



УНИВЕРЗИТЕТ У НИШУ  
ПРИРОДНО-МАТЕМАТИЧКИ ФАКУЛТЕТ



**Петра Н. Лакета**

**НЕНЕГАТИВНИ ЦЕЛОБРОЈНИ  
АУТОРЕГРЕСИВНИ ПРОЦЕСИ У  
СЛУЧАЈНОЈ СРЕДИНИ ГЕНЕРИСАНИ  
ГЕОМЕТРИЈСКИМ БРОЈАЧКИМ  
НИЗОВИМА**

**ДОКТОРСКА ДИСЕРТАЦИЈА**

Текст ове докторске дисертације ставља се на увид јавности,  
у складу са чланом 30., став 8. Закона о високом образовању  
("Сл. гласник РС", бр. 76/2005, 100/2007 – аутентично тумачење, 97/2008, 44/2010,  
93/2012, 89/2013 и 99/2014)

**НАПОМЕНА О АУТОРСКИМ ПРАВИМА:**

Овај текст сматра се рукописом и само се саопштава јавности (члан 7. Закона о ауторским  
и сродним правима, "Сл. гласник РС", бр. 104/2009, 99/2011 и 119/2012).

**Ниједан део ове докторске дисертације не сме се користити ни у какве сврхе,  
осим за упознавање са њеним садржајем пре одбране дисертације.**

Ниш, 2020.



UNIVERSITY OF NIŠ  
FACULTY OF SCIENCES AND MATHEMATICS



**Petra N. Laketa**

**RANDOM ENVIRONMENT NONNEGATIVE  
INTEGER-VALUED AUTOREGRESSIVE  
PROCESS GENERATED BY GEOMETRIC  
COUNTING SERIES**

DOCTORAL DISSERTATION


Niš, 2020.

## Подаци о докторској дисертацији

Ментор:	Др Александар С. Настић, ванредни професор, Универзитет у Нишу, Природно-математички факултет
Наслов:	Ненегативни целобројни ауторегресивни процеси у случајној средини генерисани геометријским бројачким низовима
Резиме:	<p>Проучавају се целобројни ауторегресивни процеси у случајној средини генерисани геометријским бројачким низовима. Најпре се уводи модел првог реда, док се касније посматрају модели вишег реда и њихово уопштење. На крају се уводи дводимензионални процес заснован на дводимензионалној серији услова средине. Анализирају се особине уведених модела. Спроводи се оцењивање непознатих параметара модела, а ваљаност ових оцена се испитује на симулацијама. Квалитет модела се потврђује применом над реалним подацима и упоређивањем са другим моделима.</p>
Научна област:	Математичке науке
Научна дисциплина:	Математичка статистика, статистика случајних процеса
Кључне речи:	ИНАР, НГИНАР, случајна средина, негативни биномни тининг оператор, геометријска маргинална расподела, ланац Маркова
УДК:	519.246/.8
CERIF класификација:	P160 Статистика, операционо истраживање, програмирање, актуарска математика
Тип лиценце Креативне заједнице:	<b>CC BY-NC-ND</b>

## Data on Doctoral Dissertation


Doctoral Supervisor:	PhD, associate professor Aleksandar S. Nastić, Faculty of Sciences and Mathematics, University of Niš
Title:	Random Environment Nonnegative Integer-Valued Autoregressive Processes Generated by Geometric Counting Series
Abstract:	<p>Here are analyzed integer autoregressive (INAR) processes in the random environment generated by geometric counting series. Firstly, the first order random environment INAR model is introduced. Later, random environment INAR models of higher order, as well as their general form, are defined. Finally, the bivariate model based on the bivariate random process is defined. The properties of all introduced models are analyzed. Estimation of unknown parameters is given and validate on the simulated data. Model quality is confirmed by application on the real-life data, comparing results with the competitive models.</p>
Scientific Field:	Mathematics
Scientific Discipline:	Mathematical statistics, statistics of random processes
Key Words:	INAR, NGINAR, random environment, negative binomial thinning operator, geometric marginal distribution, Markov chain
UDC:	519.246/.8
CERIF Classification:	P160 Statistics, operations research, programming, actuarial mathematics
Creative Commons License Type:	<b>CC BY-NC-ND</b>

	<b>ПРИРОДНО - МАТЕМАТИЧКИ ФАКУЛТЕТ НИШ</b>
	<b>КЉУЧНА ДОКУМЕНТАЦИЈСКА ИНФОРМАЦИЈА</b>

Редни број, <b>РБР:</b>	
Идентификациони број, <b>ИБР:</b>	
Тип документације, <b>ТД:</b>	монографска
Тип записа, <b>ТЗ:</b>	текстуални / графички
Врста рада, <b>ВР:</b>	докторска дисертација
Аутор, <b>АУ:</b>	Петра Н. Лакета
Ментор, <b>МН:</b>	Александар С. Настић
Наслов рада, <b>НР:</b>	Ненегативни целобројни ауторегресивни процеси у случајној средини генерисани геометријским
Језик публикације, <b>ЈП:</b>	српски
Језик извода, <b>ЈИ:</b>	енглески
Земља публикавања, <b>ЗП:</b>	Србија
Уже географско подручје, <b>УГП:</b>	Србија
Година, <b>ГО:</b>	2020.
Издавач, <b>ИЗ:</b>	ауторски репринт
Место и адреса, <b>МА:</b>	Ниш, Вишеградска 33.
Физички опис рада, <b>ФО:</b> <small>(поглавља/страница/ цитата/табела/слика/графика/прилога)</small>	4/114/38/33/12/0/0
Научна област, <b>НО:</b>	Математичке науке
Научна дисциплина, <b>НД:</b>	Математичка статистика, статистика случајних
Предметна одредница/Кључне речи, <b>ПО:</b>	ИНАР, НГИНАР, случајна средина, негативни биномни тининг оператор, геометријска маргинална расподела, ланац Маркова
<b>УДК</b>	519.246/.8
Чува се, <b>ЧУ:</b>	библиотека
Важна напомена, <b>ВН:</b>	

Извод, <b>ИЗ:</b>	Проучавају се целобројни ауторегресивни процеси у случајној средини генерисани геометријским бројачким низовима. Најпре се уводи модел првог реда, док се касније посматрају модели вишег реда и њихово уопштење. На крају се уводи дводимензионални процес заснован на дводимензионалној серији услова средине. Анализирају се особине уведених модела. Спроводи се оцењивање непознатих параметара модела, а ваљаност ових оцена се испитује на симулацијама. Квалитет модела се потврђује применом над реалним подацима и упоређивањем са другим моделима.
Датум прихватања теме, <b>ДП:</b>	датум прихватања теме на наставно научном већу Универзитета
Датум одбране, <b>ДО:</b>	
Чланови комисије, <b>КО:</b>	Председник:
	Члан:
	Члан, ментор:

Образац Q4.09.13 - Издање 1

	<b>ПРИРОДНО - МАТЕМАТИЧКИ ФАКУЛТЕТ НИШ</b>
	<b>KEY WORDS DOCUMENTATION</b>

Accession number, <b>ANO</b> :	
Identification number, <b>INO</b> :	
Document type, <b>DT</b> :	monograph
Type of record, <b>TR</b> :	textual / graphic
Contents code, <b>CC</b> :	doctoral dissertation
Author, <b>AU</b> :	Petra N. Laketa
Mentor, <b>MN</b> :	Aleksandar S. Nastić
Title, <b>TI</b> :	Random Environment Nonnegative Integer-Valued Autoregressive Processes Generated by Geometric Counting Series
Language of text, <b>LT</b> :	Serbian
Language of abstract, <b>LA</b> :	English
Country of publication, <b>CP</b> :	Serbia
Locality of publication, <b>LP</b> :	Serbia
Publication year, <b>PY</b> :	2020
Publisher, <b>PB</b> :	author's reprint
Publication place, <b>PP</b> :	Niš, Višegradska 33.
Physical description, <b>PD</b> : (chapters/pages/ref./tables/pictures/graphs/appendixes)	4/114/38/33/12/0/0
Scientific field, <b>SF</b> :	Mathematics
Scientific discipline, <b>SD</b> :	Mathematical statistics, statistics of random processes
Subject/Key words, <b>S/KW</b> :	INAR, NGINAR, random environment, negative binomial thinning operator, geometric marginal distribution, Markov chain
<b>UC</b>	519.246/.8
Holding data, <b>HD</b> :	library
Note, <b>N</b> :	

Abstract, <b>AB</b> :	Here are analyzed integer autoregressive (INAR) processes in the random environment generated by geometric counting series. Firstly, the first order random environment INAR model is introduced. Later, random environment INAR models of higher order, as well as their general form, are defined. Finally, the bivariate model based on the bivariate random process is defined. The properties of all introduced models are analyzed. Estimation of unknown parameters is given and validate on the simulated data. Model quality is confirmed by application on the real-life data, comparing results with the competitive models.						
Accepted by the Scientific Board on, <b>ASB</b> :							
Defended on, <b>DE</b> :							
Defended Board, <b>DB</b> :	<table border="1"> <tr> <td data-bbox="400 584 611 629">President:</td> <td data-bbox="611 584 1465 629"></td> </tr> <tr> <td data-bbox="400 629 611 674">Member:</td> <td data-bbox="611 629 1465 674"></td> </tr> <tr> <td data-bbox="400 674 611 725">Member, Mentor:</td> <td data-bbox="611 674 1465 725"></td> </tr> </table>	President:		Member:		Member, Mentor:	
President:							
Member:							
Member, Mentor:							

Образац Q4.09.13 - Издање 1



# Предговор

Човека радозналост од давнина наводи да опише свет који га окружује. Први покушаји огледали су се у настанку многобројних религија. Наши преци су послушљивали природу и тражили њене тајне сигнале. Прилагали су жртве не би ли умилостивили богове. Они који су умели да предвиде кишу или сушу били су високо вредновани. Вео мистицизма наизглед је скинут. Улогу пророка преузели су научници, чије магичне формуле и прорачуни задивљују оне који о томе не знају ништа. У науку се верује. Данас се барата чињеницама и прича о неоспорним истинама потврђеним експериментима. Међутим, чини се да никада нисмо мање знали о томе шта је истина и које су праве вредности. Неодређеност се ушуњала на мала врата кроз квантну механику. Уместо тачних вредности положаја и импулса имамо само њихове вероватноће. Извесност је заменила пробабилистичка теорија. Човек је препуштен на милост и немилост случају. Ова дисертација се бави увођењем нових математичких структура којима се могу описати одређене појаве и предвидети њихово понашање у будућности. Резултат је истраживања вођених у току докторских студија. Аутор је дубоко свестан ништавности једног оваквог подухвата у односу на науку у целини, као и ништавности науке у односу на човека и свет уопште. Међутим, како је ово ипак један значајан корак у спољашњем свету, захвалила бих онима који су допринели настанку овог текста на посредан или непосредан начин.

Најпре исказујем захвалност ментору, др Александру Настићу, ванредном професору Природно-математичког факултета у Нишу на свој подршци током ових година и свему што ме је научио. Такође бих захвалила др Мирославу Ристићу, редовном професору на истом факултету, који је у извесном смислу био коментор

мојих студија. Ова два професора су ме увела у цео свет INAR модела и несебично поделила са мном оно што је резултат њиховог вишегодишњег рада. Захваљујем и свим професорима, колегама, као и породици и пријатељима на друженом знању и подршци.

Први покушаји моделирања временских низова са ненегативним целобројним вредностима су се заснивали на примени ауторегресивних модела са непрекидним маргиналним расподелама. Јасно је да овакав приступ није могао дати довољно добре резултате, с обзиром на то да су се дискретне вредности описивале помоћу непрекидне расподеле. Због тога су Jacobs и Lewis (1978a-c) увели дискретне ауторегресивне моделе покретних средина (DARMA). Затим је уследило увођење првог целобројног ауторегресивног модела (INAR - Integer-Valued AutoRegressive), и то INAR(1) модела, првог реда, независним приступом од стране McKenzie (1985) и Al-Osh и Alzaid (1987). Овај модел има следећи облик

$$X_n = \alpha \circ X_{n-1} + \varepsilon_n, \quad n \geq 1,$$

где  $\alpha \circ$  представља оператор који је у претходним научним публикацијама на српском језику називан “тининг” оператором или оператором истањивања и дефинисан је као  $\alpha \circ X = \sum_{i=1}^X Y_i$ , при чему је  $\{Y_i\}$  низ независних и једнако расподељених случајних променљивих са параметром расподеле  $\alpha$ , који се такође зове и бројачки низ. У зависности од тога каква је маргинална расподела бројачког низа, могу се дефинисати различити тининг оператори. Њихов назив зависи од расподеле случајне променљиве  $\alpha \circ X|X$ . У случају да су  $Y_i$  Бернулијеве случајне променљиве са  $P(Y_i = 1) = 1 - P(Y_i = 0) = \alpha$ , ради се о биномном оператору истањивања. Ако су  $Y_i$  геометријски расподељене случајне променљиве са очекивањем  $\alpha$ , тада је у питању негативни биномни оператор истањивања, пошто  $\alpha \circ X|X$  има негативну биномну расподелу. У раду Ristić, Bakouch и Nastić (2009) уведен је INAR модел са геометријском маргиналном расподелом заснован на негативном биномном оператору истањивања (NGINAR(1)). Уопштење овог модела дато је у раду Nastić, Ristić и Bakouch (2012), у коме је дефинисан комбиновани геометријски INAR модел реда  $p$  (NGINAR( $p$ )). Такође, Popović, Ristić и Nastić (2016) разматрали су дводимензионални NGINAR модел.

Дисертација садржи резултате објављених радова. Најпре је, у првој глави, дата мотивација за увођење случајне средине у INAR моделе. Друга глава садржи резултате у вези са првим INAR моделом у случајној средини, који је првог реда. У трећој глави су затим анализирани уопштени INAR модели у случајној средини вишег реда. Дводимензионални INAR модел у случајној средини тема је четврте главе. Након тога следи закључак, као и списак коришћене литературе.

Петра Н. Лакета



# Садржај

<b>1</b>	<b>Увођење случајне средине у INAR моделе</b>	<b>7</b>
<b>2</b>	<b>Ненегативни целобројни ауторегресивни модел првог реда у случајној средини</b>	<b>13</b>
2.1	Конструкција модела . . . . .	13
2.2	Особине модела . . . . .	18
2.3	Оцењивање непознатих параметара . . . . .	25
2.3.1	Јул-Вокер оцене . . . . .	26
2.3.2	Оцене условне максималне веродостојности . . . . .	33
2.3.3	Анализа квалитета оцена над симулираним подацима . . . . .	34
2.4	Примена модела над реалним подацима . . . . .	41
<b>3</b>	<b>Уопштени INAR модели вишег реда у случајној средини</b>	<b>47</b>
3.1	Конструкција модела . . . . .	48
3.2	Особине модела . . . . .	53
3.3	Оцењивање непознатих параметара . . . . .	57
3.3.1	Јул-Вокер оцене . . . . .	57
3.3.2	Оцене условне максималне веродостојности . . . . .	61
3.3.3	Анализа квалитета оцена над симулираним подацима . . . . .	61
3.4	Примена модела над реалним подацима . . . . .	67
<b>4</b>	<b>Укрштени дводимензионални INAR модел првог реда у случајној средини</b>	<b>77</b>
4.1	Конструкција модела . . . . .	78

4.2	Особине модела . . . . .	82
4.3	Оцењивање параметара . . . . .	89
4.3.1	Јул-Вокер оцењивање . . . . .	89
4.3.2	Оцене условне максималне веродостојности . . . . .	92
4.4	Анализа квалитета оцена над симулираним подацима . . . . .	93
4.5	Нови метод оцењивања процеса случајне средине . . . . .	98
4.6	Примена модела над реалним подацима . . . . .	101
4.6.1	Дискусија резултата у односу на коефицијент $k$ метода оцењивања из претходне секције	101
4.6.2	Резултати примене . . . . .	104
	<b>Закључак</b>	<b>107</b>
	<b>Литература</b>	<b>109</b>
	<b>Биографија</b>	<b>113</b>

# Глава 1

## Увођење случајне средине у INAR моделе

У овој глави ће најпре бити речено са којим проблемом су се INAR модели срели, а који се може решити увођењем случајне средине. Затим ће се описати суштина идеје модела који су тема ове дисертације и дефинисаће се процес случајне средине који служи као помоћни процес у конструкцији INAR модела у случајној средини.

Досадашња пракса примењивања стационарних INAR модела у раду са подацима чији график показује одређене нестационарне карактеристике није била најадекватнија. Због тога је природно трагати за нестационарним INAR моделима, како би се могли моделовати подаци са ненегативним целобројним вредностима који немају стационаран карактер. Такође, пожељно је да ови модели буду што једноставнији за рад и у неком смислу што сличнији стационарним INAR моделима. Један од начина на који се нестационарност може укључити је увођење новог концепта конструкције модела применом случајне средине. Као што знамо, у природи је све повезано. Свака величина коју меримо налази се у некаквим условима које одређује околина и који утичу на њене вредности. Промене ових услова би, дакле, требало да доведу до промене расподеле мерене величине. Претпоставимо, примера ради, да на мерене вредности има утицај температура. Уместо да температуру одредимо тачним бројем степени (Целзијусових или Келвина), можемо поједноставити ситуацију дискретизацијом

вредности. Наиме, можемо посматрати три различита стања температуре: ниска, средња и висока. Ако је температура ниска, сматраћемо да процес има једну расподелу, уколико је средња другу, а у случају високе температуре трећу. Уколико желимо да још више нагласимо нестационарност, можемо повећати број стања, тако што ћемо посматрати стања екстремно ниске, ниске, средње, високе и екстремно високе температуре. Уопште, ради једноставности, сматраћемо да се услови средине могу поделити у  $r$  различитих стања. Притом је у конкретном случају добро да број  $r$  није, сувише мали, али ни сувише велики. Наиме, што је број стања већи, то је модел флексибилнији, па самим тим боље описује податке. Међутим, ово истовремено узрокује велики број параметара модела, што компликује рад са таквим моделом. Сада ћемо дати формалну дефиницију процеса којим моделирамо услове средине.

*ДЕФИНИЦИЈА 1.0.1 Низ случајних променљивих  $\{Z_n\}$ ,  $n \in \mathbb{N}_0$ , је процес случајне средине са  $r$  стања, за  $r \in \mathbb{N}$ , ако је он ланац Маркова првог реда који узима вредности из скупа  $E_r = \{1, 2, \dots, r\}$ . Уопште,  $\{Z_n\}$ ,  $n \in \mathbb{N}_0$ , је процес случајне средине ако је он процес случајне средине са  $r$  стања за неко  $r \in \mathbb{N}$ .*

Овде треба обратити пажњу на саму природу управо дефинисаног процеса. Приметимо да скуп  $E_r$  садржи првих  $r$  природних бројева. Можемо поставити питање шта представљају ове вредности и како се мере у експерименту. Истина је да оне нису никакве опсервирајуће величине, тј. не репрезентују никакву мерљиву карактеристику средине. Наиме, ако бисмо у експерименту регистровали вредности које желимо описати INAR моделом са случајном средином, никаквих мерења саме средине не би било. Ово може деловати збуњујуће, међутим, треба имати на уму да циљ нашег истраживања није сама случајна средина, већ искључиво њен утицај на посматрану величину чије вредности моделирамо. У том смислу нам нису важне особине средине, јер је њена једина улога да обезбеди увођење нестационарности у одређени INAR модел.

