ dhe 54\$, dhe në fund të 2021 me 107\$.

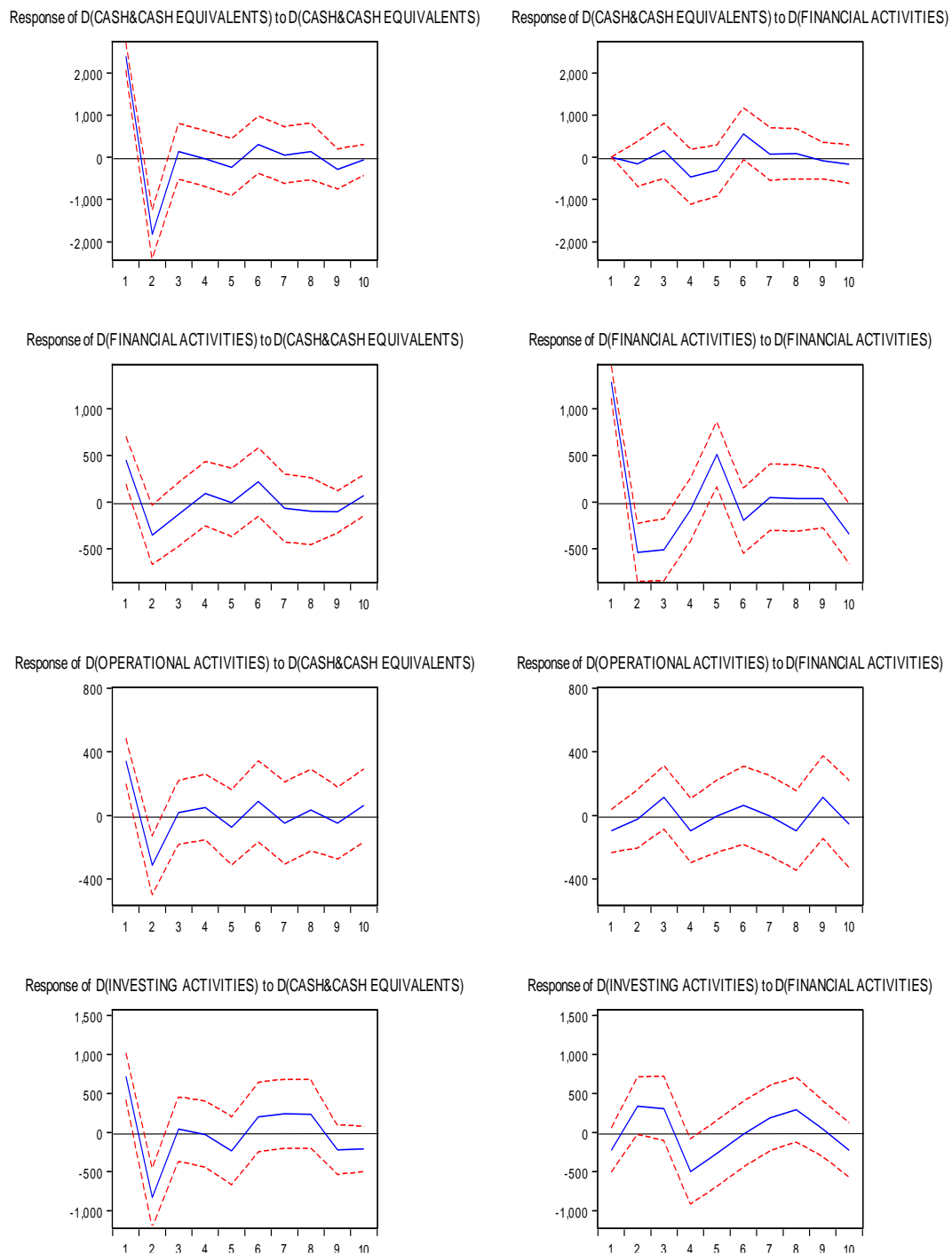
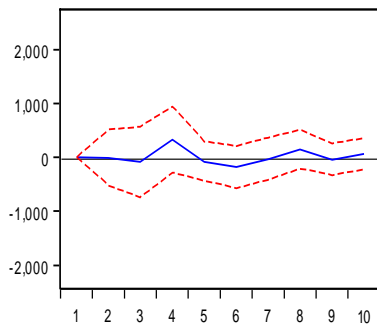
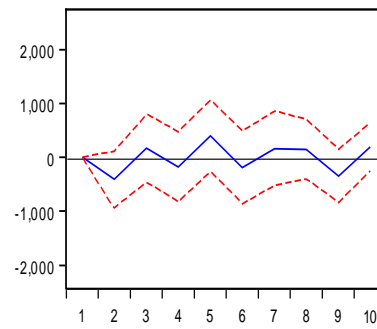


Figura 5.4.10 (a) Paraqitja grafike e Analizës Asimptotike

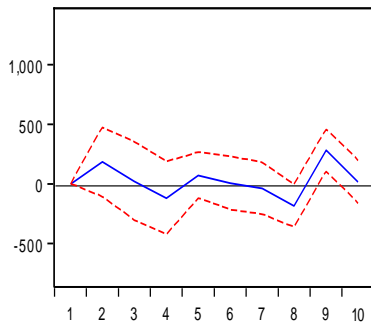
Response of D(CASH&CASH EQUIVALENTS) to D(OPERATIONAL ACTIVITIES)



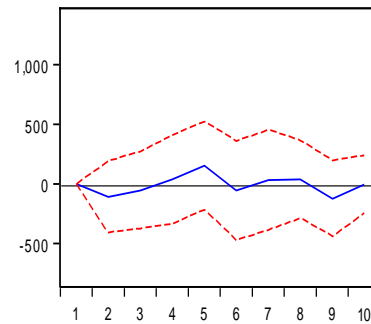
Response of D(CASH&CASH EQUIVALENTS) to D(INVESTING ACTIVITIES)



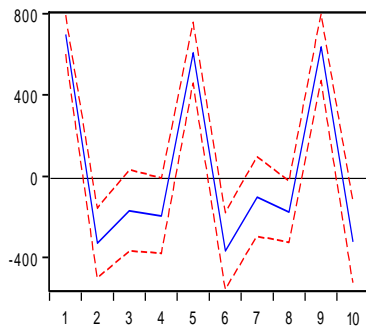
Response of D(FINANCIALACTIVITIES) to D(OPERATIONAL ACTIVITIES)



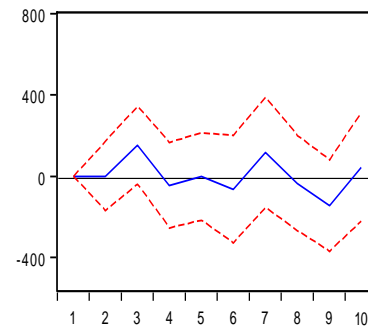
Response of D(FINANCIALACTIVITIES) to D(INVESTING ACTIVITIES)