Следеће питање које се намеће је како онда одредити вредности процеса случајне средине, ако се оне не добијају експерименталним путем. Посматрањем података које желимо моделирати може



се утврдити у колико различитих скупова се они групишу (притом овај избор није једнозначан), те у складу са тим претпоставити да има толико различитих стања средине. Ако је број ових стања  $r$ , онда је довољно да кажемо да процес случајне средине има вредности из скупа  $E_r = \{1, 2, \dots, r\}$ , при чему сваки пут када је  $Z_n = i$ , сматрамо да се средина налази у  $i$ -том стању и да у складу са тим INAR модел има  $i$ -ту расподелу из фамилије могућих расподела  $\{F(x, \theta_k), k \in E_r\}$ . Дакле, саме вредности из скупа  $E_r$  нису важне, већ само чињеница да их има  $r$  различитих и да је свако стање повезано са тачно једном расподелом посматраног процеса.

Ако се вратимо сада примеру са температуром, могуће је ниској температури придружити број 1, средњој 2, а високој 3. Међутим, могуће је нумерисати их и на било који други начин, рецимо средња 1, ниска 2, висока 3. Дакле, као што је већ речено, сама нумерација није важна, као ни стварно описивање средине. У овом конкретном примеру једино је важно моделирати на који начин промене температуре утичу на промену расподеле посматраног процеса.

Приметимо још једну значајну особину. Нека, примера ради, за реализоване вредности важи  $z_n = z_{n+1} = \dots = z_m = i$ , где  $i \in E_r$  и  $z_{n-1} \neq z_n$ ,  $z_m \neq z_{m+1}$ . Тада део процеса  $X_n, X_{n+1}, \dots, X_m$  одговара стању  $i$ , док елементи процеса суседни овом делу одговарају неком другом стању. У том смислу се овај део може посматрати као део стационарног процеса, пошто нема промене стања. На овај начин се INAR модели у случајној средини могу посматрати као процеси који су део по део стационарни. Стога се неки познати закључци о стационарним INAR моделима могу применити парцијално. Управо ово својство чини моделе који су тема ове дисертације једноставним за анализу. Величина ових “стационарних делова” је одређена матрицом вероватноћа прелаза. Што су веће вредности на дијагонали ове матрице, то је већа вероватноћа да процес остане у истом стању, што имплицира дуже делове који одговарају једном стању.

Када је реч о Јул-Вокер оценама, у досадашњој пракси INAR модела користила се особина стационарности модела за доказивање строге постојаности. Јасно је да у нашем случају нестационарних модела ову технику није могуће применити. Међутим, оцене се могу дефинисати на малопре поменути “стационарним

деловима” и показује се да су оне строго постојане (у следећој глави се налази теорема која то потврђује).

Овде ћемо ипак посматрати процес базиран на реализованим вредностима  $z_n$ . Наиме, када желимо да применимо модел над неким подацима, прво ћемо одредити низ  $\{z_n\}$ , па затим формирати модел. У ову сврху се користи кластеровање. Након што одлучимо колико стања  $r$  ћемо користити (нешто више о самом одабиру броја  $r$  биће речено касније), подаци се групишу у  $r$  различитих кластера и онда се сваки кластер посматра као једно стање. Дакле, ако неки податак припада  $i$ -том кластеру, ми ћемо сматрати да је он измерен у таквим условима средине који одговарају стању  $i$ . На тај начин је сваком податку придружено одговарајуће стање, тј. одређен је низ  $\{z_n\}$ .

Могуће је оценити и матрицу вероватноћа прелаза која одговара овом ланцу Маркова на уобичајен начин оцењивања вероватноћа као количника броја повољних и укупних исхода. На пример, ако желимо да оценимо вероватноћу преласка из стања  $i$  у стање  $j$ , као оцену ћемо користити количник броја свих оваквих реализованих прелаза и броја свих могућих прелаза из стања  $i$  у неко друго стање. У случају да нам је потребно предвиђање, а имамо узорак обима  $N$ , можемо предвидети  $Z_{N+1}$  на основу  $Z_N$  помоћу матрице вероватноћа прелаза, затим  $Z_{N+2}$  на основу  $Z_{N+1}$  итд. На тај начин можемо генерисати низ реализација  $\{z_n\}$  произвољне дужине, а онда поново применити модел базиран на овим реализацијама. Овоме се прибегава због поједностављивања самог модела, који би био комплекснији у случају да се користе случајне величине, односно процес случајне средине.

Оваква веза између INAR модела и случајне средине се може још више уопштити. Наиме, до сада смо поменули да стања средине одређују једино расподелу посматраног процеса. Може се отићи корак даље и претпоставити да осим расподеле ова стања одређују и ред модела, као и параметар тининг оператора којим се модел генерише. Више о томе се налази у четвртој глави. Такође, у случају дводимензионалног процеса, могућа је ситуација да постоје две различите средине, од којих једна више утиче на једну компоненту процеса, а друга на другу. У том циљу уведен је дводимензионални процес случајне средине на начин приказан у следећој дефиницији.

*ДЕФИНИЦИЈА 1.0.2* Низ дводимензионалних случајних променљивих  $\{(W_n, Q_n)\}$ ,  $n \in \mathbb{N}_0$ , је **дводимензионални процес случајне средине са  $(r_1, r_2)$  стања**, за  $r_1, r_2$  из скупа природних бројева  $\mathbb{N}$ , ако су  $\{W_n\}$  и  $\{Q_n\}$ ,  $n \in \mathbb{N}_0$ , процеси случајне средине са  $r_1$ , односно  $r_2$  стања. Уопште,  $\{(W_n, Q_n)\}$ ,  $n \in \mathbb{N}_0$ , је дводимензионални процес случајне средине ако је он дводимензионални процес случајне средине са  $(r_1, r_2)$  стања, за неке  $r_1, r_2 \in \mathbb{N}$ .

Наравно, могуће би било потпуно аналогно дефинисати процес случајне средине произвољне димензије, међутим није било потребе за тим узевши у обзир проблеме који се овде разматрају.



## Глава 2

# Ненегативни целобројни ауторегресивни модел првог реда у случајној средини

Први корак у дефинисању INAR модела у случајној средини учињен је радом Nastić, Laketa и Ristić (2016), који ће бити тема овог поглавља и у ком је дефинисан први модел тог типа. Најпре ће бити приказана конструкција самог модела, а затим и његове особине, као што су расподела иновационих случајних променљивих, корелациона структура процеса и условне величине. Након тога ће пажња бити посвећена оцењивању непознатих параметара. У ту сврху су коришћена два метода: метод момената (Јул-Вокер) и метод условних најмањих квадрата. Такође, квалитет оцена је потврђен тако што су симулиране вредности процеса и на основу њих израчунате реализоване оцене параметара које показују конвергенцију ка тачним вредностима параметара при повећању обима узорка. На крају, модел је примењен над подацима из стварног живота, тзв. реалним подацима, где је показао боље резултате у погледу “фитовања” података у поређењу са другим моделима који долазе у обзир за посматрани тип података.

### 2.1 Конструкција модела

Као инспирација за увођење овог модела, послужио је рад Tang

и Wang (2014), у ком је посматран процес  $\{X_n\}$ ,  $n \in \mathbb{N}_0$ , одређен релацијом

$$X_n = \phi_n \circ X_{n-1} + \sum_{j=1}^r \varepsilon_n(j) I_{\{Z_n=j\}}, \quad n \in \mathbb{N},$$

где је  $\phi_n \circ X_{n-1} = \sum_{i=1}^{X_{n-1}} U_i$ , при чему је  $\{U_i\}$  низ независних и једнако расподељених Бернулијевих случајних променљивих са  $P(U_i = 1) = \phi_n$  и  $\{\varepsilon_n(j)\}$  низ ненегативних независних и једнако расподељених случајних променљивих, док  $I_{\{Z_n=j\}}$  означава индикатор догађаја  $Z_n = j$ . Модификовањем ове формуле, може се добити модел заснован на процесу случајне средине на начин који је дат у следећој дефиницији.

**ДЕФИНИЦИЈА 2.1.1** *За низ ненегативних целобројних случајних променљивих  $\{X_n(Z_n)\}$ ,  $n \in \mathbb{N}_0$ , кажемо да је ненегативни целобројни ауторегресивни процес првог реда у случајној средини са  $r$  стања ( $RrINAR(1)$ ) ако је дат релацијом*

$$X_n(Z_n) = \sum_{i=1}^{X_{n-1}(Z_{n-1})} U_i + \varepsilon_n(Z_{n-1}, Z_n), \quad n \in \mathbb{N},$$

где је

$$X_n(Z_n) = \sum_{z=1}^r X_n(z) I_{\{Z_n=z\}},$$

$$\varepsilon_n(Z_{n-1}, Z_n) = \sum_{z_1=1}^r \sum_{z_2=1}^r \varepsilon_n(z_1, z_2) I_{\{Z_{n-1}=z_1, Z_n=z_2\}},$$

$\{U_i\}$ ,  $i \in \mathbb{N}$ , је бројачки низ независних једнако расподељених случајних променљивих који генерише тининг оператор,  $\{Z_n\}$  је процес случајне средине са  $r$  стања, дат дефиницијом 1.0.1 и  $\{\varepsilon_n(i, j)\}$ , за  $n \in \mathbb{N}$ ,  $i, j \in E_r$ , су низови независних једнако расподељених случајних променљивих, који задовољавају следеће услове:

(A1)  $\{Z_n\}$ , за  $n \in \mathbb{N}_0$  и  $\{\varepsilon_n(1, 1)\}, \{\varepsilon_n(1, 2)\}, \dots, \{\varepsilon_n(r, r)\}$ , за  $n \in \mathbb{N}$  су међусобно независни

(A2)  $X_n(l)$  не зависи од  $Z_m$  и  $\varepsilon_m(i, j)$ , за  $n < m$  и произвољне  $i, j, l \in E_r$ .

За произвољно  $n \in \mathbb{N}$ , случајне променљиве  $X_n(Z_n)$  и  $\varepsilon_n(Z_{n-1}, Z_n)$  могу се представити на следећи начин

$$X_n(Z_n) \stackrel{d}{=} \begin{cases} X_n(1), & \text{са вероватноћом } P(Z_n = 1), \\ X_n(2), & \text{са вероватноћом } P(Z_n = 2), \\ \vdots & \vdots \\ X_n(r), & \text{са вероватноћом } P(Z_n = r), \end{cases}$$

$$\varepsilon_n(Z_{n-1}, Z_n) \stackrel{d}{=} \begin{cases} \varepsilon_n(1, 1), & \text{са вероватноћом } P(Z_{n-1} = 1, Z_n = 1), \\ \varepsilon_n(1, 2), & \text{са вероватноћом } P(Z_{n-1} = 1, Z_n = 2), \\ \vdots & \vdots \\ \varepsilon_n(r, r), & \text{са вероватноћом } P(Z_{n-1} = r, Z_n = r). \end{cases}$$

У претходним формулама  $\stackrel{d}{=}$  означава једнакост у расподели.

**Теорема 2.1.1** *Дводимензионални временски низ  $\{X_n(Z_n), Z_n\}$ , дат дефиницијама 1.0.1 и 2.1.1 је ланац Маркова првог реда.*

*Доказ.* Уведимо најпре следеће ознаке  $\mathbf{Y}_n = (X_n(Z_n), Z_n)$  и  $\mathbf{y}_n = (x_n, z_n)$ . Приметимо да је

$$\mathbf{Y}_n = \mathbf{y}_n \Leftrightarrow (Z_n = z_n \wedge X_n(z_n) = x_n).$$

Нека је  $A = \{\mathbf{Y}_s = \mathbf{y}_s, 0 \leq s < n - 1\}$ . Означимо  $p_{n-1, n} = P(Z_n = z_n | Z_{n-1} = z_{n-1})$ . Посматрајмо условну вероватноћу  $P_{n-1, n} = P(\mathbf{Y}_n = \mathbf{y}_n | \mathbf{Y}_{n-1} = \mathbf{y}_{n-1}, A)$ . Користећи дефиницију процеса, имамо да је

$$\mathbf{Y}_n = \mathbf{y}_n \Leftrightarrow \left( \sum_{i=1}^{X_{n-1}(Z_{n-1})} U_i + \varepsilon_n(Z_{n-1}, z_n) = x_n \wedge Z_n = z_n \right),$$

док је

$$\mathbf{Y}_{n-1} = \mathbf{y}_{n-1} \Leftrightarrow (Z_{n-1} = z_{n-1} \wedge X_{n-1}(z_{n-1}) = x_{n-1}).$$

Одатле закључујемо да је

$$P_{n-1,n} = P \left( \sum_{i=1}^{x_{n-1}} U_i + \varepsilon_n(z_{n-1}, z_n) = x_n, Z_n = z_n \mid B \right),$$

где смо са  $B$  означили догађај  $\{Z_{n-1} = z_{n-1}\} \cap \{X_{n-1}(z_{n-1}) = x_{n-1}\} \cap A$ . Због независности случајних променљивих  $\sum_{i=1}^{x_{n-1}} U_i + \varepsilon_n(z_{n-1}, z_n)$  и  $Z_n$ , ова вероватноћа се може представити као производ

$$P \left( \sum_{i=1}^{x_{n-1}} U_i + \varepsilon_n(z_{n-1}, z_n) = x_n \mid B \right) P(Z_n = z_n \mid B).$$

Како  $\sum_{i=1}^{x_{n-1}} U_i + \varepsilon_n(z_{n-1}, z_n)$  не зависи од случајних променљивих које се јављају у  $B$ , то је први члан производа једнак вероватноћи  $P(\sum_{i=1}^{x_{n-1}} U_i + \varepsilon_n(z_{n-1}, z_n) = x_n)$ , док  $Z_n$  на основу тога што је  $\{Z_n\}$  ланац Маркова првог реда зависи само од  $Z_{n-1}$ , па је други члан једнак  $P(Z_n = z_n \mid Z_{n-1} = z_{n-1}) = p_{n-1,n}$ . Коначно, имамо да је

$$P_{n-1,n} = p_{n-1,n} P \left( \sum_{i=1}^{x_{n-1}} U_i + \varepsilon_n(z_{n-1}, z_n) = x_n \right). \quad (2.1.1)$$

Сада ћемо сличан поступак применити на  $P(\mathbf{Y}_n = \mathbf{y}_n \mid \mathbf{Y}_{n-1} = \mathbf{y}_{n-1})$ . Поново ову вероватноћу можемо приказати као производ две вероватноће и на исти начин, као мало пре, закључујемо да је

$$P(\mathbf{Y}_n = \mathbf{y}_n \mid \mathbf{Y}_{n-1} = \mathbf{y}_{n-1}) = p_{n-1,n} P \left( \sum_{i=1}^{x_{n-1}} U_i + \varepsilon_n(z_{n-1}, z_n) = x_n \right). \quad (2.1.2)$$

Како су десне стране (2.1.1) и (2.1.2) једнаке, можемо изједначити и леве стране ових израза, па је

$$P(\mathbf{Y}_n = \mathbf{y}_n \mid \mathbf{Y}_{n-1} = \mathbf{y}_{n-1}) = P(\mathbf{Y}_n = \mathbf{y}_n \mid \mathbf{Y}_{n-1} = \mathbf{y}_{n-1}, A),$$

тј.  $\{X_n(Z_n), Z_n\}$  је ланац Маркова првог реда.  $\square$

У претходној глави је било речи о томе да је у пракси корисно заменити процес случајне средине  $\{Z_n\}$  његовом реализацијом  $\{z_n\}$ . Наиме, како  $Z_n$  може узети било коју од  $r$  вредности из скупа  $E_r$ , то се вероватноћа  $P(X_n(Z_n) = x)$  рачуна на следећи



начин

$$\begin{aligned} P(X_n(Z_n) = x) &= \sum_{i=1}^r P(X_n(Z_n) = x | Z_n = i) P(Z_n = i) \\ &= \sum_{i=1}^r P(X_n(i) = x) P(Z_n = i). \end{aligned}$$

Јасно је да би оваква особина учинила модел мање подесним за рад, па је у том циљу природно у дефиницији користити реализацију  $\{z_n\}$ . Једини разлог који би био противан оваквом поједностављењу била би сумња да се губљењем случајности овог процеса може изгубити флексибилност која је била главна карактеристика модела. Међутим, ово ипак није случај. У пракси се низ  $\{z_n\}$  одређује на основу узорка и то на тај начин да се модел што више прилагоди подацима (више о самом одређивању овог низа биће речено у секцији која се бави оцењивањем непознатих параметара, а основна идеја је кластеровање података). Са друге стране, ако посматрамо узорак обима  $N$ , а желимо да предвидимо будуће вредности  $X_n(z_n)$ , за  $n > N$ , можемо се вратити процесу случајне средине  $\{Z_n\}$  и помоћу матрице вероватноћа прелаза (коју можемо оценити на основу већ одређених вредности  $z_1, z_2, \dots, z_N$ ) генерисати будуће реализације овог процеса  $z_n$ , за  $n > N$ , након чега опет можемо радити само са реализованим низом  $\{z_n\}$ .

Осим ове, направљене су додатне претпоставке. Због природе посматраних реалних података, у којима је углавном присутна овердисперзија, претпостављено је да је маргинална расподела геометријска. Осим тога, модел је заснован на негативном биномном тининг оператору. Ово се може разумети тако што су коришћени подаци који представљају број месечних криминалних активности. Зато има смисла да се овај број криминалних активности у следећем месецу проузрокује путем негативног биномног тининг оператора. Прецизна дефиниција модела који се добија тако што се моделу из дефиниције 2.1.1 наметну све поменуте претпоставке дата је у наставку.

*ДЕФИНИЦИЈА 2.1.2 Нека је  $\{z_n\}$  реализација процеса  $\{Z_n\}$  случајне средине са  $r$  стања. Кажемо да је низ  $\{X_n(z_n)\}$ ,  $n \in \mathbb{N}_0$ ,  $INAR(1)$  процес у случајној средини са  $r$  стања и геометријском маргиналном*

расподелом, заснован на негативном биномном тининг оператору ( $RrNGINAR(1)$ ), ако је случајна променљива  $X_n(z_n)$  дефинисана на следећи начин

$$X_n(z_n) = \alpha * X_{n-1}(z_{n-1}) + \varepsilon_n(z_{n-1}, z_n), \quad n \in \mathbb{N}, \quad (2.1.3)$$

где је  $\alpha \in (0, 1)$ , бројачки низ  $\{U_i\}$ ,  $i \in \mathbb{N}$ , укључен у  $\alpha*$  представља низ независних једнако расподелених случајних променљивих са следећом расподелом

$$P(U_i = u) = \frac{\alpha^u}{(1 + \alpha)^{u+1}}, \quad u \in \mathbb{N}_0,$$

случајна променљива  $\varepsilon_n(z_{n-1}, z_n)$  је дата у дефиницији 2.1.1 и важи

$$P(X_n(z_n) = x) = \frac{\mu_{z_n}^x}{(1 + \mu_{z_n})^{x+1}}, \quad x \in \mathbb{N}_0, \quad (2.1.4)$$

где  $\mu_{z_n} \in \{\mu_1, \mu_2, \dots, \mu_r\}$  и  $r \in \mathbb{N}$ .

Овакав модел је само један од многих специјалних случајева модела датог дефиницијом 2.1.1 који се могу дефинисати, што отвара бројне могућности за даља истраживања.

## 2.2 Особине модела

У овој секцији ћемо истражити неке особине управо конструисаног модела. Почнимо најпре са теоремом о расподели иновационе случајне променљиве.

**Теорема 2.2.1** Нека је  $\{X_n(z_n)\}$ ,  $n \in \mathbb{N}_0$ ,  $RrNGINAR(1)$  процес дат дефиницијом 2.1.2 и нека је  $\mu_1 > 0$ ,  $\mu_2 > 0$ ,  $\dots$ ,  $\mu_r > 0$ . Ако  $0 \leq \alpha \leq \min \left\{ \frac{\mu_l}{1 + \mu_k}; k, l \in E_r \right\}$ , тада расподеле случајних променљивих  $\varepsilon_n(1, 1)$ ,  $\varepsilon_n(1, 2)$ ,  $\dots$ ,  $\varepsilon_n(r, r)$ , могу бити представљене као мешавине две геометријске расподеле, тј. за  $z_n = j$  и  $z_{n-1} = i$ ,  $i, j \in E_r$  је

$$\varepsilon_n(i, j) \stackrel{d}{=} \begin{cases} \text{Geom} \left( \frac{\mu_j}{1 + \mu_j} \right), & \text{с.в. } 1 - \frac{\alpha \mu_i}{\mu_j - \alpha}, \\ \text{Geom} \left( \frac{\alpha}{1 + \alpha} \right), & \text{с.в. } \frac{\alpha \mu_i}{\mu_j - \alpha}. \end{cases} \quad (2.2.1)$$

Притом ознака  $\text{Geom} \left( \frac{\mu}{1 + \mu} \right)$  означава геометријску расподелу са очекивањем  $\mu$ .

*Доказ.* Доказ ћемо спровести посматрајући функцију генератрисе вероватноћа. На основу једначине (2.1.3), узимајући у обзир услове теореме  $z_n = j$  и  $z_{n-1} = i$ , важи  $X_n(j) = \alpha * X_{n-1}(i) + \varepsilon_n(i, j)$ , па је

$$\Phi_{X_n(j)}(s) = E(s^{X_n(j)}) = E(s^{\alpha * X_{n-1}(i) + \varepsilon_n(i, j)}). \quad (2.2.2)$$

Посматрајмо сада десну страну једначине (2.2.2). Из чињенице да су случајне променљиве  $\varepsilon_n(i, j)$  и  $X_{n-1}(i)$  независне, следи да важи

$$E(s^{\alpha * X_{n-1}(i) + \varepsilon_n(i, j)}) = E(s^{\alpha * X_{n-1}(i)}) E(s^{\varepsilon_n(i, j)}).$$

Даље, на основу особина негативног биномног тининг оператора важи

$$E(s^{\alpha * X_{n-1}(i)}) = E\left((Es^{U_1})^{X_{n-1}(i)}\right) = \Phi_{X_{n-1}(i)}(\Phi_{U_1}(s)).$$

Сада имамо да је

$$\Phi_{X_n(j)}(s) = \Phi_{X_{n-1}(i)}(\Phi_{U_1}(s))\Phi_{\varepsilon_n(i, j)}(s),$$

па је

$$\Phi_{\varepsilon_n(i, j)}(s) = \frac{\Phi_{X_n(j)}(s)}{\Phi_{X_{n-1}(i)}(\Phi_{U_1}(s))}. \quad (2.2.3)$$

На основу релације (2.1.4),  $X_n(j)$  има геометријску расподелу, па је на основу особина геометријске расподеле

$$\Phi_{X_n(j)}(s) = \frac{1}{1 + \mu_j - \mu_j s}. \quad (2.2.4)$$

Слично је и

$$\Phi_{X_{n-1}(i)}(s) = \frac{1}{1 + \mu_i - \mu_i s}. \quad (2.2.5)$$

Са друге стране,  $U_1$  има геометријску расподелу са очекивањем  $\alpha$ , па је

$$\Phi_{U_1}(s) = \frac{1}{1 + \alpha - \alpha s}.$$

Сада је

$$\Phi_{X_{n-1}(i)}(\Phi_{U_1}(s)) = E\left(\left(\frac{1}{1 + \alpha - \alpha s}\right)^{X_{n-1}(i)}\right) \quad (2.2.6)$$

$$= \frac{1}{1 + \mu_i - \mu_i \frac{1}{1 + \alpha - \alpha s}} = \frac{1 + \alpha - \alpha s}{1 + \alpha(1 + \mu_i) - \alpha(1 + \mu_i)s},$$

где претпоследња једнакост следи из чињенице да  $X_{n-1}(i)$  има геометријску расподелу са очекивањем  $\mu_i$ . Ако сада једначине (2.2.4), (2.2.5) и (2.2.6) уврстимо у једначину (2.2.3), добијамо

$$\begin{aligned} \Phi_{\varepsilon_n(i,j)}(s) &= \frac{1 + \alpha(1 + \mu_i) - \alpha(1 + \mu_i)s}{(1 + \alpha - \alpha s)(1 + \mu_j - \mu_j s)} = \frac{1 + \alpha - \alpha s + \alpha\mu_i + \alpha\mu_i s}{(1 + \alpha - \alpha s)(1 + \mu_j - \mu_j s)} \\ &= \frac{1}{1 + \mu_j - \mu_j s} + \frac{\alpha\mu_i}{\mu_j - \alpha} \cdot \frac{\mu_j - \mu_j s - \alpha + \alpha s}{(1 + \alpha - \alpha s)(1 + \mu_j - \mu_j s)} \\ &= \frac{1}{1 + \mu_j - \mu_j s} + \frac{\alpha\mu_i}{\mu_j - \alpha} \cdot \left( \frac{1}{1 + \alpha - \alpha s} - \frac{1}{1 + \mu_j - \mu_j s} \right) \\ &= \left( 1 - \frac{\alpha\mu_i}{\mu_j - \alpha} \right) \cdot \frac{1}{1 + \mu_j - \mu_j s} + \frac{\alpha\mu_i}{\mu_j - \alpha} \cdot \frac{1}{1 + \alpha - \alpha s}. \end{aligned}$$

На основу претходног израза којим је одређена функција генератрисе вероватноће случајне променљиве  $\varepsilon_n(i, j)$ , можемо видети да она има расподелу дату са (2.2.1), ако је задовољен услов  $0 \leq \alpha \leq \frac{\mu_j}{1 + \mu_i}$ , којим се обезбеђује да су одговарајуће вероватноће ненегативне. Како су  $i$  и  $j$  произвољне вредности из скупа  $E_r$ , то за све  $i, j \in E_r$  важи  $0 \leq \alpha \leq \frac{\mu_j}{1 + \mu_i}$ . Ово ће бити постигнуто, ако је  $\alpha \in \bigcap_{k, l \in E_r} \left[ 0, \frac{\mu_l}{1 + \mu_k} \right]$ , тј. за  $0 \leq \alpha \leq \min \left\{ \frac{\mu_l}{1 + \mu_k}; k, l \in E_r \right\}$ , што је претпостављено у формулацији теореме.  $\square$

Као и код сваког иновационог процеса, потребно је познавати његову очекивану вредност, као и степен волатилности, о чему говори следеће тврђење.

**Теорема 2.2.2** Нека је  $z_{n-1} = i$  и  $z_n = j$ , где су  $i, j \in E_r$ . Тада су математичко очекивање и дисперзија случајне променљиве  $\varepsilon_n(i, j)$  редом дати једнакостима  $E(\varepsilon_n(i, j)) = \mu_j - \alpha\mu_i$  и  $D(\varepsilon_n(i, j)) = \mu_j(1 + \mu_j) - \alpha\mu_i(1 + 2\alpha + \alpha\mu_i)$ .

*Доказ.* Нека је  $\Phi_\varepsilon(s)$  функција генератрисе вероватноћа случајне променљиве  $\varepsilon_n(i, j)$ . Тада доказ следи из чињенице да је  $E(\varepsilon_n(i, j)) = \Phi'_\varepsilon(1)$  и  $D(\varepsilon_n(i, j)) = \Phi''_\varepsilon(1) + \Phi'_\varepsilon(1)(1 - \Phi'_\varepsilon(1))$ , где су  $\Phi'_\varepsilon(1) = \mu_j - \alpha\mu_i$  и  $\Phi''_\varepsilon(1) = 2(\mu_j^2 - \alpha\mu_i\mu_j - \alpha^2\mu_i)$ .  $\square$

Приметимо да је очекивање случајне променљиве  $\varepsilon_n(i, j)$  за  $i, j \in E_r$  позитивно. Наиме, из теореме 2.2.1 имамо да је

$$0 \leq \alpha \leq \min \left\{ \frac{\mu_l}{1 + \mu_k}; k, l \in E_r \right\} \leq \frac{\mu_j}{1 + \mu_i} < \frac{\mu_j}{\mu_i}. \quad (2.2.7)$$

Са друге стране, једна од најважнијих особина сваког ауторегресивног процеса је међусобна повезаност и узајамна зависност њених појединих елемената. Ово својство је познато под називом коваријансно-корелациона структура, о чему говори наредна теорема.

**Теорема 2.2.3** *Нека је  $\{X_n(z_n)\}$ ,  $n \in \mathbb{N}_0$ ,  $\text{RrNGINAR}(1)$  модел дат дефиницијом 2.1.2 и нека су  $\mu_1 > 0$ ,  $\mu_2 > 0$ ,  $\dots$ ,  $\mu_r > 0$ . Тада су:*

- (а) *коваријансе између случајних променљивих  $X_n(z_n)$  и  $X_{n-k}(z_{n-k})$ , за  $k \in \{0, 1, \dots, n\}$  позитивне и дате формулом*

$$\gamma_n^{(k)} \equiv \text{Cov}(X_n(z_n), X_{n-k}(z_{n-k})) = \alpha^k \cdot \mu_{z_{n-k}}(1 + \mu_{z_{n-k}}),$$

- (б) *корелације између случајних променљивих  $X_n(z_n)$  и  $X_{n-k}(z_{n-k})$ , за  $k \in \{1, 2, \dots, n\}$  позитивне, мање од 1 и дате формулом*

$$\rho_n^{(k)} \equiv \text{Corr}(X_n(z_n), X_{n-k}(z_{n-k})) = \alpha^k \sqrt{\frac{\mu_{z_{n-k}}(1 + \mu_{z_{n-k}})}{\mu_{z_n}(1 + \mu_{z_n})}}.$$

*Доказ.* (а) Нека је  $\{X_n(z_n)\}$ ,  $n \in \mathbb{N}_0$ ,  $\text{RrNGINAR}(1)$  модел дат дефиницијом 2.1.2. Тада  $X_n(z_n)$  задовољава једначину (2.1.3), па је

$$\begin{aligned} \gamma_n^{(k)} &= \text{Cov}(X_n(z_n), X_{n-k}(z_{n-k})) \\ &= \text{Cov}(\alpha * X_{n-1}(z_{n-1}) + \varepsilon_n(z_{n-1}, z_n), X_{n-k}(z_{n-k})). \end{aligned}$$

Како су случајне променљиве  $\varepsilon_n(z_{n-1}, z_n)$  и  $X_{n-k}(z_{n-k})$  независне, добија се

$$\gamma_n^{(k)} = \text{Cov}(\alpha * X_{n-1}(z_{n-1}), X_{n-k}(z_{n-k})).$$

На основу познате особине негативног биномног тининг оператора да је  $\text{Cov}(\alpha * X, Y) = \alpha \text{Cov}(X, Y)$ , ако су случајне променљиве  $X$

и  $Y$  независне од случајних променљивих укључених у оператор негативног биномног тининг оператора, имамо да је

$$\gamma_n^{(k)} = \alpha \cdot Cov(X_{n-1}(z_{n-1}), X_{n-k}(z_{n-k})) = \alpha \cdot \gamma_{n-1}^{(k-1)}.$$

Примењујући последњу једнакост још  $k - 1$  пута, показује се да је коваријанса случајних променљивих  $X_n(z_n)$  и  $X_{n-k}(z_{n-k})$  дата са  $\gamma_n^{(k)} = \alpha^k \gamma_{n-k}^{(0)}$ . Како  $\gamma_{n-k}^{(0)}$  представља дисперзију случајне променљиве  $X_{n-k}(z_{n-k})$  која има геометријску расподелу са параметром очекивања  $\mu_{z_{n-k}}$ , добија се  $\gamma_n^{(k)} = \alpha^k \cdot \mu_{z_{n-k}}(1 + \mu_{z_{n-k}})$ .

(б) Корелација се може представити помоћу одговарајуће коваријансе као

$$\rho_n^{(k)} = \frac{\gamma_n^{(k)}}{\sqrt{\gamma_n^{(0)} \gamma_{n-k}^{(0)}}}.$$

Сада, користећи резултат (а), корелација случајних променљивих  $X_n(z_n)$  и  $X_{n-k}(z_{n-k})$  дата је са

$$\rho_n^{(k)} = \alpha^k \sqrt{\frac{\mu_{z_{n-k}}(1 + \mu_{z_{n-k}})}{\mu_{z_n}(1 + \mu_{z_n})}}.$$

Докажимо сада да су корелације  $\rho_n^{(k)}$  увек мање од 1. На основу неједнакости (2.2.7), за  $i = z_{n-k}$  и  $j = z_n$ , имамо да је

$$\alpha < \frac{\mu_{z_n}}{1 + \mu_{z_{n-k}}} < \frac{\mu_{z_n}}{\mu_{z_{n-k}}} < \frac{1 + \mu_{z_n}}{\mu_{z_{n-k}}},$$

одакле следи да је

$$\alpha < \sqrt{\frac{\mu_{z_n}(1 + \mu_{z_n})}{\mu_{z_{n-k}}(1 + \mu_{z_{n-k}})}}. \quad (2.2.8)$$

Пошто је  $\alpha < 1$ , то је  $\alpha^k < \alpha$ , па из (2.2.8) следи

$$\rho_n^{(k)} < \alpha \sqrt{\frac{\mu_{z_{n-k}}(1 + \mu_{z_{n-k}})}{\mu_{z_n}(1 + \mu_{z_n})}} < 1. \quad \square$$

Регресиони потенцијал сваког модела, односно могућност апроксимирања наредне, па и непознате, вредности процеса огледа се кроз особине, као и саме изразе који одређују условне величине. Са овим нас упознаје наредно тврђење.

**Теорема 2.2.4** Нека је  $\{X_n(z_n)\}$ ,  $n \in \mathbb{N}_0$ ,  $\text{RrNGINAR}(1)$  модел дат дефиницијом 2.1.2 и нека су  $\mu_1 > 0$ ,  $\mu_2 > 0$ ,  $\dots$ ,  $\mu_r > 0$ . Тада:

(а) условно очекивање случајне променљиве  $X_{n+k}(z_{n+k})$  за дато  $X_n(z_n)$  је

$$E(X_{n+k}(z_{n+k})|X_n(z_n)) = \alpha^k [X_n(z_n) - \mu_{z_n}] + \mu_{z_{n+k}}, \quad k \in \mathbb{N}_0, \quad (2.2.9)$$

(б) условна дисперзија случајне променљиве  $X_{n+k}(z_{n+k})$  за дато  $X_n(z_n)$  је

$$D(X_{n+k}(z_{n+k})|X_n(z_n)) = \frac{\alpha^k(1-\alpha^k)(1+\alpha)}{1-\alpha} [X_n(z_n) - \mu_{z_n}] + \mu_{z_{n+k}}(1 + \mu_{z_{n+k}}) - \alpha^{2k} \mu_{z_n}(1 + \mu_{z_n}).$$

*Доказ.* (а) У циљу једноставнијег означавања, нека је  $\mu_{n+k|n} = E(X_{n+k}(z_{n+k})|X_n(z_n))$  и  $\mu_{\varepsilon_n} = E(\varepsilon_n(z_{n-1}, z_n))$ . На основу дефиниције  $\text{RrNGINAR}(1)$  модела и независности случајних променљивих  $\varepsilon_{n+k}(z_{n+k-1}, z_{n+k})$  и  $X_{n+k-1}(z_{n+k-1})$ , следи да је

$$\begin{aligned} \mu_{n+k|n} &= E(\alpha * X_{n+k-1}(z_{n+k-1}) + \varepsilon_{n+k}(z_{n+k-1}, z_{n+k})|X_n(z_n)) \\ &= E(\alpha * X_{n+k-1}(z_{n+k-1})|X_n(z_n)) + E(\varepsilon_{n+k}(z_{n+k-1}, z_{n+k})). \end{aligned}$$

Познато је да важи  $E(\alpha * X|Y) = \alpha E(X|Y)$  за случајне променљиве  $X$  и  $Y$  које су независне од бројачког низа укљученог у оператор, па имамо коначно да је  $\mu_{n+k|n} = \alpha \mu_{n+k-1|n} + \mu_{\varepsilon_{n+k}}$ . Примењујући последњу једначину још  $k-1$  пута, добија се

$$\begin{aligned} \mu_{n+k|n} &= \alpha(\alpha \mu_{n+k-2|n} + \mu_{\varepsilon_{n+k-1}}) + \mu_{\varepsilon_{n+k}} \\ &= \alpha^k \mu_{n|n} + \sum_{l=0}^{k-1} \alpha^l \mu_{\varepsilon_{n+k-l}}. \end{aligned}$$