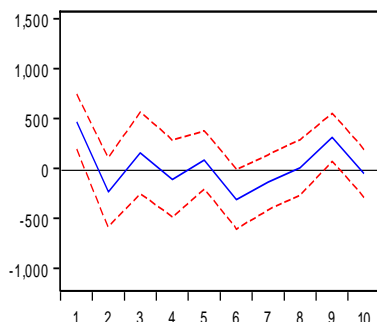
Response of D(OPERATIONALACTIVITIES) to D(OPERATIONAL ACTIVITIES)



Response of D(OPERATIONALACTIVITIES) to D(INVESTING ACTIVITIES)



Response of D(INVESTING ACTIVITIES) to D(OPERATIONAL ACTIVITIES)



Response of D(INVESTING ACTIVITIES) to D(INVESTING ACTIVITIES)

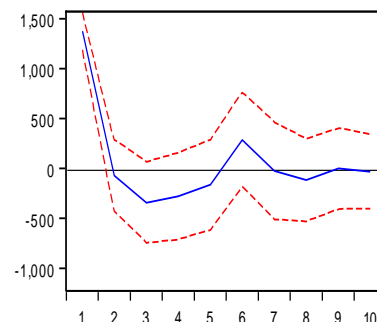


Figura 5.4.10 (b) Paraqitja grafike e Analizës Asimptotike

Figura 5.4.10 (a) dhe Figura 5.4.10 (b) paraqesin grafikisht rezultatet e Tabelës 5.4.12, të cilat i interpretuam me detaje.

Tabela 5.4.13 Rezultatet e Metodës Monte Carlo, për 1000 rizgjedhje për identifikimin e përgjigjeve përkatësisht për çdo seri

| Response of D(CASH&CASH EQUIVALENTS): | | | | |
|--------------------------------------------------|-------------------------------------|---------------------------------|--------------------------------|----------------------------------|
| Period | D(CASH&CASH EQUIVALENTS) | D)(FINANCIAL ACTIVITIES) | D(INVESTING ACTIVITIES) | D(OPERATIONAL ACTIVITIES) |
| 1 | 2394.841 (170.596) | 0.000000 (0.000000) | 0.000000 (0.000000) | 0.000000 (0.000000) |
| 2 | 568.2679 (299.914) | -157.7709 (275.133) | -10.38674 (275.943) | -411.8117 (271.467) |
| 3 | 702.6332 (300.923) | -3.894493 (304.405) | -97.59456 (285.315) | -246.6277 (300.308) |
| 4 | 672.7592 (310.196) | -470.8568 (322.099) | 233.0849 (302.059) | -426.3975 (307.083) |
| 5 | 439.4490 (329.784) | -785.6302 (349.972) | 156.9372 (272.865) | -27.38035 (332.738) |
| 6 | 736.7667 (350.220) | -230.9969 (354.627) | -26.52665 (271.063) | -218.5284 (335.095) |
| 7 | 786.1820 (334.468) | -156.0496 (368.168) | -54.90778 (241.888) | -56.70502 (364.191) |
| 8 | 921.8115 (285.626) | -73.17848 (296.676) | 89.36438 (226.182) | 84.60340 (296.107) |
| 9 | 637.1619 (291.697) | -151.2552 (323.543) | 47.83586 (231.283) | -264.6837 (268.044) |
| 10 | 574.3262 (273.179) | -317.4087 (332.872) | 108.8015 (229.498) | -67.36506 (277.699) |

| Response of D(Financial Activities): | | | | |
|---------------------------------------------|-------------------------------------|---------------------------------|--------------------------------|----------------------------------|
| Period | D(CASH&CASH EQUIVALENTS) | D)(FINANCIAL ACTIVITIES) | D(INVESTING ACTIVITIES) | D(OPERATIONAL ACTIVITIES) |
| 1 | 452.1291 (134.964) | 1288.179 (87.7315) | 0.000000 (0.000000) | 0.000000 (0.000000) |
| 2 | 101.0596 (175.728) | 751.0235 (170.563) | 183.6154 (153.147) | -106.5981 (153.740) |
| 3 | -29.63566 (184.872) | 243.0635 (194.814) | 207.2283 (181.139) | -157.3587 (197.890) |
| 4 | 63.32674 (189.137) | 163.3950 (195.861) | 88.68121 (187.774) | -120.0178 (198.129) |
| 5 | 58.56586 (205.685) | 675.2093 (218.795) | 162.9359 (180.737) | 33.48376 (200.104) |
| 6 | 276.3703 (203.677) | 479.7182 (239.033) | 167.6653 (166.719) | -20.76574 (213.051) |
| 7 | 214.1503 (198.085) | 531.6361 (227.604) | 130.6273 (141.865) | 11.33875 (235.374) |
| 8 | 116.3856 (166.907) | 573.4527 (211.207) | -51.78321 (151.198) | 49.72144 (206.345) |
| 9 | 13.69447 (178.023) | 610.7718 (223.638) | 230.2707 (153.458) | -71.78050 (173.802) |
| 10 | 88.03166 (181.130) | 271.7297 (240.313) | 247.6626 (170.350) | -73.57136 (180.248) |

| Response of D(Operational Activities): | | | | |
|-----------------------------------------------|-------------------------------------|---------------------------------|--------------------------------|----------------------------------|
| Period | D(CASH&CASH EQUIVALENTS) | D)(FINANCIAL ACTIVITIES) | D(INVESTING ACTIVITIES) | D(OPERATIONAL ACTIVITIES) |
| 1 | 344.1629 (71.5141) | -98.12310 (69.0335) | 699.6536 (47.3397) | 0.000000 (0.000000) |

| | | | | |
|----|-----------------------|------------------------|------------------------|------------------------|
| 2 | 29.13426 (98.5558) | -120.5300 (96.7324) | 369.9468 (89.5173) | -0.254375 (89.0068) |
| 3 | 45.63403 (97.1030) | -7.668934 (103.718) | 200.5067 (101.969) | 152.8427 (105.068) |
| 4 | 96.66188 (102.578) | -103.9431 (113.316) | 4.262022 (109.148) | 106.2483 (110.518) |
| 5 | 21.22117 (125.202) | -109.7512 (135.273) | 612.3433 (116.433) | 103.8137 (108.598) |
| 6 | 109.1326 (133.020) | -48.00116 (141.208) | 241.8991 (114.388) | 38.53418 (130.332) |
| 7 | 60.43096 (116.985) | -54.00299 (133.104) | 139.4091 (101.312) | 155.8878 (145.947) |
| 8 | 93.14285 (98.3919) | -150.4174 (122.388) | -38.36987 (100.969) | 120.5862 (121.972) |
| 9 | 44.00179 (114.155) | -35.60591 (136.234) | 598.9447 (104.093) | -25.97544 (107.473) |
| 10 | 105.7675 (120.855) | -92.95895 (143.787) | 275.4370 (109.605) | 15.85959 (120.295) |