Ако сада заменимо  $\mu_{n|n} = X_n(z_n)$  и искористимо резултате теореме 2.2.2 за очекивање случајних променљивих  $\varepsilon_{n+k-l}$ ,  $l \in \{0, 1, \dots, k-1\}$ , добија се

$$\mu_{n+k|n} = \alpha^k X_n(z_n) + \sum_{l=0}^{k-1} \alpha^l (\mu_{z_{n+k-l}} - \alpha \mu_{z_{n+k-l-1}})$$

$$\begin{aligned}
&= \alpha^k X_n(z_n) + \sum_{l=0}^{k-1} \alpha^l \mu_{z_{n+k-l}} - \sum_{l=0}^{k-1} \alpha^{l+1} \mu_{z_{n+k-l-1}} \\
&= \alpha^k X_n(z_n) + \sum_{l=0}^{k-1} \alpha^l \mu_{z_{n+k-l}} - \sum_{l=1}^k \alpha^l \mu_{z_{n+k-l}} \\
&= \alpha^k X_n(z_n) + \mu_{z_{n+k}} - \alpha^k \mu_{z_n}.
\end{aligned}$$

(б) Посматрајмо сада условну дисперзију. Означимо са  $\sigma_{n+k|n}^2$  условну дисперзију случајне променљиве  $X_{n+k}(z_{n+k})$  за дато  $X_n(z_n)$ . Такође означимо  $D(\varepsilon_n(z_{n-1}, z_n))$  са  $\sigma_{\varepsilon_n}^2$ . На основу независности  $\varepsilon_{n+k}(z_{n+k-1}, z_{n+k})$  од  $X_n(z_n)$  и  $X_{n+k-1}(z_{n+k-1})$ , имамо да је

$$\begin{aligned}
E(X_{n+k}^2(z_{n+k})|X_n(z_n)) &= E([\alpha * X_{n+k-1}(z_{n+k-1})]^2 | X_n(z_n)) \\
&\quad + 2E(\alpha * X_{n+k-1}(z_{n+k-1}) \\
&\quad \cdot \varepsilon_{n+k}(z_{n+k-1}, z_{n+k}) | X_n(z_n)) \\
&\quad + E(\varepsilon_{n+k}^2(z_{n+k-1}, z_{n+k}) | X_n(z_n)) \\
&= E([\alpha * X_{n+k-1}(z_{n+k-1})]^2 | X_n(z_n)) \\
&\quad + 2E(\alpha * X_{n+k-1}(z_{n+k-1}) | X_n(z_n)) \\
&\quad \cdot E(\varepsilon_{n+k}(z_{n+k-1}, z_{n+k})) \\
&\quad + E(\varepsilon_{n+k}^2(z_{n+k-1}, z_{n+k})).
\end{aligned}$$

Тако је

$$\begin{aligned}
\sigma_{n+k|n}^2 &= \alpha^2 (\sigma_{n+k-1|n}^2 + \mu_{n+k-1|n}^2) + \alpha(1 + \alpha)\mu_{n+k-1|n} \\
&\quad + 2\alpha\mu_{n+k-1|n}\mu_{\varepsilon_{n+k}} + \sigma_{\varepsilon_{n+k}}^2 + \mu_{\varepsilon_{n+k}}^2 - \mu_{n+k|n}^2 \\
&= \alpha^2 \sigma_{n+k-1|n}^2 + \alpha(1 + \alpha)\mu_{n+k-1|n} + \sigma_{\varepsilon_{n+k}}^2 \\
&\quad + (\alpha\mu_{n+k-1|n} + \mu_{\varepsilon_{n+k}})^2 - \mu_{n+k|n}^2
\end{aligned}$$

због познате особине  $E(\alpha * X|Y) = \alpha E(X|Y)$  и  $E((\alpha * X)^2|Y) = \alpha^2 E(X^2|Y) + \alpha(1 + \alpha)E(X|Y)$ , те је

$$\begin{aligned}
E(X_{n+k}^2(z_{n+k})|X_n(z_n)) &= \alpha^2 E(X_{n+k-1}(z_{n+k-1})^2 | X_n(z_n)) \\
&\quad + \alpha(1 + \alpha)\mu_{n+k-1|n} \\
&\quad + 2\alpha\mu_{n+k-1|n}\mu_{\varepsilon_{n+k}} + \sigma_{\varepsilon_{n+k}}^2 + \mu_{\varepsilon_{n+k}}^2.
\end{aligned}$$

Узимајући у обзир да је  $\mu_{n+k|n} = \alpha\mu_{n+k-1|n} + \mu_{\varepsilon_{n+k}}$ , добијамо рекурентну релацију

$$\sigma_{n+k|n}^2 = \alpha^2 \sigma_{n+k-1|n}^2 + \alpha(1 + \alpha)\mu_{n+k-1|n} + \sigma_{\varepsilon_{n+k}}^2.$$



Применом ове једначине још  $k - 1$  пута, добија се

$$\sigma_{n+k|n}^2 = \alpha^{2k} \sigma_{n|n}^2 + \alpha(1 + \alpha) \sum_{l=0}^{k-1} \alpha^{2l} \mu_{n+k-1-l|n} + \sum_{l=0}^{k-1} \alpha^{2l} \sigma_{\varepsilon_{n+k-l}}^2.$$

Коначно, користећи чињеницу да је  $\sigma_{n|n}^2 = 0$ , резултат теореме 2.2.2 за дисперзију случајне променљиве  $\varepsilon_{n+k-l}$ ,  $l \in \{0, 1, \dots, k-1\}$  и једначину (2.2.9), добија се тражена релација.  $\square$

Због  $|\alpha| < 1$  је, за довољно велико  $k$ ,

$$E(X_{n+k}(z_{n+k})|X_n(z_n)) \approx \mu_{z_{n+k}} = E(X_{n+k}),$$

$$D(X_{n+k}(z_{n+k})|X_n(z_n)) \approx \mu_{z_{n+k}}(1 + \mu_{z_{n+k}}) = D(X_{n+k}),$$

где  $\approx$  означава “приближно”. Како  $z_{n+k}$  није константно, не можемо говорити о конвергенцији условних величина. Међутим, уколико посматрамо поднизове који одговарају истом стању, онда можемо сматрати да условно очекивање и условна дисперзија конвергирају ка очекивању и дисперзији маргиналне расподеле која одговара том стању.

### 2.3 Оцењивање непознатих параметара

У досадашњем тексту дисертације истакнуто је да INAR модели у случајној средини имају ту погодност да се “лако прилагођавају” подацима, што их чини применљивим у реалним ситуацијама. Ову своју особину дугују управо процесу случајне средине, те је од великог значаја сам поступак одређивања реализације  $\{z_n\}$ , који ћемо сада описати. Претпоставимо да располажемо подацима  $x_1, x_2, \dots, x_N$  које желимо да опишемо RiNGINAR(1) моделом. Присетимо се сада да вредност  $z_n$  одређује какву расподелу има случајна променљива  $X_n(z_n)$ . Због тога је очекивано да оне вредности  $x_i$ ,  $i \in \{1, 2, \dots, N\}$  које су међусобно блиске одговарају истом стању  $z_i$ . У складу са таквим закључивањем, може се очекивати да ће кластеровање података  $x_1, x_2, \dots, x_N$  у  $r$  различитих кластера заправо поделити податке на тај начин да они који припадају истом кластеру одговарају истом стању. Подсетимо се опет тога да

вредности скупа  $E_r$  које узимају елементи низа  $\{z_n\}$  саме по себи нису важне, пошто је њихова улога једино да повежу  $X_n(z_n)$  са одговарајућим параметром расподеле  $\mu_{z_n}$ . Из истог разлога није важно коју вредност из скупа  $E_r$  ћемо придружити којем кластеру. Дакле, можемо нумерисати кластере на произвољан начин, а затим за свако  $n \in \{1, 2, \dots, N\}$  сматрати да је  $z_n = i$  уколико  $x_n$  припада  $i$ -том кластеру. На овај начин је у потпуности одређен низ  $z_1, z_2, \dots, z_N$ .

### 2.3.1 Јул-Вокер оцене

Посматрајмо случајни узорак  $X_1(z_1), X_2(z_2), \dots, X_N(z_N)$  обима  $N$  из  $\text{RrNGINAR}(1)$  процеса  $\{X_n(z_n)\}$ . Нека су  $k$  и  $n$  природни бројеви, такви да је  $z_k \neq i$ ,  $z_{k+1} = z_{k+2} = \dots = z_n = i$ ,  $z_{n+1} \neq i$ . Ако посматрамо подузорак  $X_{k+1}(i), X_{k+2}(i), \dots, X_n(i)$ , можемо приметити да сви његови елементи одговарају стању  $i$ , као и да је максималан у смислу да се не може проширити ни левим ни десним суседом, а да се не дода елемент који не одговара стању  $i$ . Како сви елементи овог подузорка одговарају истом стању, он се може посматрати као узорак из  $\text{NGINAR}(1)$  процеса са очекивањем  $\mu_i$ . Пошто је  $\text{NGINAR}(1)$  процес стационаран, одговарајућа узорачка коваријанса је строго постојана. Стога има смисла користити управо ове коваријансе у оцењивању.

Нека су  $\hat{\mu}_i$ ,  $\hat{\gamma}_0^{(i)}$  и  $\hat{\gamma}_1^{(i)}$  редом узорачка средина, узорачка дисперзија и узорачка коваријанса првог реда подузорка  $X_{k+1}(i), X_{k+2}(i), \dots, X_n(i)$ , тј. нека су дефинисане на следећи начин

$$\hat{\mu}_i = \frac{1}{n-k} \sum_{j=1}^{n-k} X_{k+j}(i), \quad \hat{\gamma}_0^{(i)} = \frac{1}{n-k} \sum_{j=1}^{n-k} (X_{k+j}(i) - \hat{\mu}_i)^2,$$

$$\hat{\gamma}_1^{(i)} = \frac{1}{n-k} \sum_{j=1}^{n-k-1} (X_{k+j+1}(i) - \hat{\mu}_i)(X_{k+j}(i) - \hat{\mu}_i).$$

На основу ергодичности  $\text{NGINAR}(1)$  процеса следи строга постојаност ових оцена.

Вратимо се сада узорку  $X_1(z_1), X_2(z_2), \dots, X_N(z_N)$ . Релација “одговарати истом стању” представља релацију еквиваленције на скупу елемената узорка. Због тога се може извршити партиција

узорка у односу на ову релацију. За произвољно  $k \in \{1, 2, \dots, r\}$  нека је

$$\begin{aligned} I_k &= \{i \in \{1, 2, \dots, N\} | z_i = k\}, \\ \bigcup_{k=1}^r I_k &= \{1, 2, \dots, N\}, \quad |I_k| = n_k, \quad n_1 + n_2 + \dots + n_r = N, \\ S_k &= (X_{k_1}(k), X_{k_2}(k), \dots, X_{k_{n_k}}(k)), \end{aligned}$$

где је  $k_i \in I_k$ ,  $k_i < k_{i+1}$ ,  $\forall i \in \{1, 2, \dots, n_k - 1\}$ . Дакле,  $S_k$  је подузорак почетног узорка који садржи све елементе који одговарају стању  $k$  и ниједне друге. Може се приметити да се  $S_k$  заправо састоји од свих максималних подузорака који одговарају стању  $k$ . Нека су ови максимални подузорци  $S_{k,1}, S_{k,2}, \dots, S_{k,i_k}$ , где је  $i_k$  њихов број. Нека је даље  $J_{k,l} = \{i \in \{1, 2, \dots, N\} | X_i(z_i) \in S_{k,l}\}$ ,  $|J_{k,l}| = n_{k,l}$ , и  $n_{k,1} + n_{k,2} + \dots + n_{k,i_k} = n_k$ . Као што смо већ закључили, оцене добијене на основу узорака  $S_{k,l}$  су строго постојане. У том циљу, нека је

$$\begin{aligned} \hat{\mu}_{k,l} &= \frac{1}{n_{k,l}} \sum_{i \in J_{k,l}} X_i(k), \quad \hat{\gamma}_{0,l}^{(k)} = \frac{1}{n_{k,l}} \sum_{i \in J_{k,l}} (X_i(k) - \hat{\mu}_{k,l})^2, \\ \hat{\gamma}_{1,l}^{(k)} &= \frac{1}{n_{k,l}} \sum_{\{i,i+1\} \subseteq J_{k,l}} (X_{i+1}(k) - \hat{\mu}_{k,l})(X_i(k) - \hat{\mu}_{k,l}). \end{aligned}$$

Дефинишимо сада оцене на основу целих подузорака  $S_k$ , не узимајући у обзир то да се они састоје од делова који нису повезани, тј. из максималних подузорака.

**ДЕФИНИЦИЈА 2.3.1** *Оцене добијене на основу подузорка  $S_k$  који одговара стању  $k$  дефинисане су следећим изразима*

$$\begin{aligned} \hat{\mu}_k &= \frac{1}{n_k} \sum_{i \in I_k} X_i(k), \quad \hat{\gamma}_0^{(k)} = \frac{1}{n_k} \sum_{i \in I_k} (X_i(k) - \hat{\mu}_k)^2, \\ \hat{\gamma}_1^{(k)} &= \frac{1}{n_k} \sum_{\{i,i+1\} \subseteq I_k} (X_{i+1}(k) - \hat{\mu}_k)(X_i(k) - \hat{\mu}_k). \end{aligned} \quad (2.3.1)$$

Важи следеће

$$\begin{aligned}\widehat{\mu}_k &= \frac{1}{n_k} \sum_{i \in I_k} X_i(k) = \frac{1}{n_k} \sum_{l=1}^{i_k} \sum_{i \in J_{k,l}} X_i(k) \\ &= \sum_{l=1}^{i_k} \frac{n_{k,l}}{n_k} \frac{1}{n_{k,l}} \sum_{i \in J_{k,l}} X_i(k) = \sum_{l=1}^{i_k} \frac{n_{k,l}}{n_k} \widehat{\mu}_{k,l}.\end{aligned}$$

Оцене  $\widehat{\mu}_{k,l}$  су строго постојане, тј.  $P(\lim_{n_{k,l} \rightarrow \infty} \widehat{\mu}_{k,l} = \mu_k) = 1$ . Дакле,  $\lim_{n_{k,l} \rightarrow \infty} \widehat{\mu}_{k,l} = \mu_k$  свугде осим на скупу  $\Omega_{k,l}$ , где је  $P(\Omega_{k,l}) = 0$ . Стога је  $\widehat{\mu}_{k,l} = \mu_k + o(n_{k,l})$ ,  $n_{k,l} \rightarrow \infty$ , свугде осим на скупу  $\Omega_{k,l}$ .

На основу претходног разматрања, можемо записати израз за  $\widehat{\mu}_k$  на следећи начин

$$\begin{aligned}\widehat{\mu}_k &= \sum_{l=1}^{i_k} \frac{n_{k,l}}{n_k} (\mu_k + o(n_{k,l})) = \sum_{l=1}^{i_k} \frac{n_{k,l}}{n_k} \mu_k + \sum_{l=1}^{i_k} \frac{n_{k,l}}{n_k} o(n_{k,l}) \\ &= \mu_k \frac{\sum_{l=1}^{i_k} n_{k,l}}{n_k} + \sum_{l=1}^{i_k} \frac{n_{k,l}}{n_k} o(n_{k,l}).\end{aligned}$$

Како је  $\frac{\sum_{l=1}^{i_k} n_{k,l}}{n_k} = 1$  и  $\forall l \in \{1, 2, \dots, i_k\} \lim_{n_{k,l} \rightarrow \infty} \frac{n_{k,l}}{n_k} < \infty$ , важи  $\widehat{\mu}_k \rightarrow \mu_k$ ,  $n_{k,i} \rightarrow \infty$ ,  $\forall i \in \{1, 2, \dots, i_k\}$  свугде осим на скупу  $\Omega_k = \cup_{l=1}^{i_k} \Omega_{k,l}$ .

На основу  $P(\Omega_k) = P(\cup_{l=1}^{i_k} \Omega_{k,l}) \leq \sum_{l=1}^{i_k} P(\Omega_{k,l})$  и особине ненегативности вероватноће, имамо да је  $P(\Omega_k) = 0$ . Сада је

$$P(\widehat{\mu}_k \rightarrow \mu_k, n_{k,i} \rightarrow \infty, \forall i \in \{1, 2, \dots, i_k\}) = 1.$$

Ово тврђење јако подсећа на строго постојаност. Наиме, једина разлика је у томе што строга постојаност захтева конвергенцију када обим узорка тежи бесконачности, док у нашем случају имамо конвергенцију када обими свих максималних подузорака теже бесконачности. Ипак, могуће је показати и строго постојаност у смислу тачне дефиниције. Овим се бави следећа теорема.

**Теорема 2.3.1** *Оцене  $\widehat{\mu}_k, \widehat{\gamma}_0^{(k)}, \widehat{\gamma}_1^{(k)}$  дате дефиницијом 2.3.1 су строго постојане.*

*Доказ.* Докажимо најпре да је  $\hat{\mu}_k$  строго постојана оцена. У том циљу треба показати да је  $P(\hat{\mu}_k \rightarrow \mu_k, n_k \rightarrow \infty) = 1$ . Важи  $n_k = n_{k,1} + n_{k,2} + \dots + n_{k,i_k}$ . Нека  $n_k \rightarrow \infty$ . Ако су обими  $n_{k,j}$  свих подузорака коначни, тада је и  $n_k$  коначно. Дакле, да би  $n_k$  тежило бесконачности, исто мора да важи бар за једно  $n_{k,j}$ . Претпоставимо да је за неко  $d \in \{1, 2, \dots, i_k\}$ ,

$$\begin{aligned} n_{k,l} &\rightarrow \infty, \quad \text{за свако } l \in \{1, 2, \dots, d\}, \\ n_{k,j} &\rightarrow c_j < \infty, \quad \text{за свако } j \in \{d+1, d+2, \dots, i_k\}. \end{aligned} \quad (2.3.2)$$

Уколико ово не важи, могу се пренумерисати подузорци. Сада имамо

$$\hat{\mu}_k = \sum_{l=1}^{i_k} \frac{n_{k,l}}{n_k} \hat{\mu}_{k,l} = \sum_{l=1}^d \frac{n_{k,l}}{n_k} \hat{\mu}_{k,l} + \sum_{l=d+1}^{i_k} \frac{n_{k,l}}{n_k} \hat{\mu}_{k,l}.$$

Посматрајмо коефицијенте из последње једнакости. На основу (2.3.2) је

$$\lim_{n_k \rightarrow \infty} \frac{n_{k,j}}{n_k} = 0, \quad \text{за } j \in \{d+1, d+2, \dots, i_k\}.$$

Због тога је

$$\lim_{n_k \rightarrow \infty} \hat{\mu}_k = \lim_{n_k \rightarrow \infty} \sum_{l=1}^d \frac{n_{k,l}}{n_k} \hat{\mu}_{k,l},$$

где је

$$n_k = n_{k,1} + n_{k,2} + \dots + n_{k,i_k} = \sum_{l=1}^d n_{k,l} + \sum_{j=d+1}^{i_k} n_{k,j} \rightarrow \sum_{l=1}^d n_{k,l} + \sum_{j=d+1}^{i_k} c_j,$$

када  $n_k \rightarrow \infty$ . Сума  $\sum_{j=d+1}^{i_k} c_j$  је коначна, па је занемарљива у односу на  $n_{k,l}$ , за  $l \in \{1, 2, \dots, d\}$ . Сада можемо записати  $n_k = n_{k,1} + n_{k,2} + \dots + n_{k,d}$  када посматрамо граничну вредност при  $n_k \rightarrow \infty$ .