**Response of
D(Investing
Activities):**

| Period | D(CASH&CASH EQUIVALENTS) | D(FINANCIAL ACTIVITIES) | D(INVESTING ACTIVITIES) | D(OPERATIONAL ACTIVITIES) |
|--------|-----------------------------|----------------------------|----------------------------|------------------------------|
| 1 | 721.3286 (147.927) | -226.4929 (148.659) | 470.3973 (141.400) | 1373.819 (96.2577) |
| 2 | -111.2956 (224.226) | 113.8001 (235.185) | 235.3369 (225.861) | 1302.645 (205.996) |
| 3 | -68.42523 (254.162) | 420.3981 (264.313) | 390.9020 (260.061) | 961.4687 (264.973) |
| 4 | -92.43861 (273.604) | -79.02140 (284.494) | 283.5033 (290.718) | 682.9834 (286.524) |
| 5 | -327.0289 (290.705) | -344.0374 (324.814) | 366.6240 (301.664) | 518.5893 (302.668) |
| 6 | -128.9231 (298.297) | -366.0815 (337.033) | 55.52404 (305.522) | 806.1561 (316.404) |
| 7 | 112.9342 (284.857) | -179.8340 (355.210) | -84.09303 (276.658) | 779.5804 (348.108) |
| 8 | 349.6960 (224.383) | 111.5027 (307.111) | -75.24612 (244.770) | 659.4689 (331.971) |
| 9 | 128.9787 (221.602) | 154.7615 (293.891) | 236.9854 (246.248) | 656.6727 (291.401) |
| 10 | -81.86372 (237.247) | -70.38999 (299.921) | 182.8260 (239.909) | 623.1245 (264.664) |

Cholesky Ordering: D(CASH_CASH_EQUIVALENTS)
D(FINANCIAL_ACTIVITIES) D(OPERATIONAL_ACTIVITIES)
D(INVESTING_ACTIVITIES)

Standard Errors: Monte Carlo (1000 repetitions)

Tabela 5.4.13 paraqet rezultatet e simulimeve Monte Carlo në termat e përgjigjes së një serie kohore pas ndryshimit të papritur në serinë tjetër influencuese. Këto rezultate nuk kanë diferenca të konsiderueshme nëse krahasohen me rezultatet e analizës asimptotike, por një diferencë të vogël saktësie bën Metoda e Monte Carlos, për shkak të zbatimit të një numri të madh rizgjedhjesh.