Због (2.3.2) је

$$\left( n_k \rightarrow \infty \right) \iff \left( n_{k,l} \rightarrow \infty \quad \forall l \in \{1, 2, \dots, d\} \right).$$

Сада имамо

$$\lim_{n_k \rightarrow \infty} \hat{\mu}_k = \lim_{n_{k,l} \rightarrow \infty, \forall l \in \{1, 2, \dots, d\}} \hat{\mu}_k = \mu_k.$$

Коначно, важи  $\hat{\mu}_k \rightarrow \mu_k, n_k \rightarrow \infty$ , свугде осим на скупу  $\Omega_k = \cup_{l=1}^{i_k} \Omega_{k,l}$ . Ово можемо записати на следећи начин  $P(\hat{\mu}_k \rightarrow \mu_k, n_k \rightarrow \infty) = 1$ , што заправо представља строгу постојаност узорачке средине.

Сада ћемо доказану строгу постојаност узорачке средине користити за доказивање строге постојаности узорачке коваријансе. Како је  $\hat{\gamma}_{0,l}^{(k)}$  строго постојана за све  $l \in \{1, 2, \dots, i_k\}$ , важи  $\hat{\gamma}_{0,l}^{(k)} \rightarrow \gamma_0^{(k)}, n_{k,l} \rightarrow \infty$  свугде осим на скупу  $\Omega'_{k,l}$  мере нула. Како смо доказали да је  $\hat{\mu}_k$  строго постојана, имамо да је  $\hat{\mu}_k \rightarrow \mu_k, n_k \rightarrow \infty$ , свугде сем на скупу  $\Omega_k$ , где је  $P(\Omega_k) = 0$ . Ако покажемо следећу импликацију

$$\begin{aligned} & \left( (\hat{\gamma}_{0,l}^{(k)} \rightarrow \gamma_0^{(k)}, n_{k,l} \rightarrow \infty (\forall l \in \{1, 2, \dots, i_k\})) \wedge (\hat{\mu}_k \rightarrow \mu_k, n_k \rightarrow \infty) \right) \\ & \Rightarrow \left( \hat{\gamma}_0^{(k)} \rightarrow \gamma_0^{(k)}, n_k \rightarrow \infty \right), \end{aligned} \quad (2.3.3)$$

тада ће бити  $\hat{\gamma}_0^{(k)} \rightarrow \gamma_0^{(k)}, n_k \rightarrow \infty$  свугде осим на скупу  $\Omega'_k = \cup_{l=1}^{i_k} \Omega'_{k,l} \cup \Omega_k$ , при чему  $P(\Omega_k) = 0$ , па ће важити  $P(\hat{\gamma}_0^{(k)} \rightarrow \gamma_0^{(k)}, n_k \rightarrow \infty) = 1$ , што је еквивалентно строгој постојаности оцене  $\hat{\gamma}_0^{(k)}$ . Стога можемо доказ свести на доказивање поменуते импликације.

Имамо да је

$$\begin{aligned} \hat{\gamma}_0^{(k)} &= \frac{1}{n_k} \sum_{i \in I_k} (X_i(k) - \hat{\mu}_k)^2 = \frac{1}{n_k} \sum_{l=1}^{i_k} \sum_{i \in J_{k,l}} (X_i(k) - \hat{\mu}_k)^2 \\ &= \sum_{l=1}^{i_k} \frac{n_{k,l}}{n_k} \frac{1}{n_{k,l}} \sum_{i \in J_{k,l}} (X_i(k) - \hat{\mu}_k)^2 = \sum_{l=1}^{i_k} \frac{n_{k,l}}{n_k} \hat{\gamma}_0^{(k)}. \end{aligned}$$

Нека је  $n_k \rightarrow \infty$ . Понављајући исти поступак као малопре, претпоставимо да важе услови (2.3.2). Овим се претходна сума своди на збир само првих  $d$  сабирака

$$\hat{\gamma}_0^{(k)} = \sum_{l=1}^d \frac{n_{k,l}}{n_k} \hat{\gamma}_0^{(k)}.$$

Даљим модификовањем ове суме добија се

$$\begin{aligned}\widehat{\gamma}_0^{(k)} &= \sum_{l=1}^d \frac{n_{k,l}}{n_k} \frac{1}{n_{k,l}} \sum_{i \in J_{k,l}} (X_i(k) - \widehat{\mu}_k)^2 \\ &= \sum_{l=1}^d \frac{n_{k,l}}{n_k} \frac{1}{n_{k,l}} \sum_{i \in J_{k,l}} (X_i(k) - \widehat{\mu}_{k,l} + (\widehat{\mu}_{k,l} - \widehat{\mu}_k))^2.\end{aligned}$$

Притом, изразе  $\widehat{\mu}_{k,l} \rightarrow \mu_k$ ,  $n_{k,l} \rightarrow \infty$ ,  $\forall l \in \{1, 2, \dots, i_k\}$  и  $\widehat{\mu}_k \rightarrow \mu_k$ ,  $n_k \rightarrow \infty$ , можемо представити на другачији начин  $\widehat{\mu}_{k,l} = \mu_k + o(n_{k,l})$ , када  $n_{k,l} \rightarrow \infty$ ,  $\forall l \in \{1, 2, \dots, i_k\}$  и  $\widehat{\mu}_k = \mu_k + o(n_k)$ ,  $n_k \rightarrow \infty$ . Зато можемо записати  $\widehat{\mu}_{k,l} - \widehat{\mu}_k = o(n_{k,l})$ ,  $n_{k,l} \rightarrow \infty$  и

$$\begin{aligned}\frac{1}{n_{k,l}} \sum_{i \in J_{k,l}} (X_i(k) - \widehat{\mu}_{k,l} + o(n_{k,l}))^2 &= \frac{1}{n_{k,l}} \sum_{i \in J_{k,l}} (X_i(k) - \widehat{\mu}_{k,l})^2 \\ &\quad + \frac{1}{n_{k,l}} \sum_{i \in J_{k,l}} o(n_{k,l}^2) \\ &\quad + \frac{2}{n_{k,l}} \sum_{i \in J_{k,l}} (X_i(k) - \widehat{\mu}_{k,l}) \cdot o(n_{k,l}).\end{aligned}$$

Последњи члан је једнак нули због

$$\frac{1}{n_{k,l}} \sum_{i \in J_{k,l}} (X_i(k) - \widehat{\mu}_{k,l}) = \frac{1}{n_{k,l}} \sum_{i \in J_{k,l}} X_i(k) - \widehat{\mu}_{k,l} = 0.$$

Други члан је

$$\frac{1}{n_{k,l}} \sum_{i \in J_{k,l}} o(n_{k,l}^2) = \frac{1}{n_{k,l}} o(n_{k,l}^2) |J_{k,l}| = \frac{1}{n_{k,l}} o(n_{k,l}^2) n_{k,l} = o(n_{k,l}^2),$$

док је први

$$\frac{1}{n_{k,l}} \sum_{i \in J_{k,l}} (X_i(k) - \widehat{\mu}_{k,l})^2 = \widehat{\gamma}_{0,l}^{(k)}.$$

Дакле,

$$\begin{aligned}\widehat{\gamma}_0^{(k)} &= \sum_{l=1}^d \frac{n_{k,l}}{n_k} (\widehat{\gamma}_{0,l}^{(k)} + o(n_{k,l}^2)) = \sum_{l=1}^d \frac{n_{k,l}}{n_k} (\gamma_0^{(k)} + o(n_{k,l}) + o(n_{k,l}^2)) \\ &= \gamma_0^{(k)} + \sum_{l=1}^d \frac{n_{k,l}}{n_k} o(n_{k,l}).\end{aligned}$$

Када  $n_k \rightarrow \infty$ , онда  $\frac{n_{k,l}}{n_k} \rightarrow c_l < \infty$  и  $o(n_{k,l}) \rightarrow 0, \forall l \in \{1, 2, \dots, d\}$ , па  $\widehat{\gamma}_0^{(k)} \rightarrow \gamma_0^{(k)}$ . Овим смо доказали импликацију (2.3.3), а самим тим и строгу постојаност оцене  $\widehat{\gamma}_0^{(k)}$ .

Доказ за  $\widehat{\gamma}_1^{(k)}$  је аналоган. Довољно је доказати

$$\begin{aligned} & \left( (\widehat{\gamma}_{1,l}^{(k)} \rightarrow \gamma_{1,l}^{(k)}, n_{k,l} \rightarrow \infty (\forall l \in \{1, 2, \dots, i_k\})) \wedge (\widehat{\mu}_k \rightarrow \mu_k, n_k \rightarrow \infty) \right) \\ & \Rightarrow \left( \widehat{\gamma}_1^{(k)} \rightarrow \gamma_1^{(k)}, n_k \rightarrow \infty \right). \end{aligned}$$

Сада понављамо поступак

$$\begin{aligned} \widehat{\gamma}_1^{(k)} &= \sum_{l=1}^d \frac{n_{k,l}}{n_k} \frac{1}{n_{k,l}} \sum_{\{i,i+1\} \subseteq J_{k,l}} (X_{i+1}(k) - \widehat{\mu}_k)(X_i(k) - \widehat{\mu}_k) \\ &= \sum_{l=1}^d \frac{n_{k,l}}{n_k} \frac{1}{n_{k,l}} \cdot \\ & \quad \sum_{\{i,i+1\} \subseteq J_{k,l}} (X_{i+1}(k) - \widehat{\mu}_{k,l} + (\widehat{\mu}_{k,l} - \widehat{\mu}_k))(X_i(k) - \widehat{\mu}_{k,l} + (\widehat{\mu}_{k,l} - \widehat{\mu}_k)) \\ &= \sum_{l=1}^d \frac{n_{k,l}}{n_k} \frac{1}{n_{k,l}} \cdot \\ & \quad \sum_{\{i,i+1\} \subseteq J_{k,l}} (X_{i+1}(k) - \widehat{\mu}_{k,l} + o(n_{k,l}))(X_i(z_i) - \widehat{\mu}_{k,l} + o(n_{k,l})) \end{aligned}$$

$$\begin{aligned} \widehat{\gamma}_1^{(k)} &= \sum_{l=1}^d \frac{n_{k,l}}{n_k} \frac{1}{n_{k,l}} \sum_{\{i,i+1\} \subseteq J_{k,l}} (X_{i+1}(k) - \widehat{\mu}_{k,l})(X_i(k) - \widehat{\mu}_{k,l}) \\ & \quad + \sum_{l=1}^d \frac{n_{k,l}}{n_k} \frac{1}{n_{k,l}} \sum_{\{i,i+1\} \subseteq J_{k,l}} (X_{i+1}(k) - \widehat{\mu}_{k,l}) \cdot o(n_{k,l}) \\ & \quad + \sum_{l=1}^d \frac{n_{k,l}}{n_k} \frac{1}{n_{k,l}} \sum_{\{i,i+1\} \subseteq J_{k,l}} o(n_{k,l}) \cdot (X_i(k) - \widehat{\mu}_{k,l}) \\ & \quad + \sum_{l=1}^d \frac{n_{k,l}}{n_k} \frac{1}{n_{k,l}} \sum_{\{i,i+1\} \subseteq J_{k,l}} o(n_{k,l}^2) \end{aligned}$$



$$\begin{aligned}
\widehat{\gamma}_1^{(k)} &= \sum_{l=1}^d \frac{n_{k,l}}{n_k} \frac{1}{n_{k,l}} \sum_{\{i,i+1\} \subseteq J_{k,l}} (X_{i+1}(k) - \widehat{\mu}_{k,l})(X_i(k) - \widehat{\mu}_{k,l}) \\
&\quad + \sum_{l=1}^d \frac{n_{k,l}}{n_k} o(n_{k,l}^2) \\
&= \sum_{l=1}^d \frac{n_{k,l}}{n_k} (\widehat{\gamma}_{1,l}^{(k)} + o(n_{k,l}^2)) \\
&= \sum_{l=1}^d \frac{n_{k,l}}{n_k} (\gamma_1^{(k)} + o(n_{k,l}) + o(n_{k,l}^2)) = \gamma_1^{(k)} + \sum_{l=1}^d \frac{n_{k,l}}{n_k} o(n_{k,l}).
\end{aligned}$$

Дакле,  $\widehat{\gamma}_1^{(k)} \rightarrow \gamma_1^{(k)}$ ,  $n_k \rightarrow \infty$ , чиме је доказ завршен.  $\square$

Једини преостао неоцењен параметар је  $\alpha$ . Њега ћемо оценити на следећи начин

$$\widehat{\alpha} = \sum_{k=1}^r \frac{n_k}{N} \widehat{\alpha}_k, \quad \text{где је} \quad \widehat{\alpha}_k = \frac{\widehat{\gamma}_1^{(k)}}{\widehat{\gamma}_0^{(k)}}.$$

На основу дефиниције строге постојаности, субадитивности вероватноће и особина граничне вредности, лако се показује да су  $\widehat{\alpha}_k$  и  $\widehat{\alpha}$  строго постојане оцене.

### 2.3.2 Оцене условне максималне веродостојности

Нека је  $X_1(z_1), X_2(z_2), \dots, X_N(z_N)$  случајни узорак обима  $N$  из  $\text{RrNGINAR}(1)$  процеса. Сада ћемо дефинисати функцију логаритамске условне веродостојности у циљу оцењивања непознатих параметара модела  $\mu_1, \mu_2, \dots, \mu_r, \alpha$ . Како се ради о ланцу Маркова првог реда, очигледно је

$$\begin{aligned}
\log L &= \sum_{i=2}^N \log P(X_i(z_i) = x_i | X_{i-1}(z_{i-1}) = x_{i-1}) \\
&= \sum_{i=2}^N \log P(\alpha * X_{n-1}(z_{n-1}) + \varepsilon_n(z_{n-1}, z_n) = x_i | X_{i-1}(z_{i-1}) = x_{i-1}) \\
&= \sum_{i=2}^N \log \left\{ \sum_{k=0}^{x_i} P(\alpha * X_{n-1}(z_{n-1}) = k) P(\varepsilon_n(z_{n-1}, z_n) = x_i - k) \right\}
\end{aligned}$$

$$\cdot I_{\{x_{i-1} \neq 0\}} + P(\varepsilon_n(z_{n-1}, z_n) = x_i) I_{\{x_{i-1} = 0\}} \Big\}.$$

Како је, на основу расподеле случајне променљиве  $\varepsilon_n(z_{n-1}, z_n)$ ,

$$P(\varepsilon_n(z_{n-1}, z_n) = x) = \left(1 - \frac{\alpha \mu_{z_{i-1}}}{\mu_{z_i} - \alpha}\right) \frac{\mu_{z_i}^x}{(1 + \mu_{z_i})^{x+1}} + \frac{\alpha \mu_{z_{i-1}}}{\mu_{z_i} - \alpha} \cdot \frac{\alpha^x}{(1 + \alpha)^{x+1}},$$

то је

$$\log L = \sum_{i=2}^N \log \left\{ \sum_{k=0}^{x_i} D_i [(1 - C_i)C(x_i - k) + C_i C(x_i - k)] I_{\{x_{i-1} \neq 0\}} \right. \\ \left. + [(1 - C_i)C(x_i) + C_i C(x_i)] I_{\{x_{i-1} = 0\}} \right\},$$

при чему су коришћене следеће ознаке  $D_i = \binom{x_{i-1} + k - 1}{x_{i-1} - 1} \frac{\alpha^k}{(1 + \alpha)^{k + x_{i-1}}}$ ,  $C_i = \frac{\alpha \mu_{z_{i-1}}}{\mu_{z_i} - \alpha}$ ,  $C(x) = \frac{\mu_{z_i}^x}{(1 + \mu_{z_i})^{x+1}}$ . Због сложености ове функције, у пракси се оцене методом условне максималне веродостојности могу добити коришћењем неких нумеричких алгоритама за максимизирање ове функције који су саставни део многих статистичких софтверских пакета.

### 2.3.3 Анализа квалитета оцена над симулираним подацима

Сада ћемо испитати квалитет Јул-Вокер (YW) оцена, као и оцена условне максималне веродостојности (CML), дефинисаних у претходним подсекцијама. Наиме, циљ је показати да оцене непознатих параметара у пракси конвергирају ка њиховим тачним вредностима при повећању обима узорка. Ово се постиже на тај начин што се симулирају узорци жељених обима из процеса који посматрамо и над таквим узорцима се онда израчунају реализације оцена. У том циљу смо симулирали узорак обима 500 у 10000 понављања. Наравно, потребно је прво симулирати процес случајне средине, а затим његовим коришћењем и одговарајући INAR процес. У пракси је често довољно посматрати процесе са два или три различита стања, те су ови случајеви засебно разматрани у даљем тексту. Притом је потребно задати следеће параметре:

- број стања  $r$
- вектор очекивања  $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_r)'$
- параметар тининг оператора  $\alpha$
- вектор  $\mathbf{p}_{vec}$  дужине  $r$  чије компоненте су вероватноће са којима се бирају иницијална стања
- матрица вероватноћа прелаза  $\mathbf{p}_{mat}$  димензија  $r \times r$  која се дефинише за ланце Маркова првог реда и садржи вероватноће којима се у једном кораку прелази из једног стања у друго, тј.  $[\mathbf{p}_{mat}]_{ij}$  представља вероватноћу са којом ће процес из стања  $i$  прећи у стање  $j$  у једном кораку.

## 1. Случај са три стања

У овом случају сматрамо да процес случајне средине узима вредности из скупа  $E_3 = \{1, 2, 3\}$ . Да бисмо видели на који начин вредности параметара утичу на понашање оцена, разликоваћемо три различита и можемо рећи репрезентативна случаја.

1.1. Вектор очекивања је  $\boldsymbol{\mu} = (1, 2, 3)'$ . На основу теореме 2.2.1 сада можемо израчунати границу коју  $\alpha$  не сме да пређе и она износи 0,25, па бирамо  $\alpha = 0,2$ . Одабрали смо  $\mathbf{p}_{vec} = (0,33, 0,34, 0,33)'$  тако да се у првом кораку стања бирају са скоро једнаким вероватноћама. Што се тиче матрице вероватноћа прелаза, она је јако значајна с обзиром на то да одређује изглед процеса случајне средине, а самим тим и RrNGINAR(1) процеса. У овом случају је

$$\mathbf{p}_{mat} = \begin{bmatrix} 0,4 & 0,3 & 0,3 \\ 0,3 & 0,4 & 0,3 \\ 0,3 & 0,3 & 0,4 \end{bmatrix}.$$

Вредности на дијагонали представљају вероватноће остајања у истом стању. Видимо да су оне само незнатно веће од вероватноћа да се из једног стања пређе у неко од друга два стања.

- 1.2. Овде је коришћен вектор очекивања  $\boldsymbol{\mu} = (4, 5, 6)'$ . Овакав избор допушта већу корелацију, тј. вишу границу за  $\alpha$ , те бирамо  $\alpha = 0,5$ . Осим тога, претпостављамо да је у почетном тренутку друго стање вероватније од осталих, што је одређено са  $\mathbf{p}_{vec} = (0,2, 0,6, 0,2)'$ . За матрицу прелаза је одабрана

$$\mathbf{p}_{mat} = \begin{bmatrix} 0,5 & 0,4 & 0,1 \\ 0,3 & 0,4 & 0,3 \\ 0,1 & 0,4 & 0,5 \end{bmatrix}.$$

И у овом случају су највеће вероватноће остајања у истом стању, с тим што је вероватноћа да се из стања које је “на једном крају” дође у стање “на другом крају” најмања (тј. вероватноћа прелаза из првог у треће или из трећег у прво стање).

- 1.3. У последњем случају је  $\boldsymbol{\mu} = (1, 2, 8)'$ . Овакав избор вектора очекивања имплицира нижу границу за  $\alpha$ , те бирамо  $\alpha = 0,1$ . За вектор иницијалних вероватноћа је изабран  $\mathbf{p}_{vec} = (0,6, 0,3, 0,1)'$ , док је матрица вероватноћа прелаза

$$\mathbf{p}_{mat} = \begin{bmatrix} 0,1 & 0,6 & 0,3 \\ 0,2 & 0,3 & 0,5 \\ 0,1 & 0,3 & 0,6 \end{bmatrix}.$$

Ако се процес налази у првом, највероватније ће прећи у друго стање, а ако је у другом, највероватније ће прећи у треће стање. Када доспе у треће стање, највероватније је да у њему остане, што значи да има тенденцију ка већим вредностима.

## 2. Случај са два стања

И у овом случају посматраћемо три различите комбинације параметара.

- 2.1. Најпре ћемо за вектор очекивања одабрати  $\boldsymbol{\mu} = (1, 2)'$  и у складу са границом коју он одређује  $\alpha = 0,3$ . Притом за  $\mathbf{p}_{vec}$  бирамо  $(0,5, 0,5)'$ , што значи да се иницијално стања бирају

Табела 2.1: YW оцене за случај 1.1.

$n_1$	$\hat{\mu}_1^{YW}$	$\hat{\mu}_2^{YW}$	$\hat{\mu}_3^{YW}$	$\hat{\alpha}^{YW}$
100	0,99589	2,00557	3,00441	0,19408
ст. гр.	2,68695	4,69168	6,65246	1,39636
200	0,99119	2,00586	3,00355	0,19701
ст. гр.	1,34039	2,32784	3,31118	0,67834
500	0,99861	2,00232	2,99752	0,19863
ст. гр.	0,53974	0,92978	1,31754	0,26985

Табела 2.2: YW оцене за случај 1.2.

$n_1$	$\hat{\mu}_1^{YW}$	$\hat{\mu}_2^{YW}$	$\hat{\mu}_3^{YW}$	$\hat{\alpha}^{YW}$
100	4,00237	5,01148	5,99324	0,49084
ст. гр.	11,22251	11,50781	15,97502	1,04790
200	4,00264	5,01621	6,00560	0,49456
ст. гр.	5,54954	5,73101	8,00416	0,51069
500	4,00129	5,00423	6,00566	0,49676
ст. гр.	2,23714	2,30200	3,23006	0,20601

Табела 2.3: YW оцене за случај 1.3.

$n_1$	$\hat{\mu}_1^{YW}$	$\hat{\mu}_2^{YW}$	$\hat{\mu}_3^{YW}$	$\hat{\alpha}^{YW}$
100	0,99854	1,99324	7,99212	0,09723
ст. гр.	3,98321	4,31584	12,37995	4,46355
200	1,00127	1,99868	7,99392	0,09892
ст. гр.	1,98700	2,17395	6,15311	2,04681
500	0,99981	2,00035	7,99983	0,10001
ст. гр.	0,78869	0,86622	2,44374	0,75179

Табела 2.4: CML оцене за случај 1.1.

$n_1$	$\hat{\mu}_1^{CML}$	$\hat{\mu}_2^{CML}$	$\hat{\mu}_3^{CML}$	$\hat{\alpha}^{CML}$
100	0,97372	2,00921	3,03007	0,21393
ст. гр.	2,86161	4,74395	5,70047	0,94502
200	0,99188	2,00378	3,01200	0,20512
ст. гр.	1,34182	2,32191	2,78148	0,40580
500	0,99721	2,00090	3,00001	0,20079
ст. гр.	0,51135	0,92146	1,11075	0,15303

Табела 2.5: YW оцене за случај 1.2.

$n_1$	$\hat{\mu}_1^{CML}$	$\hat{\mu}_2^{CML}$	$\hat{\mu}_3^{CML}$	$\hat{\alpha}^{CML}$
100	3,89194	4,99390	6,04319	0,51053
ст. гр.	8,85908	8,42631	8,94478	0,74102
200	3,98092	5,00306	6,02398	0,50327
ст. гр.	4,42773	4,06375	4,63689	0,33986
500	3,99761	4,99303	5,99855	0,50056
ст. гр.	1,59450	1,57352	1,78594	0,12960

Табела 2.6: YW оцене за случај 1.3.

$n_1$	$\hat{\mu}_1^{CML}$	$\hat{\mu}_2^{CML}$	$\hat{\mu}_3^{CML}$	$\hat{\alpha}^{CML}$
100	0,92033	1,99812	8,00413	0,11217
ст. гр.	4,42923	4,55641	4,39353	0,50847
200	0,95213	2,00017	7,99959	0,10684
ст. гр.	2,12835	2,14075	2,17734	0,22146
500	0,97631	2,00232	8,01027	0,10534
ст. гр.	0,81158	0,82962	0,85594	0,08487

са једнаким вероватноћама. матрица вероватноћа прелаза је  $\mathbf{P}_{mat} = \begin{bmatrix} 0,6 & 0,4 \\ 0,4 & 0,6 \end{bmatrix}$ , те са вероватноћом 0,6 процес остаје у истом стању, док са вероватноћом 0,4 мења стање.

2.2. Сада је  $\mu = (4, 5)'$ , што омогућује већу корелацију и у складу са тим бирамо  $\alpha = 0,6$ . Поново је  $\mathbf{p}_{vec} = (0,5, 0,5)'$ . У погледу матрице вероватноћа прелаза посматрамо супротну ситуацију у односу на претходну комбинацију параметара, у смислу да процес са већом вероватноћом мења стање него што остаје у истом стању. Дакле,  $\mathbf{P}_{mat} = \begin{bmatrix} 0,3 & 0,7 \\ 0,7 & 0,3 \end{bmatrix}$ .

2.3. Сада се посматра случај када се очекивања у различитим стањима разликују значајно. Вектор очекивања је  $\mu = (1, 4)'$ . За параметар тининг оператора бирамо  $\alpha = 0,2$ . Што се тиче иницијалног избора стања, прво стање се бира са значајно већом вероватноћом, тј.  $\mathbf{p}_{vec} = (0,8, 0,2)'$ . Матрица вероватноћа прелаза је  $\mathbf{P}_{mat} = \begin{bmatrix} 0,3 & 0,7 \\ 0,3 & 0,7 \end{bmatrix}$ , те се из било ког стања у друго стање прелази или остаје са вероватноћом 0,7, док се у прво стање прелази или остаје са вероватноћом 0,3. Дакле, процес има тенденцију да буде у другом стању.

Табела 2.7: YW и CML оцене за случај 2.1.

$n_1$	$\hat{\mu}_1^{YW}$	$\hat{\mu}_2^{YW}$	$\hat{\alpha}^{YW}$	$\hat{\mu}_1^{CML}$	$\hat{\mu}_2^{CML}$	$\hat{\alpha}^{CML}$
100	0,99831	1,99878	0,29073	0,96818	2,09529	0,32119
ст. гр.	2,45566	4,25409	1,24102	2,45727	5,07878	1,27816
200	1,00080	1,99801	0,29470	0,98583	2,03518	0,30994
ст. гр.	1,23721	2,11595	0,62401	1,18132	2,20676	0,54138
500	0,99871	2,00032	0,29692	0,99443	2,00853	0,30257
ст. гр.	0,49793	0,85208	0,25356	0,46512	0,82278	0,19269

Табела 2.8: YW и CML оцене за случај 2.2.

$n_1$	$\hat{\mu}_1^{YW}$	$\hat{\mu}_2^{YW}$	$\hat{\alpha}^{YW}$	$\hat{\mu}_1^{CML}$	$\hat{\mu}_2^{CML}$	$\hat{\alpha}^{CML}$
100	4,00418	5,01148	0,58778	3,87326	5,08567	0,61746
ст. гр.	9,48699	11,50429	0,97049	6,29138	7,59957	0,93556
200	3,99947	5,00049	0,59148	3,95615	5,00041	0,60260
ст. гр.	4,71575	5,77812	0,49295	3,04303	3,62314	0,39850
500	3,99595	4,99948	0,59580	3,98387	4,97811	0,59815
ст. гр.	1,89895	2,33979	0,20194	1,15382	1,35172	0,13665

Табела 2.9: YW и CML оцене за случај 2.3.

$n_1$	$\hat{\mu}_1^{YW}$	$\hat{\mu}_2^{YW}$	$\hat{\alpha}^{YW}$	$\hat{\mu}_1^{CML}$	$\hat{\mu}_2^{CML}$	$\hat{\alpha}^{CML}$
100	0,99960	4,00079	0,20024	0,85078	4,12198	0,24634
ст. гр.	2,77164	6,23327	2,81257	3,11731	5,61843	0,88345
200	1,00030	4,00321	0,19836	0,92049	4,07302	0,23072
ст. гр.	1,38526	3,10897	1,33556	1,45406	2,73887	0,41757
500	0,99929	4,00045	0,19816	0,96903	4,02906	0,21658
ст. гр.	0,55628	1,23384	0,52684	0,53306	0,99794	0,16318



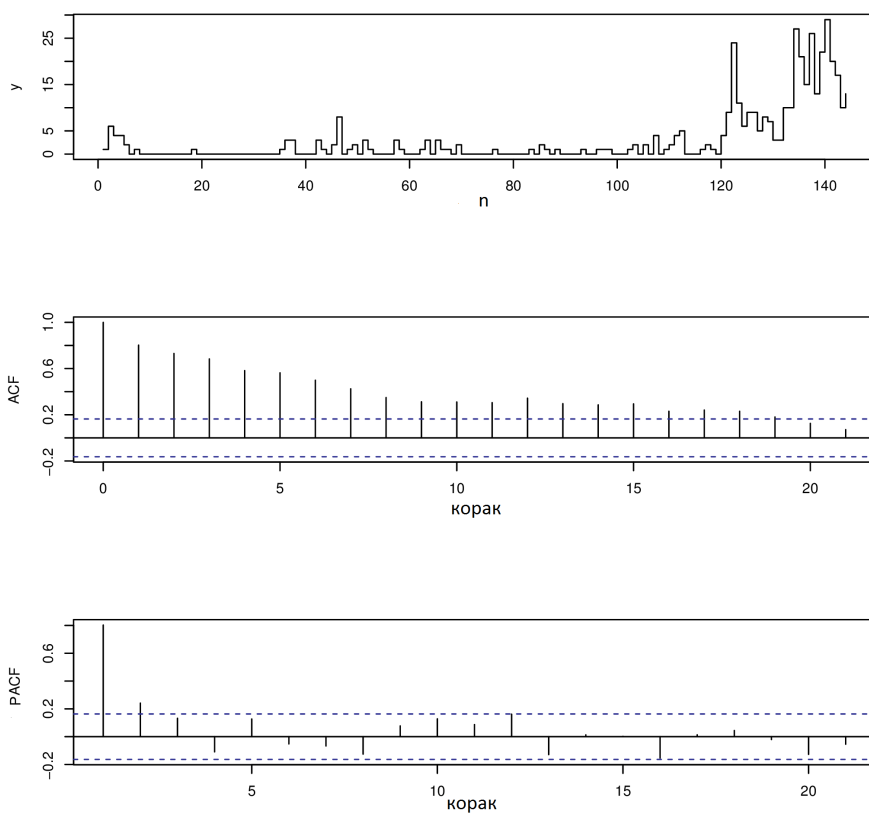
## Резултати оцењивања

У свим претходно размотреним случајевима оценили смо непознате параметре модела  $\alpha$  и  $\mu$  Јул-Вокеровим методом, као и методом условне максималне веродостојности. Како су подаци симулирани у 10000 понављања узорака обима 500, а циљ је испитати како се понашају оцене при повећању обима узорка, то ћемо реализације оцена израчунати најпре на основу подузорака од првих 100 елемената из сваког понављања, затим на основу подузорака које чине првих 200 елемената и на крају на основу целих узорака, обима 500. Резултати оцењивања, заједно са одговарајућим стандардним грешкама приказани су у табелама 2.3, 2.6 и 2.9, при чему прве две одговарају случају са 3, а трећа са 2 стања. Стварне вредности параметара дате су непосредно изнад резултата за сваки случај. Може се приметити да у свим случајевима реализације оцена конвергирају тачним вредностима параметара, док се стандардне грешке смањују. Што се тиче матрице вероватноћа прелаза која одговара процесу случајне средине, њу нисмо оцењивали, с обзиром на то да не представља параметар  $RrNGINAR(1)$  модела и да нам није била неопходна за остале израчунате вредности. Међутим, њена улога је јако важна у креирању самог процеса, пошто она заправо одређује утицај случајне средине на посматрани процес. У случају да желимо да предвидимо вредности  $RrNGINAR(1)$  процеса, тада би требало оценити ову матрицу.

## 2.4 Примена модела над реалним подацима

Већ је било речи о томе да су до сада  $INAR$  модели примењивани и код неких нестационарних серија. Притом овакве серије немају тренд или сезонски карактер, већ различити делови серије имају приметно различите средње вредности. Иако стационарни  $INAR$  модели не одговарају у потпуности оваквим подацима, до сада су коришћени, у одсуству бољих модела. Очекује се да у овом случају  $INAR$  процес у случајној средини покаже предност у односу на остале моделе, те је ова секција посвећена управо примени  $RrNGINAR(1)$  модела у случају конкретних реалних података.

У том циљу смо користили податке са интернет стране Forecasting Principles (<http://www.forecastingprinciples.com>), и то конкретно оне који представљају колико људи је ухапшено у периоду од месец дана због препродавања дроге. Ова реализација временског низа је преузета из 27. полицијске станице у Питсбургу, у периоду од јануара 1990. до децембра 2001. и садржи укупно 144 опсервације.

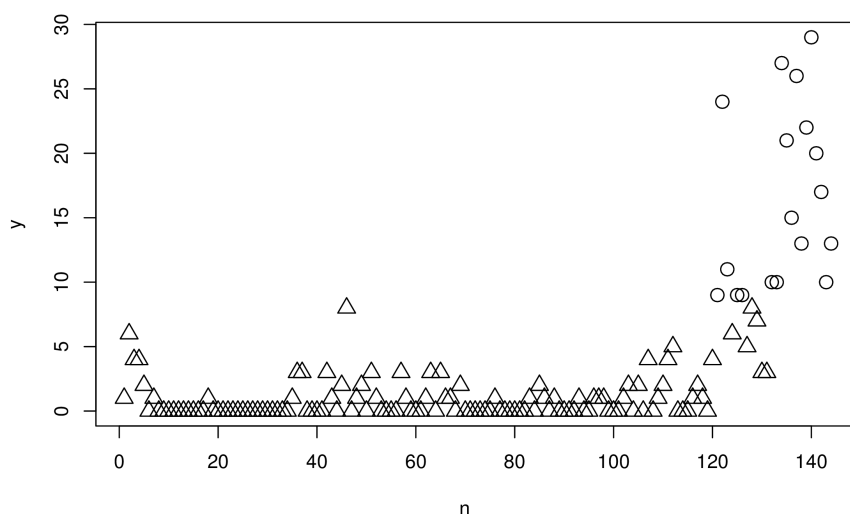


Слика 2.1: График коришћених података, као и графици одговарајућих аутокорељација и парцијалних аутокорељација.

Како би се закључило нешто о карактеру посматране серије, природно је најпре посматрати њен график, као и графике узорачке аутокорељационе (ACF) и парцијалне аутокорељационе функције.

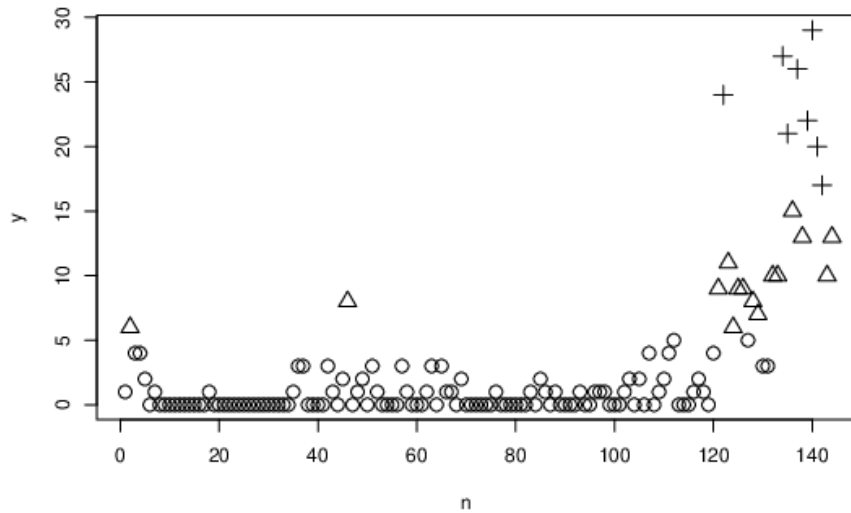
ције (PACF), што је приказано на слици 2.1. На основу тога, може се приметити да је адекватно применити INAR(1) моделе. Као могуће конкурентне моделе, ми смо посматрали следеће: целобројне ауторегресивне моделе са прагом SETINAR(2,1) и SETINAR(3,2) дате у Monteiro, Scotto и Pereira (2012), Möller и Weiß (2015), INAR(1) модел са Пуасоном маргиналном расподелом PoINAR(1), уведен у Al-Osh и Alzaid (1987), квази-биномни INAR(1) модел са генерализованом Пуасоном маргиналном расподелом GPoINAR(1), дат у Alzaid и Al-Osh (1993), геометријски INAR(1) модел дефинисан у Alzaid и Al-Osh (1988), нови геометријски INAR(1) у Ristić, Bakouch и Nastić (2009), негативни биномни INAR(1) из Zhu и Joe (2006, 2010), итеративни INAR(1) модел са негативном биномном маргиналном расподелом, дефинисан у Al-Osh и Aly (1992), INAR(1) модел са случајним коефицијентима и негативном биномном маргиналном расподелом, дат у Zheng, Basawa и Datta (2007). Посматрајући график података, могу се уочити скокови на почетку, а затим у току четврте године, док се у последње две године примећује значајно повећање посматраних криминалних активности. Може се претпоставити да је ово повећање изазвано променом неких услова у Питсбургу која се одиграла у последње две године. Одатле се природно намеће идеја да се покуша са применом RrNGINAR(1) модела. Први корак је кластеровање података. О овоме је било речи у првој глави. Наиме, сматрамо да је  $z_n = i$ , тј. да је реализована вредност процеса случајне средине једнака  $i$ , за  $i \in E_r$ , уколико вредност  $y_n$  припада  $i$ -том кластеру, за свако  $n \in \{1, 2, \dots, 144\}$ , где је  $E_r = \{1, 2, \dots, r\}$  одговарајући скуп могућих стања процеса случајне средине. С обзиром на то да сада имамо кластере који одговарају стањима, центри ових кластера се могу схватити као оцене средине  $\mu_1, \mu_2, \dots, \mu_r$ . Због тога се центри кластера користе као почетне вредности при нумеричкој процедури израчунавања оцена условне максималне веродостојности ових параметара. За оцењивање параметра  $\alpha$ , као почетна вредност се може узети средина интервала којем вредност овог параметра мора припадати. Како сада знамо низ реализација  $\{z_n\}$ , за свако  $n \in \{1, 2, \dots, 144\}$ , знамо да  $X_n(z_n)$  има геометријску расподелу са параметром  $\mu_{z_n}$ . Како су параметри расподела непознати, ми их оцењујемо помоћу оцена условне максималне веродостојности  $\hat{\mu}_1, \hat{\mu}_2, \dots, \hat{\mu}_r$ . Притом примењујемо два модела, у односу на број

стања, и то **R2NGINAR(1)** и **R3NGINAR(1)**. На сликама 2.2 и 2.3 је приказано на који начин су кластеровани подаци у случајевима  $r = 2$  и  $r = 3$ , респективно.