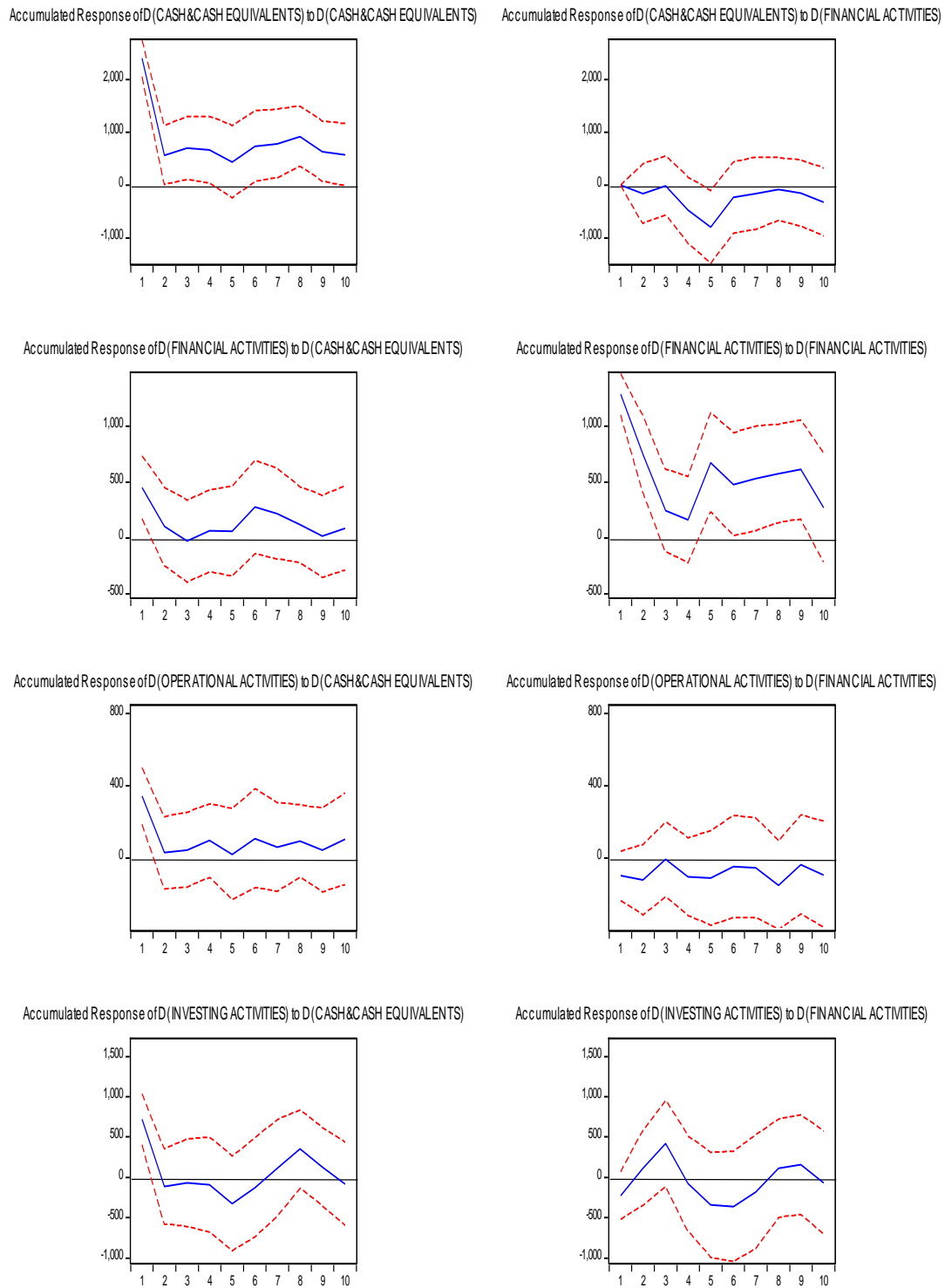


Figura 5.4.11 (a) Paraqitja grafike e simulimeve Monte Carlo

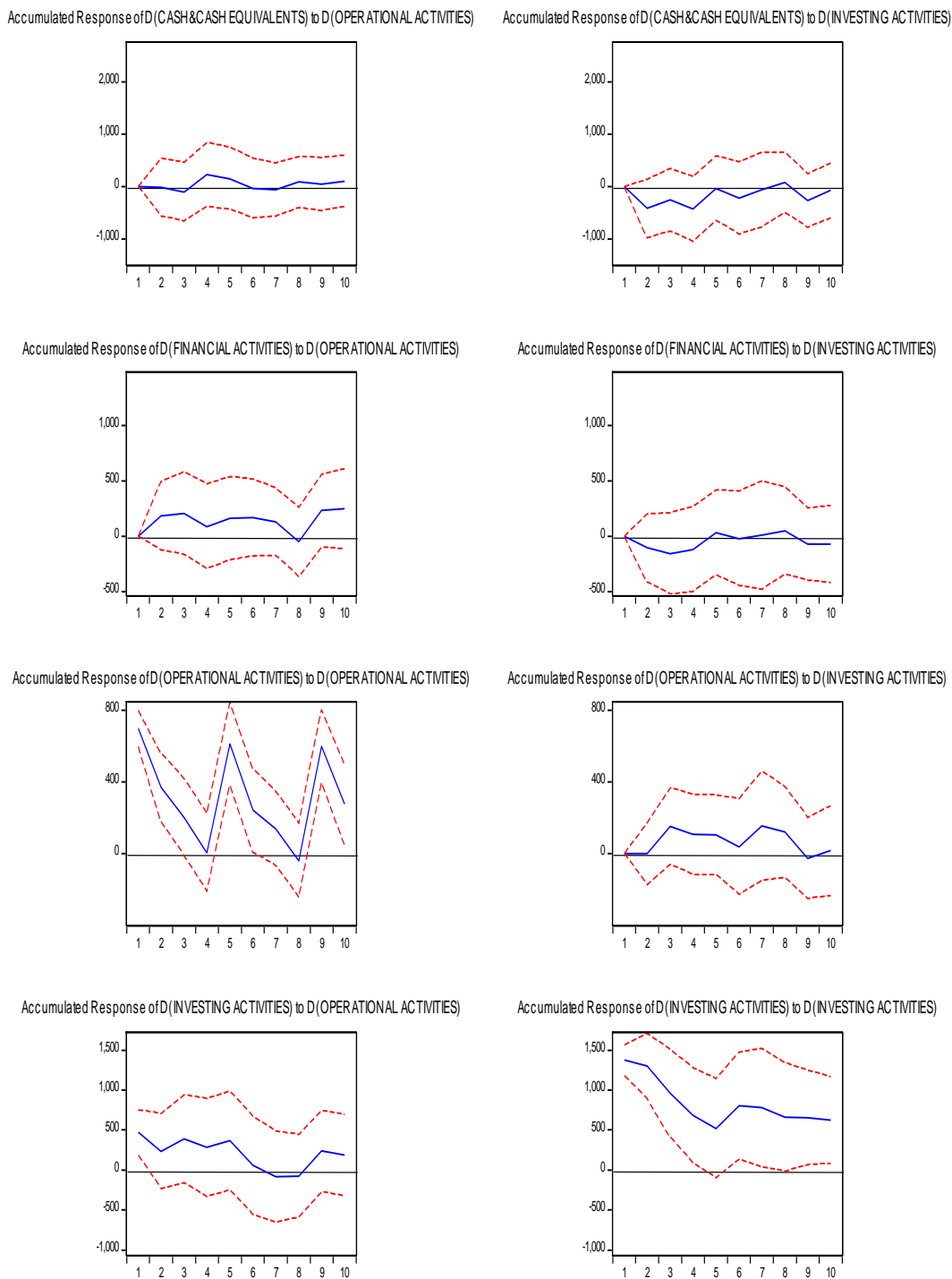


Figura 5.4.11 (b) Paraqitja grafike e simulimeve Monte Carlo

Figurat 5.4.11 (a) dhe 5.4.11 (b) paraqesin grafikisht rezultatet e gjeneruara në Tabelën 5.4.13. Pra, sërisht evidentohet tendenca përafërsisht e njëjtë e përgjigjes së serive kohore të marra në studim pas ndryshimit imediat të serive të tjera influencuese, sikurse në rastin e analizës asimptotike në termat e përafërimit Edgeworth.