Слика 2.2: Кластеровање података у два стања.

Коначно, потребно је упоредити у којој мери модели успешно описују податке, па је у том циљу рачуната стандардна грешка предвиђања **RMS** (root mean squares). Она представља квадратни корен суме квадрата одступања предвиђењих од тачних вредности. У табели 2.10 су дате вредности за **RMS** заједно са оценама параметара добијених методом условне максималне веродостојности. Посматрајући вредности у тој табели закључујемо да је **R3NGINAR(1)** модел, заснован на процесу случајне средине са три стања показао најбоље резултате у смислу најмање вредности **RMS**. Примећујемо да је он бољи од модела **R2NGINAR(1)**. Ово се може описати тиме што подаци који су на слици 2.3. означени троуглом указују на скокове у првој и четвртој години мањег интензитета и велики број скокова у последњој години, од којих су они мањег интензитета означени троуглом, а већег крстом. У случају са два стања, сви



Слика 2.3: Кластеровање података у три стања.

ови скокови су посматрани као подаци из једне расподеле, док случај са три стања раздваја податке означене троуглом од оних означених крстом.

Табела 2.10: CML оцене параметара и RMS за примену INAR(1) модела.

Model	CML	RMS
SETINAR(2,1)	$\hat{\alpha} = 0,2990$ $\hat{\lambda}_1 = 0,8417$ $\hat{\lambda}_2 = 11,5794$	3,0372
SETINAR(3,1)	$\hat{\alpha} = 0,2845$ $\hat{\lambda}_1 = 0,8027$ $\hat{\lambda}_2 = 9,1617$ $\hat{\lambda}_3 = 10,4424$	3,3569
PoINAR(1)	$\hat{\lambda} = 1,237$ $\hat{\alpha} = 0,5948$	3,6613
GPQINAR(1)	$\hat{\lambda} = 0,5505$ $\hat{\theta} = 0,6108$ $\hat{\rho} = 0,392$	4,3398
GINAR(1)	$\hat{q} = 0,7596$ $\hat{\alpha} = 0,4809$	3,9456
NGINAR(1)	$\hat{\mu} = 3,3014$ $\hat{\alpha} = 0,7308$	3,4595
NBINAR(1)	$\hat{q} = 0,2173$ $\hat{\theta} = 0,834$ $\hat{\alpha} = 0,4563$	4,0185
NBIINAR(1)	$\hat{n} = 0,323$ $\hat{p} = 0,5335$ $\hat{\rho} = 0,8107$	3,4211
NBRCINAR(1)	$\hat{n} = 0,5435$ $\hat{p} = 0,1854$ $\hat{\rho} = 0,46$	4,0232
R2NGINAR(1)	$\hat{\mu}_1 = 1,1085$ $\hat{\mu}_2 = 12,9138$ $\hat{\alpha} = 0,052$	3,1090
R3NGINAR(1)	$\hat{\mu}_1 = 9,5906$ $\hat{\mu}_2 = 0,821$ $\hat{\mu}_3 = 23,249$ $\hat{\alpha} = 0,028$	1,6628

## Глава 3

# Уопштени INAR модели вишег реда у случајној средини

Модели вишег реда су бољи за описивање података са значајном парцијалном аутокорељацијом вишег реда, односно јаком зависношћу међу удаљеним елементима. Стога је природно увести моделе вишег реда у случајној средини зарад проширења поља примене. Први рад који се бави овим је Nastić, Laketa и Ristić (2019). У њему су разматрана два модела који се разликују у начину на који се достиже максимални (унапред задати) ред. Они су добијени као извесна комбинација  $RrNGINAR(1)$  модела и комбинованог геометријског  $INAR(p)$  модела базираног на негативном биномном тининг оператору ( $CGINAR(p)$ ) уведеног у раду Nastić, Ristić и Bakouch (2012). Након тога су уведени уопштени модели вишег реда у случајној средини, у раду Laketa, Nastić и Ristić (2018). Уопштење се састоји у томе да се допушта да стања средине, поред маргиналне расподеле, одређују и параметар тининг оператора, као и максимални ред процеса. Управо ти процеси су тема ове секције, чија је организација слична претходној. Најпре су поменути модели конструисани, а затим су дате њихове особине. Након тога су дефинисане оцене параметара, чији квалитет је испитан помоћу симулираних података. На крају су модели примењени над реалним подацима у циљу показивања применљивости модела.

### 3.1 Конструкција модела

У циљу увођења модела, дефинишимо следеће скупове:

- $\mathcal{M} = \{\mu_1, \mu_2, \dots, \mu_r\}$  - скуп који садржи могуће вредности параметра расподеле,
- $\mathcal{A} = \{\alpha_1, \alpha_2, \dots, \alpha_r\}$  - скуп који садржи могуће вредности параметра тининг оператора,
- $\mathcal{P} = \{p_1, p_2, \dots, p_r\}$  - скуп могућих редова процеса,

где  $r$  представља број различитих стања средине. Ови скупови садрже  $r$  вредности, при чему вредност са индексом  $i \in \{1, 2, \dots, r\}$  одговара  $i$ -том стању.

Уведимо следеће ознаке:

$$X_n(Z_n) = \sum_{z=1}^r X_n(z) I_{\{Z_n=z\}},$$

$$\varepsilon_n(Z_{n-1}, Z_n) = \sum_{z_1=1}^r \sum_{z_2=1}^r \varepsilon_n(z_1, z_2) I_{\{Z_{n-1}=z_1, Z_n=z_2\}},$$

$$\alpha_{Z_n} = \sum_{z=1}^r \alpha_z I_{\{Z_n=z\}}, \quad p_{Z_n} = \sum_{z=1}^r p_z I_{\{Z_n=z\}},$$

$$\phi_{k, P_n}^{(Z_n)} = \sum_{z=1}^r \sum_{p=1}^{p_z} \phi_{k,p}^{(z)} I_{\{Z_n=z, P_n=p\}}, \quad k \in \{1, 2, \dots, P_n\}.$$

Модел вишег реда можемо дефинисати на следећи начин

$$X_n(z_n) = \begin{cases} \alpha_{z_n} * X_{n-1}(z_{n-1}) + \varepsilon_n(z_{n-1}, z_n), & \text{с.в. } \phi_{1, P_n}^{(z_n)}, \\ \alpha_{z_n} * X_{n-2}(z_{n-2}) + \varepsilon_n(z_{n-2}, z_n), & \text{с.в. } \phi_{2, P_n}^{(z_n)}, \\ \vdots & \vdots \\ \alpha_{z_n} * X_{n-P_n}(z_{n-P_n}) + \varepsilon_n(z_{n-P_n}, z_n), & \text{с.в. } \phi_{P_n, P_n}^{(z_n)}. \end{cases}$$

Ако би важило  $P_n = p_{z_n}$ , било би јако компликовано одредити расподелу иновационе случајне променљиве. У том циљу се  $P_n$  дефинише на другачији начин. Ово је учињено на два различита



начина и оба ефикасно решавају поменути проблем. Примећено је да се у случају да свих  $P_n$  претходника одговарају истом стању довољно поједностављује израчунавање, па је ова идеја искоришћена. У том циљу уводимо број

$$p_n^* = \max \{i \in \{1, 2, \dots, n\} : z_{n-1} = z_{n-2} = \dots = z_{n-i}\},$$

који одређује колико претходника одговара истом стању. Први начин је да ред буде одређен минимумом  $P_n = \min\{p_{z_n}, p_n^*\}$ . Дакле, када дође до промене стања  $z_{n-1} \neq z_n$ , ред процеса постаје једнак јединици. Након тога, ред процеса расте све док процес остаје у истом стању до оног тренутка када достигне максималну вредност  $p_{z_n}$ , након чега је ред процеса константан и једнак овој максималној вредности све до наредне промене стање, где се ситуација понавља. Притом, може да се деси и да се стање промени пре него што процес достигне маскимални ред за то стање када се ред процеса поново враћа на јединичну вредност. Други начин је да дефинишемо

$$P_n = \begin{cases} p_{z_n}, & p_n^* \geq p_{z_n}, \\ 1, & p_n^* < p_{z_n}. \end{cases}$$

У овом случају, када дође до промене стања и ред процеса постане једнак јединици, он не расте до максималног реда, већ остаје једнак јединици све док се не испуне услови за максимални ред у том стању, када нагло скочи до те максималне вредности. Дакле, ред оваквог процеса у стању  $z_n$  је једнак 1 или  $p_{z_n}$ . Овај модел има мањи број параметара него претходни, па је једноставнији за рад. Како бисмо илустровали овако дефинисан низ редова, посматраћемо реализацију низа  $\{z_n\}$  дужине 20. У табели 3.1 су приказане одговарајуће вредности за  $P_n$  у случају оба модела са параметрима  $r = 2$  и  $\mathcal{P} = \{2, 3\}$ .

Дефинишимо сада прецизно ове моделе.

**ДЕФИНИЦИЈА 3.1.1** *Нека је  $\{Z_n\}$  процес случајне средине са  $r$  могућих стања из скупа  $E_r = \{1, 2, \dots, r\}$ ,  $r \in \mathbb{N}$ . Нека је  $\mathcal{M} = \{\mu_1, \mu_2, \dots, \mu_r\}$ ,  $\mathcal{A} = \{\alpha_1, \alpha_2, \dots, \alpha_r\}$  и  $\mathcal{P} = \{p_1, p_2, \dots, p_r\}$ , при чему  $\mu_i > 0$ ,  $\alpha_i \in (0, 1)$ ,  $p_i \in \mathbb{N}$ , за свако  $i \in E_r$ . Кажемо да је  $\{X_n\}$  INAR процес у случајној средини са  $r$  стања, скупом параметара расподеле  $\mathcal{M}$ , скупом параметара тининг*

Табела 3.1: Пример који илуструје дефинисање редова за оба модела за реализовани низ дужине 20 са параметрима  $r = 2$  и  $\mathcal{P} = \{2, 3\}$ .

$z_n$	1	1	1	1	1	2	2	2	1	1
$P_n$ за први модел	1	2	2	2	2	1	2	3	1	2
$P_n$ за други модел	1	2	2	2	2	1	1	3	1	2
$z_n$	2	1	2	2	2	2	2	2	1	1
$P_n$ за први модел	1	1	1	2	3	3	3	3	1	2
$P_n$ за други модел	1	1	1	1	3	3	3	3	1	2

оператора  $\mathcal{A}$  и скупом максималних редова  $\mathcal{P}$  ( $\text{RrINAR}_{\max}(\mathcal{M}, \mathcal{A}, \mathcal{P})$ ) ако је случајна променљива  $X_n(Z_n)$  за  $n \geq 1$  дефинисана са

$$X_n(Z_n) = \begin{cases} \alpha_{Z_n} * X_{n-1}(Z_{n-1}) + \varepsilon_n(Z_{n-1}, Z_n), & \text{с.в. } \phi_{1, P_n}^{(Z_n)}, \\ \alpha_{Z_n} * X_{n-2}(Z_{n-2}) + \varepsilon_n(Z_{n-2}, Z_n), & \text{с.в. } \phi_{2, P_n}^{(Z_n)}, \\ \vdots & \vdots \\ \alpha_{Z_n} * X_{n-P_n}(Z_{n-P_n}) + \varepsilon_n(Z_{n-P_n}, Z_n), & \text{с.в. } \phi_{P_n, P_n}^{(Z_n)} \end{cases} \quad (3.1.1)$$

где је  $P_n = \min\{p_{Z_n}, p_n^*\}$ , при чему важе следећи услови:

1. За свако  $i \in E_r$ ,  $p \in \{1, 2, \dots, p_i\}$  и  $k \in \{1, 2, \dots, p\}$  важи  $\phi_{k,p}^{(i)} \in [0, 1]$ . Такође,  $\sum_{k=1}^p \phi_{k,p}^{(i)} = 1$  за све  $i \in E_r$  и  $p \in \{1, 2, \dots, p_i\}$ .
2. За  $i \in E_r$  бројачки низ укључен у негативни биномни тининг оператор  $\alpha_i *$  је низ независних и једнако расподељених случајних променљивих са геометријском расподелом са очекивањем  $\alpha_i$ . Претпостављамо да је бројачки низ независан од свих осталих променљивих које се јављају у овој дефиницији.
3. За фиксиране вредности  $i, j \in E_r$ , низ  $\{\varepsilon_n(i, j)\}_{n \in \mathbb{N}}$  садржи независне и једнако расподељене случајне променљиве.
4.  $\{Z_n\}, \{\varepsilon_n(1, 1)\}, \{\varepsilon_n(1, 2)\}, \dots, \{\varepsilon_n(r, r)\}$  су низови међусобно независних случајних променљивих.

5. Случајна променљива  $X_n(l)$  не зависи од  $Z_m$  и  $\varepsilon_m(i, j)$  за све  $n < m$  и све  $i, j, l \in E_r$ .

У следећој дефиницији увешћемо други модел вишег реда у случајној средини.

**ДЕФИНИЦИЈА 3.1.2** Нека је  $\{Z_n\}$  процес случајне средине са  $r$  могућих стања из скупа  $E_r = \{1, 2, \dots, r\}$ ,  $r \in \mathbb{N}$ . Нека је  $\mathcal{M} = \{\mu_1, \mu_2, \dots, \mu_r\}$ ,  $\mathcal{A} = \{\alpha_1, \alpha_2, \dots, \alpha_r\}$  и  $\mathcal{P} = \{p_1, p_2, \dots, p_r\}$ , при чему  $\mu_i > 0$ ,  $\alpha_i \in (0, 1)$ ,  $p_i \in \mathbb{N}$  за све  $i \in E_r$ . Кажемо да је  $\{X_n\}$  INAR процес у случајној средини са  $r$  стања, скупом параметара расподеле  $\mathcal{M}$ , скупом параметара тининг оператора  $\mathcal{A}$  и скупом редова  $\mathcal{P}$  ( $\text{RrINAR}_1(\mathcal{M}, \mathcal{A}, \mathcal{P})$ ), ако за  $n \geq 1$  случајна променљива  $X_n(Z_n)$  задовољава једначину (3.1.1), где је

$$P_n = \begin{cases} p_{Z_n}, & p_n^* \geq p_{Z_n}, \\ 1, & p_n^* < p_{Z_n}, \end{cases}$$

и испуњени су услови 1-5 претходне дефиниције.

Како су дефинисани процеси јако комплексни, посматраћемо поједностављене моделе. Као у претходној глави, претпоставићемо да знамо реализацију  $\{z_n\}$  процеса случајне средине, као и маргиналну расподелу. У складу са тим, дефинисаћемо одговарајуће моделе на основу претходних дефиниција.

**ДЕФИНИЦИЈА 3.1.3** Нека је  $\{z_n\}$  реализација процеса случајне средине  $\{Z_n\}$  са  $r$  могућих стања из скупа  $E_r = \{1, 2, \dots, r\}$ ,  $r \in \mathbb{N}$  и нека је  $\mathcal{M} = \{\mu_1, \mu_2, \dots, \mu_r\}$ ,  $\mathcal{A} = \{\alpha_1, \alpha_2, \dots, \alpha_r\}$ ,  $\mathcal{P} = \{p_1, p_2, \dots, p_r\}$ , при чему  $\mu_i > 0$ ,  $\alpha_i \in (0, 1)$ ,  $p_i \in \mathbb{N}$ , за свако  $i \in E_r$ . Кажемо да је  $\{X_n\}$  геометријски INAR процес у случајној средини са  $r$  стања, скупом параметара расподеле  $\mathcal{M}$ , скупом параметара тининг оператора  $\mathcal{A}$  и скупом максималних редова  $\mathcal{P}$  ( $\text{RrNGINAR}_{\max}(\mathcal{M}, \mathcal{A}, \mathcal{P})$ ) ако за случајну променљива  $X_n(z_n)$  важи

$$X_n(z_n) = \begin{cases} \alpha_{z_n} * X_{n-1}(z_{n-1}) + \varepsilon_n(z_{n-1}, z_n), & \text{с.в. } \phi_{1, P_n}^{(z_n)}, \\ \alpha_{z_n} * X_{n-2}(z_{n-2}) + \varepsilon_n(z_{n-2}, z_n), & \text{с.в. } \phi_{2, P_n}^{(z_n)}, \\ \vdots & \vdots \\ \alpha_{z_n} * X_{n-P_n}(z_{n-P_n}) + \varepsilon_n(z_{n-P_n}, z_n), & \text{с.в. } \phi_{P_n, P_n}^{(z_n)}, \end{cases} \quad (3.1.2)$$

за  $n \geq 1$ , где  $P_n = \min\{p_{z_n}, p_n^*\}$ , задовољени су услови 1-5 из дефиниције 3.1.1 и случајна променљива  $X_n(z_n)$  има геометријску расподелу са очекивањем  $\mu_{z_n} \in \mathcal{M}$ , за свако  $n \geq 0$ .