Response to Cholesky One S.D. (d.f. adjusted) Innovations

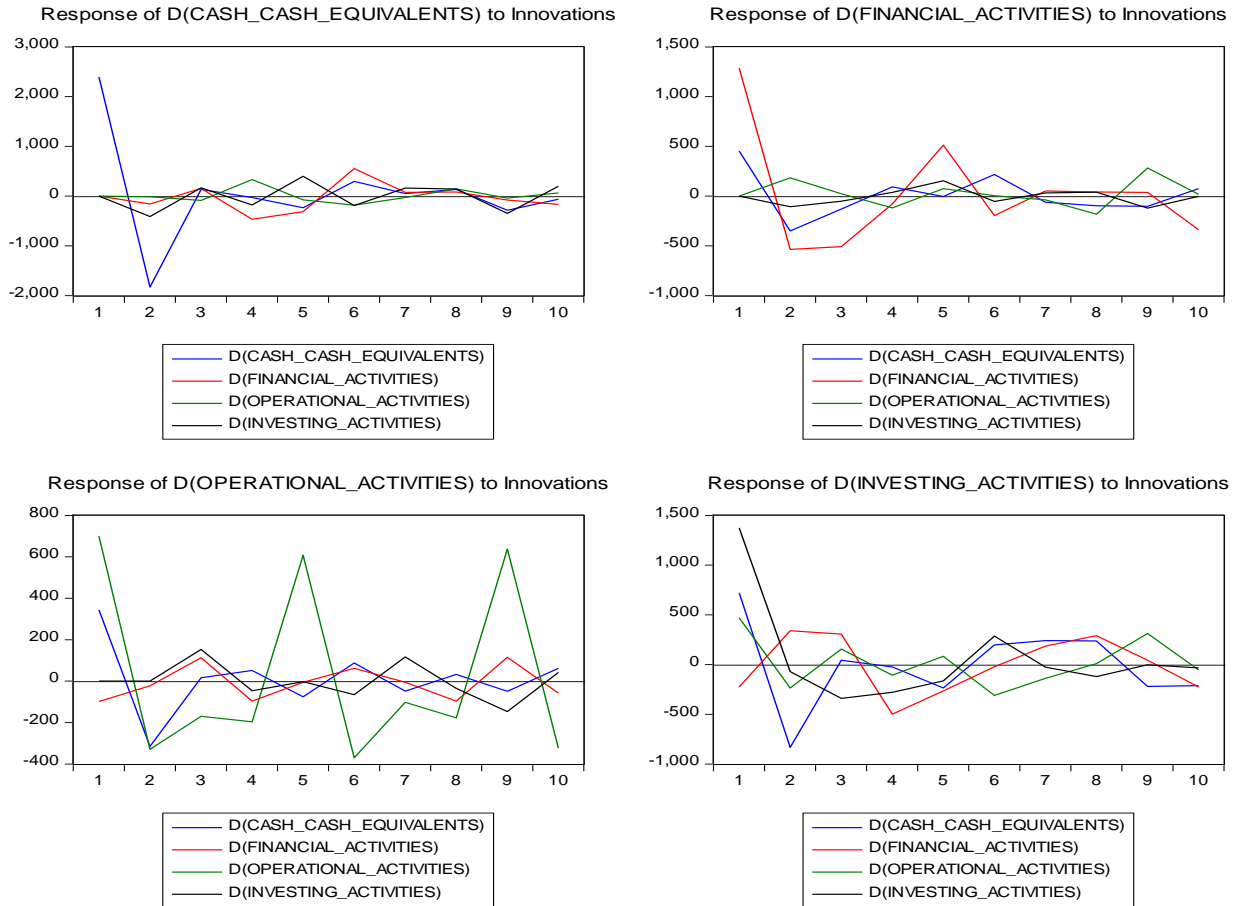


Figura 5.4.12 Rezultate përmblendhëse të Përafritit Edgeworth dhe Simulimeve Monte Carlo

Figura 5.4.12 tregon mbivendosjen e rezultateve përafërsisht të njëjta të analizës asimptotike dhe simulimeve Monte Carlo për përgjigjen e serive kohore të marra në studim për çdo rritje me një njësi në devijimet standarde ose ndryshimeve të rastësishme të serive kohore, të cilat kanë ndikim të konsiderueshëm në seritë që po studiojmë.

Tabela 5.4.14 Versioni Bootstrap në Modeli VAR

```
Model: Bootstrap in VAR model
Sample (adjusted): 8 114
Solve Options:
  Static-Stochastic Simulation
  Solver: Broyden
  Max iterations = 5000, Convergence = 1e-08
  Requested repetitions = 1000, Allow up to 5 percent failures
  Solution does not account for coefficient uncertainty in linked equations
  Track endogenous: mean, standard deviation, 95% confidence interval

Calculating Bootstrap Innovations
Sample: @ALL
107 residuals per equation calculated for bootstrap

Scenarios: Baseline & Baseline
Solve begin 14:42:02
  Repetitions 1-200: successful 14:42:02
  Repetitions 201-400: successful 14:42:03
  Repetitions 401-600: successful 14:42:03
  Repetitions 601-800: successful 14:42:03
  Repetitions 801-1000: successful 14:42:03
Solve complete 14:42:03
1000 successful repetitions, 0 failure(s)
```

Në Tabela 5.4.14 janë paraqitur rezultatet pas aplikimit të Versionit Bootstrap të Modelit Baseline, i cili përdoret për studimin e marrëdhënieve Stokastike, në termat e konsiderimit të gabimeve e realizuara gjatë vlerësimit të modelit VAR. Për modelin Bootstrap janë përdorur 1000 rizgjedhje, nivel 95% të intervalit të besimi, 5% nivel gabimi i lejuar.

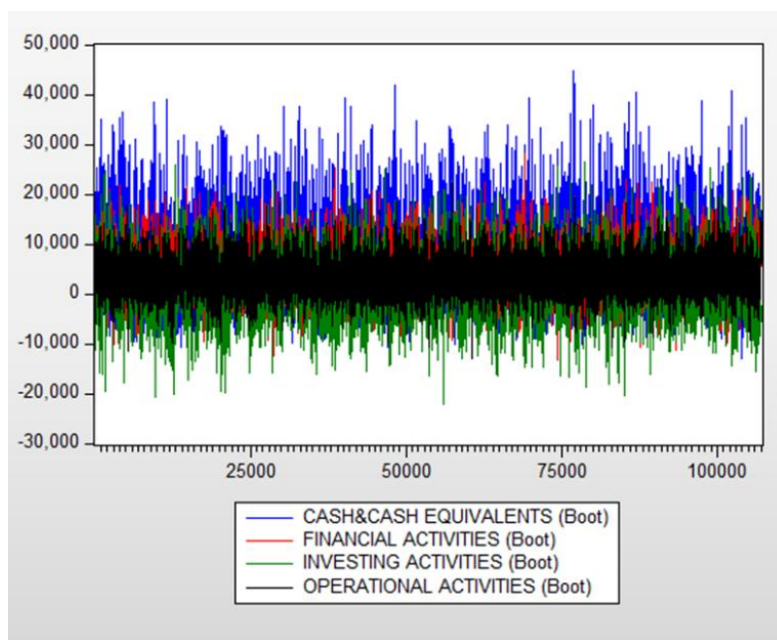


Figura 5.4.13 Mbivendosje e rezultateve të përgjigjeve të çdo serie

Në Figurën 5.4.13, paraqitet vizualizimi i mbivendosjes së rezultateve grafike të përgjigjes së serive ndaj ndryshimit rastësor në njësinë e gabimit të serive influencuese. Kjo paraqitje përputhet plotësisht me grafikun që do të gjenerohej nëse

të gjithë grafikët përbërës të simulimeve të metodave Edgeworth dhe Monte Carlo do të paraqiteshin në një grafik të vetëm.

Pra, konkluzioni i këtij studimi konsiston në faktin se Përafrimi Edgeworth, Metoda Bootstrap dhe simulimet Monte Carlo janë metodat më të mira, të cilat optimalizojnë modelin shumëpërmasor të krijuar. Konkretisht, modeli VAR që ne ndërtuam me të dhënat për Paratë (Ekuivalentët e Parave), Aktivitetet Financiare, Aktivitetet Investuese dhe Aktivitetet e Shfrytëzimit të biznesit Coca- Cola, u studiu në termat e prognozës së saktë kur Përafrimi Edgeworth, Metoda Bootstrap dhe simulimet Monte Carlo japin përafërsisht të njëjtin rezultat.

PËRFUNDIME

1. Nga simulimet që kemi kryer vumë re se përafrimi Edgeworth është ndjeshëm më i saktë se përafrimi normal. Nëse numri i simulimeve dhe vëllimit të zgjedhjes rritet, tendenca e përafrimit Edgeworth ka ndjeshëm më tepër epërsi se përafrimi normal.
2. Performancat asimptotike të polinomeve të rendit të parë dhe të dytë, për shpërndarjet e rëndësishme që morëm në shqyrtim, janë përafërsisht të njëjta dhe nuk ka diferenca të konsiderueshme për të dyja rastet me implementimin e versionit bootstrap dhe pa përdorur versionin bootstrap.
3. Intervali i Besimit është një ndër aspektet statistikore i cili lidhet me saktësinë e modelit të marrë në shqyrtim. Saktësia e modelit rritet kur përqindja e mbulimit të intervalit rritet dhe diapazoni i intervalit është i ngushtë. Versioni Bootstrap kontribuon në rritjen e probabilitetit të mbulimit duke rritur nivelin e besueshmërisë së rastit që morëm në studim.
4. Zbatimi i versionit Bootstrap kontribuon tek saktësia e paraqitjes grafike së shpërndarjeve kryesisht me pjerrësi pozitive. Implementimi i metodave Bootstrap rritin korrektësinë e rendit të dytë të shpërndarjeve me pjerrësi pozitive.
5. Kontributi i përafrimit Edgeworth në metodat e rizgjedhjes Bootstrap vërrehet në zbatime ekonometrike. Një ndër rastet ekonometrike që morëm në studim është marrëdhënia regressive, ku lehtësisht evidentohet se përafrimi Edgeworth, metodat e rizgjedhjes Bootstrap dhe përafrimi Edgeworth i versionit Bootstrap reduktojnë devijimet standarde. Megjithatë gjatë analizës krahasimore të kontributit të metodave statistikore në reduktimin e termit të gabimit, vërrehet se përafrimi Edgeworth i versionit Bootstrap është metoda më e mirë statistikore e cila optimizon modelin e regresit të krijuar.
6. Implementimi i zbërthimit Edgeworth, metodave të rizgjedhjes Bootstrap dhe simulimeve Monte Carlo gjejnë konkretizim praktik në prognozimin e serive kohore. Analiza e serive kohore që morëm në studim është realizuar nëpërmjet ndërtimit të Vektorit të Autoregresit (Modeli VAR). Nga simulimet e këtij rasti studimor konkluduar që meqënëse rezultatet e Përafrimit Edgeworth, gjenerimeve Bootstrap dhe simulimeve Monte Carlo përafrohen, parashikimi i Modelit VAR për seritë kohore, është kryer saktë.

BIBLIOGRAFIA

- [1] Abramowitz, M. and Stegun, I.A. (1972). Handbook of Mathematical Functions, 9th ed. Dover, New York.
- [2] Anderson, D., Sweeney, D., Williams, T. (2013), Statistics for Business and Economics, United States, Cengage Learning, Inc.
- [3] Baltagi, B. (2011), Econometrics, Berlin, Springer.
- [4] Basna, R. (2010): Edgeworth Expansion and Saddlepoint Approximation for Discrete Data with Applications in the Chance Games. Linnéuniversitetet.
- [5] Bassett, G. J., Koenker, R., (1982). An Empirical Quantile Function for Linear Models with i.i.d. Errors, Journal of the American Statistical Association, 77 (378), 407-415.
- [6] Bhattacharya, R. N., Ghosh, J. K., (1978). On The Validity of The Formal Edgeworth Expansion, The Annals of Statistics, Vol.6, No.2, 434-451.
- [7] Bickel, P.J. and Freedman, D.A. (1981). Some asymptotic theory for the bootstrap. The Annals of Statistics. Vol.9, No.6.
- [8] Bloch, D. A., Gastwirth, J. L., (1968). On a simple estimate of the reciprocal of the density function, Ann. Math. Statist., 39, 1083-1085.
- [9] Bose, A., (1988). Edgeworth Correction by Bootstrap in Autoregressions, The Annals of Statistics, Vol.16, No.4, 1709-1722.
- [10] Buchinsky, M., (1995). Estimating the Asymptotic Covariance Matrix for Quantile Regression Models: A Monte Carlo Study, Journal of Econometrics, 68, 303-338.
- [11] Butler, R.W. (2007). Saddlepoint Approximation with Applications. Cambridge University Press, New York, USA, 145.
- [12] Casella, G. and Berger, R. L. (2002): Statistical Inference, second edition. The Wadsworth Group, USA. 83-87.
- [13] Chebyshev, P.L. (1890). Sur deux theorems relatifs aux probabilités. Acto Math. 14, 305-315.
- [14] Chernick, M.R., LaBudde, R.A. (2011). An introduction to bootstrap methods with applications to R, first edition. Published by John Wiley & Sons, Inc.
- [15] Cornish, E.A. and Fisher, R.A. (1937). Moments and cumulants in the specification of distributions. Internat. Statist. Rev. 5, 307-322.
- [16] Cramer, H. (1928). On the composition of elementary errors. Skand. Aktuarietidskr. 11, 13-74, 141-180.
- [17] Cramer, H. (1946). Mathematical Methods of Statistics. Princeton University Press, Princeton, NJ.
- [18] Cramér, H. (1970). Random Variables and Probability Distributions, 3rd ed. Cambridge University Press, Cambridge, UK.

- [19] David H. A., (1981). Order Statistics, 2nd ed. Wiley, New York.
- [20] Davison A. C. and Hinkley D. V. (1997). Bootstrap Methods and their Applications. Cambridge, England: Cambridge University Press.
- [21] DiCiccio T. J., & Efron B. (1996). Bootstrap Confidence Intervals (with discussion). Statistical Science, 11(3), 189-228.
- [22] Donefski, E., Donefski. T., Milo E. (2021). The contribution of Edgeworth approximation in a real study case. Applied Mathematical Sciences, Vol. 15, no.3, 119-130, ISSN 1314-7552, <https://doi.org/10.12988/ams.2021.914419>
- [23] Donefski, E. (2014). Edgeworth approximation and Cornish-Fisher expansion for estimation errors in confidence interval. SPNA, Statistics, probability & numerical analysis. Methods and applications, 5-6 December 2014, Proceeding book, p.146-150, ISBN 978-9928-4252-4-9.
- [24] Donefski, E., Braho, D. (2016). Edgeworth expansions for bootstrap confidence intervals correctness. Aktet, Journal of Institute Alb-Shkenca, Vol. 9, No. 1, 2016, pp. 40-45, ISSN 2073-2244.
- [25] Donefski, E., Ekonomi, L. (2021). Application of Edgeworth Expansions on some important distributions, European Academic Research, Volume VIII, Issue 10, ISSN 2286-4822, p 6567-6576, <http://www.euacademic.org/receivedArticle.aspx>
- [26] Donefski, E., Ekonomi.L., Braho, D. (2017). Edgeworth approximation for Chi-squared distribution, Seventh International Conference “Modern Trends in Science” 14 – 18 june , Blagoevgrad, Bulgaria.
- [27] Donefski, E., Milo E., Margo L., Braho D. (2015). Edgeworth approximation for asymmetric distribution, Third International Conference Research and Education in Natural Sciences: Harmonization of Environment Research and Teaching with Sustainable Policy, HERTSPO 2015, Shkodër, 6-8 nëntor 2015, botuar në librin e Procëdings, Volumi I, fq. 69-77, ISBN 978-9928-4135-9-8.
- [28] Donefski,E., Donefski, T., Ekonomi, L. (2021). The Contribution of Edgeworth, Bootstrap and Monte Carlo Methods in Financial Data. XV. International Conference on Computational Mathematics Statistics and Information Applications, ICCMSIA 2021, 08-09 april, Athens, Greece, IRC 2021, X. International Research Conference Proceedings, part IV, p.308-317, online ISSN 1307-6892.
- [29] Donefski,E., Ekonomi, L., Donefski, T. (2021). The profit trend of cosmetics products using bootstrap Edgeworth approximation. XV. International Conference on Economics, Mathematics and Statistics, ICEMS 2021, 12-13 april, Venice, Italy, IRC 2021, X. International Research Conference Proceedings, part V, p.429-436, online ISSN 1307-6892.
- [30] Edgerton, D., Shukur, G., (1999). Testing autocorrelation in a system perspective, Econometric Reviews, 18, 343–386.
- [31] Edgeworth, F.Y. (1896). The assymetrical probability curve. Philos. Mag., 5th Ser. 41, 90-99.

- [32] Edgeworth, F.Y. (1905). The law of error. Proc. Cambridge Philos. Soc. 20, 36-65.
- [33] Efron B., Tibshirani R.J.,(1993): An introduction to the bootstrap, Chapman and Hall, New York
- [34] Efron, B. (1979). Bootstrap methods: Another look at the jackknife. The Annals of Statistics, 7, 1-26.
- [35] Efron, B. (1980). The Jackknife, the Bootstrap, and Other Resampling Plans. National Science Foundation Grant, California,105-132.
- [36] Efron, B. (1982). The jackknife, the bootstrap, and other resampling plans. Society of Industrial and Applied Mathematics CBMS-NSF Monographs, 38.
- [37] Ekonomi L,(2003) “Kontribut për vlerësimet jackknife dhe bootstrap në vlerësimet lineare”, dezertacion.
- [38] Ekonomi L., Margo L., Milo E., Donefski E., Palla I. (2013). Bootstrap with cycled blocks in stationary time series, The Fifth International Scientific Conference (FMNS), South-West University “Neofit Rilski”, Blagoevgrad, Bulgaria, Volume 1 p.21-26.
- [39] Ekonomi, L., Margo, L., Milo, E. (2010). Jackknife dhe bootstrap për vlerësimin e UCL në lidhje me pritjen matematike në shpërndarjet me pjerësi pozitive. Buletini No.20, Universiteti i Korçës.
- [40] Engel, J. (2010). On teaching bootstrap confidence intervals. International Association of Statistical Education (IASE), www.stat.auckland.ac.nz/~iase/
- [41] Essen, C. G. (1968). On the Remainder Term in the Central Limit Theorem. Arkiv Mathematic 8, 7-15.
- [42] Eviews 10 User’s Guide I (2017).
- [43] [EViews Help: Multiple Equation Analysis](#)
- [44] Fisher, R.A. and Cornish, E.A. (1960). The percentile points of distributions having known cumulants. Technometrics 2, 209-226.
- [45] Geisser, S. (1975). The predictive sample reuse method with applications. J. Amer. Statist. Assoc., 70, pp.320-328.
- [46] Gnedenko, B.V. and Kolmogorov, A.N. (1954). Limit Distributions for Sums of Independent Random Variables. Addison-Wesley, Reading, MA.
- [47] Götze, F., Hipp, C., (1983). Asymptotic expansions for sums of weakly dependent random vectors, Z. Wahrsch. Verw.Gebiete, 64, 211-239.
- [48] Gut, A. (2005). Probability a Graduate Course. Springer.
- [49] Hall, P. (1987a). On the bootstrap and continuity correction. J. Roy. Statist. Soc. Ser. B 49, 82-89.
- [50] Hall, P. (1992). The Bootstrap and Edgeworth Expansion. Springer-Verlag, New York, USA.

- [51] Hall, P., Martin M. A., (1991). On the error incurred using the bootstrap variance estimate when constructing confidence intervals for quantiles, *J. Mult. Anal.*, 38, 70-81.
- [52] Hall, P., Sheather, S. J., (1988). On the Distribution of the Studentized Quantile, *Journal of the Royal Statistical Society, Series B*, 50 (3), 381-391.
- [53] He, X., Hu, F., (2002). Markov Chain Marginal Bootstrap, *Journal of the American Statistical Association*, 97 (459), 783-795.
- [54] http://fbemoodle.emu.edu.tr/pluginfile.php/47097/mod_resource/content/1/Workshop%204%20Part%202%20Advanced%20Time%20Series%20Econometrics%20with%20EViews.pdf
- [55] <http://www.eviews.com/help/helpintro.html#page/content%2Fquantreg-Background.html%23>
- [56] <https://idoc.pub/documents/eviews-8-users-guide-ii-on2397q8dyl0>
- [57] Johansen, S., (1988). *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford: Oxford University Press, 1995. A. H. Welsh, Asymptotically Efficient Estimation of the Sparsity Function at a Point, *Statistics & Probability Letters*, 6, 427-432.
- [58] Johnson, N. L., Kotz, S., Balakrishnan, N. (1994). *Chi-Squared Distributions including Chi and Rayleigh*, In: *Continuous Univariate Distributions. 1 (Second ed.)*. John Wiley and Sons pp. 415-493.
- [59] Jones, M. C., (1992). Estimating Densities, Quantiles, Quantile Densities and Density Quantiles, *Annals of the Institute of Statistical Mathematics*, 44 (4), 721-727.
- [60] Kocherginsky, M., He, X., Mu, Y., (2005). Practical Confidence Intervals for Regression Quantiles, *Journal of Computational and Graphical Statistics*, 14 (1), 41-55.
- [61] Koenker, R., (2005). *Quantile Regression*, New York: Cambridge University Press.
- [62] Koenker, R., Bassett, G.J., (1978). Regression Quantiles, *Econometrica*, 46 (1), 33-50.
- [63] Koenker, R., Mandl, P., Huskova, M., (1994). Confidence Intervals for Regression Quantiles, in *Asymptotic Statistics*, eds., New York: Springer-Verlag, 349-359.
- [64] Lee, C. F. et al. (1990). *Security Analysis and Portofolio Management*. Glenview, III: Scott, Foresman/Little, Brown, 754-760.
- [65] Lee, S. M. S. and Young, G. A. (1995). Asymptotic iterated bootstrap confidence intervals. *Ann. Statist.*, 23, 1301-1330.
- [66] Lütkepohl, H., (1991). *Introduction to Multiple Time Series Analysis*, New York: Springer-Verlag.
- [67] Lütkepohl, H., (2007). *New Introduction to Multiple Time Series Analysis*, New York: Springer-Verlag.

- [68] Magnus, W., Oberhettinger, F., and Soni, R.P. (1966). *Formulas and Theorems for the Special Functions of Mathematical Physics*. Springer-Verlag, Berlin.
- [69] Mooney, C.Z. & Duval, R.D. (1993). *Bootstrapping: A nonparametric approach to statistical inference*. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-095. Newbury Park, CA: Sage.
- [70] NIST, (2006). *Engineering Statistics Handbook Chi-Squared Distribution*.
- [71] Petrov, V.V. (1975). *Sums of Independent Random Variables*. Springer-Verlag, Berlin.
- [72] Powell, J., (1986). Censored Regression Quantiles, *Journal of Econometrics*, 32, 143-155.
- [73] Quenouille, M. (1949). Approximate tests of correlation in time series. *Journal of the Royal Statistical Society, Soc. Series B*, 11, 18-84.
- [74] Schenker N. (1985). Qualms about bootstrap confidence intervals. *J. Amer. Statist. Assoc.* 80.
- [75] Shao J. (2003). *Mathematical Statistics*, revised ed, Springer: P70-76, Sec1.5.6.
- [76] Siddiqui, M. M., (1960). Distribution of Quantiles in Samples from a Bivariate Population, *Journal of Research of the National Bureau of Standards–B*, 64 (3), 145-150.
- [77] Sims, C., (1980). *Macroeconomics and Reality*, *Econometrica*, 48, 1-48.
- [78] Singh, K. (1981). On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* 9, 1187-1195.
- [79] Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *J. Roy. Statist. Soc. Ser. B*, 36, p.111-147.
- [80] Stuart, A. and Ord, K. (1994). *Distribution theory*, vol.1. Oxford University press Inc., New York, USA, 74-162.
- [81] Tukey, J. W. (1958). Bias and confidence in not quite large samples. *Annals of Mathematical Statistics*, 29, 614.
- [82] Wackerly, D. D., Mendenhall, W.III and Scheaffer, R. L. (2002): *Mathematical Statistics with Applications*. Duxbury, CA, USA.
- [83] Wahba, G. and Wold, S. (1975). A completely automatic French curve: fitting spline functions by cross-validation. *Comm. Statist.*, 4, pp.1-17.
- [84] Welsh, A. H., (1988). Asymptotically Efficient Estimation of the Sparsity Function at a Point, *Statistics & Probability Letters*, 6, 427-432.
- [85] www.uncg.edu/mat/qms/Resampling%20Methods%20Using%20R.pdf
- [86] www.cran.r-project.org/doc/contrib/Verzani-SimpleR.pdf

ABSTRAKT

Puna kërkimore e kësaj teze doktrature, bazohet në kontributin e zërthimit Edgeworth tek metodat e rizgjedhjes Bootstrap. Duke qenë se nuk është e mundur të ndërtohet një model studimor ideal, ku termi i gabimit të jetë zero, synimi i çdo studiuesi është zbatimi i një metode statistikore, për të minimizuar devijimet standarde të rastit studimor të marrë në shqyrtim. Meqë zgjedhja e teknikave statistikore që do të mundësojnë reduktimin e gabimeve është një sfidë në procesin e kërkimit shkencor, punimi i këtij disertacioni mëshon në gjetjen e metodave të cilat lehtësojnë këtë proces. Për të evidentuar se cila prej metodave të: përafrimit Edgeworth, metodave të rizgjedhjes Bootstrap dhe përafrimit Edgeworth të versionit Bootstrap është optimale në gjetjen e zgjidhjes më të mirë të problemit kërkimor, kemi shfrytëzuar burime të gjera literaturash dhe kemi ndërmarrë studime praktike, për të konkretizuar bazën teorike të përvehtësuar. Gjuhët e programimit të cilat i kemi përdorur në gjenerimin e simulimeve konkrete të punës janë: Programi R dhe Eviews 10. Logjika induktive që kemi ndjekur përgjatë kësaj teze, fillon me një këndvështrim të kombinuar teorik dhe praktik mbi shkallën e rëndësisë së mëvetësishme të zërthimit Edgeworth dhe metodave të rizgjedhjes Bootstrap. Punimin e kemi shtjelluar në vijim me identifikimin e nivelit të rëndësisë së ndërthurjes së zërthimit Edgeworth dhe Metodave rizgjedhëse Bootstrap, ku për të cilin kemi kryer disa simulime konkretizuese. Vijimësia e punës kërkimore finalizohet me fushat ekonometrike ku metodat statistikore që kemi marre në shqyrtim aplikohen dhe kontribuojnë në optimizimin e modeleve me natyrë ekonomike dhe financiare. Për të evidentuar metodat më të mira statistikore kemi ndërtuar modele studimore praktike, të cilat bazohen në të dhëna reale të grumbulluara prej bizneseve me sukses kombëtar dhe ndërkombëtar. Marrëdhëniet e regresit dhe Vektori i Autoregresionit janë fushat ekonometrike ku kemi testuar saktësinë e metodave statistikore të marra në shqyrtim.

Fjalët kyç: Zërthimi Edgeworth, Metoda e rizgjedhjes Bootstrap, Përafrimi Edgeworth i versionit Bootstrap

ABSTRACT

The research work of this dissertation is based on the contribution of Edgeworth expansion in Bootstrap resampling methods. Since it is not possible to build an ideal study model, where the error term is zero, each researcher's aim is applying a statistical method, that can minimize the standard deviations of the study case that is under consideration. Since the choice of statistical techniques that will enable the reduction of errors is a challenge in the research process, the work of this dissertation consists in finding methods that facilitate this process. To identify which of the methods like: Edgeworth approximation, Bootstrap resampling methods, or Edgeworth approximation of the Bootstrap version, is optimal in finding the best solution to the research problem, we have used a variety of literatures and we have done some practical studies undertaken for concretizing the acquired theoretical basis. The programming languages that we have used in generating simulations are R Programming and Eviews 10. The inductive logic that we have followed throughout this thesis begins with a combination of theoretical and practical perspectives on the level of importance of Edgeworth expansion and Bootstrap resampling methods. Our following work consists in identifying the importance level of Edgeworth expansion and Bootstrap resampling methods, for which some concretizing simulations have been generated by us. The continuation of this research work is finalized with some econometric fields where we have applied the considered statistical methods that lead us to the optimization of economic models. We have built concrete study models for identifying the best statistical methods, which are based on real data collected from businesses with national and international success. Regression relationships and the Vector of Autoregression are the econometric fields where we have tested the accuracy of considered statistical methods.

Key Words: Edgeworth Expansion, Bootstrap resampling methods, Edgeworth approximation of the Bootstrap version