**ДЕФИНИЦИЈА 3.1.4** Нека је  $\{z_n\}$  реализација процеса случајне средине  $\{Z_n\}$  са  $r$  могућих стања из скупа  $E_r = \{1, 2, \dots, r\}$ ,  $r \in \mathbb{N}$  и нека је  $\mathcal{M} = \{\mu_1, \mu_2, \dots, \mu_r\}$ ,  $\mathcal{A} = \{\alpha_1, \alpha_2, \dots, \alpha_r\}$ ,  $\mathcal{P} = \{p_1, p_2, \dots, p_r\}$ , при чему  $\mu_i > 0$ ,  $\alpha_i \in (0, 1)$ ,  $p_i \in \mathbb{N}$ , за свако  $i \in E_r$ . Кажемо да је  $\{X_n\}$  геометријски INAR процес у случајној средини са  $r$  стања, скупом параметара расподеле  $\mathcal{M}$ , скупом параметара тининг оператора  $\mathcal{A}$  и скупом редова  $\mathcal{P}$  ( $\text{RrNGINAR}_1(\mathcal{M}, \mathcal{A}, \mathcal{P})$ ), ако случајна променљива  $X_n(z_n)$  задовољава једначину (3.1.2) за  $n \geq 1$ , где је

$$P_n = \begin{cases} p_{z_n}, & p_n^* \geq p_{z_n} \\ 1, & p_n^* < p_{z_n}, \end{cases}$$

задовољени су услови 1–5 из дефиниције 3.1.1 и случајна променљива  $X_n(z_n)$  има геометријску расподелу са очекивањем  $\mu_{z_n} \in \mathcal{M}$ .

У случају да се нешто односи на оба модела користићемо општији акроним  $\text{RrNGINAR}(\mathcal{M}, \mathcal{A}, \mathcal{P})$ .

Као што је већ поменуто,  $\text{RrNGINAR}_{max}(p)$  и  $\text{RrNGINAR}_1(p)$  модели су само специјални случајеви модела датих дефиницијама 3.1.1 и 3.1.2. Могуће је посматрати још специфичније случајеве, тј. моделима  $\text{RrNGINAR}_{max}(p)$  и  $\text{RrNGINAR}_1(p)$  наметнути још неке услове. Примера ради, ако су сви елементи скупа  $\mathcal{M}$  једнаки, тј.  $\mu_1 = \mu_2 = \dots = \mu_r = \mu$ , тада свим стањима одговара иста расподела, па средина не утиче на маргиналну расподелу. Како би се истакла ова особина, овакав модел ћемо означити са  $\text{RrNGINAR}(\mu, \mathcal{A}, \mathcal{P})$ . Аналогно, у случају да је  $\alpha_1 = \alpha_2 = \dots = \alpha_r = \alpha$ , тј. да средина не утиче на тининг оператор, означићемо  $\text{RrNGINAR}(\mathcal{M}, \alpha, \mathcal{P})$ , а када средина не утиче на ред процеса,  $p_1 = p_2 = \dots = p_r = p$ , имамо  $\text{RrNGINAR}(\mathcal{M}, \mathcal{A}, p)$  модел. Специјално,  $\text{RrNGINAR}(\mathcal{M}, \alpha, p)$  се односи на моделе  $\text{RrNGINAR}_{max}(p)$  и  $\text{RrNGINAR}_1(p)$ , који су разматрани у раду Nastić, Laketa и Ristić (2019), а  $\text{RrNGINAR}(\mathcal{M}, \alpha, 1)$  представља  $\text{RrNGINAR}(1)$  модел. Такође је интересантно приметити да се модели  $\text{RrNGINAR}_{max}(\mathcal{M}, \mathcal{A}, \mathcal{P})$  и  $\text{RrNGINAR}_1(\mathcal{M}, \mathcal{A}, \mathcal{P})$  не разликују уколико се скуп  $\mathcal{P}$  састоји само из јединица и двојки. У пракси је некад погодно користити неки од ових специјалних

случајева, због смањења броја непознатих параметара модела, уколико се закључи да средина не утиче на неки од параметара  $\mu_i$ ,  $\alpha_i$ ,  $p_i$ .

## 3.2 Особине модела

Дефинисани процеси одређују у потпуности расподелу иновационе случајне променљиве  $\varepsilon_n(i, j)$  за све  $n \geq 1$  и све  $i, j \in E_r$ . У следећој теорему је дата ова расподела.

**Теорема 3.2.1** *Нека је  $\{X_n(z_n)\}$  RrNGINAR( $\mathcal{M}, \mathcal{A}, \mathcal{P}$ ) процес. Претпоставимо да  $z_n = j$  и  $z_{n-1} = i$  за неке  $i, j \in E_r$ . Ако  $0 \leq \alpha_j \leq \frac{\mu_j}{1 + \max_{k \in E_r} \mu_k}$ , тада расподела случајне променљиве  $\varepsilon_n(i, j)$  може да се запише као мешавина две геометријски расподељене случајне променљиве са очекивањима  $\mu_j$  и  $\alpha_j$  на следећи начин*

$$\varepsilon_n(i, j) \stackrel{d}{=} \begin{cases} \text{Geom} \left( \frac{\mu_j}{1 + \mu_j} \right), & \text{с.в. } 1 - \frac{\alpha_j \mu_i}{\mu_j - \alpha_j}, \\ \text{Geom} \left( \frac{\alpha_j}{1 + \alpha_j} \right), & \text{с.в. } \frac{\alpha_j \mu_i}{\mu_j - \alpha_j}. \end{cases} \quad (3.2.1)$$

*Доказ.* Нека је  $\{U_m^{(z_n)}\}$  бројачки низ укључен у негативни биномни тининг оператор  $\alpha_{z_n}^*$ . Са  $\phi_{U^{(z_n)}}(s)$  означимо функцију генератрице вероватноћа ове случајне променљиве. На основу дефиниције модела имамо

$$E(s^{X_n(z_n)}) = \sum_{l=1}^{P_n} \left[ \phi_{l, P_n}^{(z_n)} E(s^{\alpha_{z_n}^* X_{n-l}(z_{n-l}) + \varepsilon_n(z_{n-l}, z_n)}) \right].$$

На исти начин као у претходној глави добија се релација

$$E(s^{\alpha_{z_n}^* X_{n-l}(z_{n-l}) + \varepsilon_n(z_{n-l}, z_n)}) = \Phi_{X_{n-l}(z_{n-l})}(\Phi_{U^{(z_n)}}(s)) \Phi_{\varepsilon_n(z_{n-l}, z_n)}(s).$$

Зато је

$$\Phi_{X_n(z_n)}(s) = \sum_{l=1}^{P_n} \left[ \phi_{l, P_n}^{(z_n)} \Phi_{X_{n-l}(z_{n-l})}(\Phi_{U^{(z_n)}}(s)) \Phi_{\varepsilon_n(z_{n-l}, z_n)}(s) \right].$$

Претпоставимо да је  $z_n = j$  и  $z_{n-1} = i$  за неке  $i, j \in E_r$ . Како је  $z_{n-1} = z_{n-2} = \dots = z_{n-P_n} = i$ , важи

$$\begin{aligned}\Phi_{X_n(j)}(s) &= \sum_{l=1}^{P_n} \left[ \phi_{l, P_n}^{(j)} \Phi_{X_{n-1}(i)}(\Phi_{U^{(j)}}(s)) \Phi_{\varepsilon_n(i,j)}(s) \right] \\ &= \sum_{l=1}^{P_n} \phi_{l, P_n}^{(j)} \Phi_{X_{n-1}(i)}(\Phi_{U^{(j)}}(s)) \Phi_{\varepsilon_n(i,j)}(s) \\ &= \Phi_{X_{n-1}(i)}(\Phi_{U^{(j)}}(s)) \Phi_{\varepsilon_n(i,j)}(s).\end{aligned}$$

Одавде добијамо једначину

$$\Phi_{\varepsilon_n(i,j)}(s) = \frac{\Phi_{X_n(j)}(s)}{\Phi_{X_{n-1}(i)}(\Phi_{U^{(j)}}(s))}.$$

Случајне променљиве  $X_n(j)$  и  $X_{n-1}(i)$  имају геометријску расподелу са очекивањима  $\mu_j$  и  $\mu_i$ , респективно, док случајна променљива  $U^{(j)}$  има геометријску расподелу са очекивањем  $\alpha_j$ , па се на аналоган начин као у претходној глави добија

$$\Phi_{\varepsilon_n(i,j)}(s) = \frac{\alpha_j \mu_i}{\mu_j - \alpha_j} \cdot \frac{1}{1 + \alpha_j - \alpha_j s} + \left( 1 - \frac{\alpha_j \mu_i}{\mu_j - \alpha_j} \right) \cdot \frac{1}{1 + \mu_j - \mu_j s},$$

што доказује (3.2.1). Видимо да је последња формула готово иста као (2.2.7), једино уместо  $\alpha$  имамо  $\alpha_j$ .  $\square$

Сада ћемо анализирати коваријансну и корелациону структуру уведених модела. Почнимо најпре са  $\text{RrNGINAR}_{max}(\mathcal{M}, \mathcal{A}, \mathcal{P})$  моделом. На основу дефиниције модела и особина негативног биномног тининг оператора имамо да је коваријанса између случајних променљивих  $X_n(z_n)$  и  $X_{n-i}(z_{n-i})$ ,  $i \in \{1, 2, \dots, P_n\}$  дата изразом

$$\begin{aligned}\text{Cov}(X_n(z_n), X_{n-i}(z_{n-i})) &= \sum_{l=1}^{P_n} \phi_{l, P_n}^{(z_n)} \text{Cov}(\alpha_{z_n} * X_{n-l}(z_{n-l}), X_{n-i}(z_{n-i})) \\ &= \alpha_{z_n} \sum_{l=1}^{P_n} \phi_{l, P_n}^{(z_n)} \text{Cov}(X_{n-l}(z_{n-l}), X_{n-i}(z_{n-i})).\end{aligned}$$

Ако означимо  $\gamma_n^{(i)} \equiv \text{Cov}(X_n(z_n), X_{n-i}(z_{n-i}))$ , тада претходну једначину можемо записати на следећи начин

$$\gamma_n^{(i)} = \alpha_{z_n} \sum_{l=1}^{P_n} \phi_{l, P_n}^{(z_n)} \gamma_{\max(n-i, n-l)}^{(|i-l|)}. \quad (3.2.2)$$

Означимо са  $\rho_n^{(i)}$  корелацију између случајних променљивих  $X_n(z_n)$  и  $X_{n-i}(z_{n-i})$ . Дељењем обеју страна једначине (3.2.2) са  $\sqrt{\gamma_n^{(0)}\gamma_{n-i}^{(0)}}$ , добијамо

$$\begin{aligned}\rho_n^{(i)} &= \alpha_{z_n} \sum_{l=1}^{P_n} \phi_{l,P_n}^{(z_n)} \rho_{\max(n-i,n-l)}^{(|i-l|)} \sqrt{\frac{\gamma_{\max(n-i,n-l)}^{(0)} \gamma_{\max(n-i,n-l)-|i-l|}^{(0)}}{\gamma_n^{(0)} \gamma_{n-i}^{(0)}}} \\ &= \alpha_{z_n} \sum_{l=1}^{P_n} \phi_{l,P_n}^{(z_n)} \rho_{\max(n-i,n-l)}^{(|i-l|)} \sqrt{\frac{\gamma_{n-l}^{(0)}}{\gamma_n^{(0)}}}.\end{aligned}$$

Из конструкције модела следи да је  $\gamma_{n-1}^{(0)} = \gamma_{n-l}^{(0)}$  за  $l \in \{1, 2, \dots, P_n\}$ , одакле следи

$$\begin{aligned}\rho_n^{(i)} &= \alpha_{z_n} \sqrt{\frac{\gamma_{n-1}^{(0)}}{\gamma_n^{(0)}}} \sum_{l=1}^{P_n} \phi_{l,P_n}^{(z_n)} \rho_{\max(n-i,n-l)}^{(|i-l|)} \\ &= \alpha_{z_n} \sqrt{\frac{\mu_{z_{n-1}}(1 + \mu_{z_{n-1}})}{\mu_{z_n}(1 + \mu_{z_n})}} \sum_{l=1}^{P_n} \phi_{l,P_n}^{(z_n)} \rho_{\max(n-i,n-l)}^{(|i-l|)}.\end{aligned}$$

Овај резултат се може користити рекурзивно како би се израчунала корелација  $\rho_n^{(i)}$  произвољног реда на основу корелације  $\rho_n^{(0)} = 1$ , за свако  $n \geq 1$ . Интересантно је приметити да у случају да је  $z_n = z_{n-1}$  и  $P_n = 1$  за неко  $n \geq 1$ , корелација првог реда је  $\rho_n^{(1)} = \alpha_{z_n}$ . Стога,  $\alpha_{z_n}$  представља корелацију првог реда између случајних променљивих које одговарају истом стању  $z_n = z_{n-1}$  и реду  $P_n = 1$ .

У следећој теорему дати су условно очекивање и условна дисперзија дефинисаних процеса.

**Теорема 3.2.2** Нека  $\{X_n(z_n)\}$  представља  $\text{RrNGINAR}(\mathcal{M}, \mathcal{A}, \mathcal{P})$  процес дат дефиницијом 3.1.3, односно 3.1.4. Претпоставимо да је  $0 \leq \alpha_l \leq \frac{\mu_i}{1 + \max_{k \in E_r} \mu_k}$  за све  $l \in E_r$  и  $z_{n+1} = j$ ,  $z_n = i$  за  $i, j \in E_r$ . Тада су условно очекивање и условна дисперзија случајне променљиве  $X_{n+1}$  за дате  $X_n, X_{n-1}, \dots, X_{n+1-P_{n+1}}$  одређени формулама

$$E(X_{n+1}|H_n) = \mu_j - \alpha_j \mu_i + \alpha_j \sum_{l=1}^{P_{n+1}} \phi_{l,P_{n+1}}^{(z_{n+1})} X_{n+1-l},$$

$$\begin{aligned}
D(X_{n+1}|H_n) &= \mu_j(\mu_j + 1) - \alpha_j\mu_i(1 + 2\alpha_j + \alpha_j\mu_i) \\
&\quad + \alpha_j(1 + \alpha_j) \sum_{l=1}^{P_{n+1}} \phi_{l,P_{n+1}}^{(z_{n+1})} X_{n+1-l} \\
&\quad + \alpha_j^2 \sum_{l=1}^{P_{n+1}} \phi_{l,P_{n+1}}^{(z_{n+1})} X_{n+1-l}^2 - \alpha_j^2 \left( \sum_{l=1}^{P_{n+1}} \phi_{l,P_{n+1}}^{(z_{n+1})} X_{n+1-l} \right)^2,
\end{aligned}$$

где  $H_n = \sigma(X_n, X_{n-1}, \dots, X_{n-p_{n+1}})$  представља историју процеса.

*Доказ.* Ради поједностављења нотације, користићемо  $X_n$  уместо  $X_n(z_n)$ , за  $n \geq 0$  и  $\varepsilon_n$  уместо  $\varepsilon_n(z_{n-1}, z_n)$ , за  $n \geq 1$ . На основу дефиниције негативног биномног тининг оператора и особина условног очекивања, функција генератрисе условних вероватноћа је

$$\begin{aligned}
\Phi_{X_{n+1}|H_n}(s) &\equiv E(s^{X_{n+1}}|H_n) = \Phi_{\varepsilon_{n+1}}(s) \sum_{l=1}^{P_{n+1}} \phi_{l,P_{n+1}}^{(z_{n+1})} E(s^{\alpha_j * X_{n+1-l}}|H_n) \\
&= \Phi_{\varepsilon_{n+1}}(s) \sum_{l=1}^{P_{n+1}} \phi_{l,P_{n+1}}^{(z_{n+1})} \Phi_U^{X_{n+1-l}}(s).
\end{aligned}$$

Условно очекивање и условну дисперзију израчунаћемо помоћу ове функције користећи следеће једнакости

$$E(X_{n+1}|H_n) = \Phi'_{X_{n+1}|H_n}(1), \quad (3.2.3)$$

$$Var(X_{n+1}|H_n) = \Phi''_{X_{n+1}|H_n}(1) + \Phi'_{X_{n+1}|H_n}(1) - (\Phi'_{X_{n+1}|H_n}(1))^2. \quad (3.2.4)$$

Притом имамо да је

$$\Phi_U(s) = \frac{1}{1 + \alpha_j - \alpha_j s},$$

$$\Phi_{\varepsilon_{n+1}}(s) = \frac{\alpha_j\mu_i}{\mu_j - \alpha_j} \cdot \frac{1}{1 + \alpha_j - \alpha_j s} + \left(1 - \frac{\alpha_j\mu_i}{\mu_j - \alpha_j}\right) \cdot \frac{1}{1 + \mu_j - \mu_j s}.$$

Ако сада потражимо први и други извод функције  $\Phi_{X_{n+1}|H_n}(s)$  у односу на  $s$  и искористимо следеће резултате

$$\Phi_U(1) = 1, \quad \Phi'_U(1) = \alpha_j, \quad \Phi''_U(1) = 2\alpha_j^2,$$



$$\Phi_{\varepsilon_{n+1}}(1) = 1, \quad \Phi'_{\varepsilon_{n+1}}(1) = \mu_j - \alpha_j \mu_i, \quad \Phi''_{\varepsilon_{n+1}}(1) = 2\mu_j^2 - 2\alpha_j \mu_i (\mu_j + \alpha_j),$$

добићемо коначно

$$\Phi'_{X_{n+1}|H_n}(1) = \mu_j - \alpha_j \mu_i + \alpha_j \sum_{l=1}^{P_{n+1}} \phi_{l, P_{n+1}}^{(z_{n+1})} X_{n+1-l}$$

док је

$$\begin{aligned} \Phi''_{X_{n+1}|H_n}(1) &= 2\mu_j^2 - 2\alpha_j \mu_i (\mu_j + \alpha_j) + \alpha_j^2 \sum_{l=1}^{P_{n+1}} \phi_{l, P_{n+1}}^{(z_{n+1})} X_{n+1-l}^2 \\ &\quad + \alpha_j (2\mu_j - 2\alpha_j \mu_i + \alpha_j) \sum_{l=1}^{P_{n+1}} \phi_{l, P_{n+1}}^{(z_{n+1})} X_{n+1-l}. \end{aligned}$$

Последње две формуле уврстимо у (3.2.3) и (3.2.4), и добијамо тражени резултат.  $\square$

### 3.3 Оцењивање непознатих параметара

У овој секцији позабавићемо се оценама непознатих параметара добијених методом момената, које ћемо затим израчунати за серије симулираних података у циљу испитивања валидности оцена. Поново ћемо разматрати оцене добијене на основу метода Јул-Вокера и метода условне максималне веродостојности. Нека је  $X_1(z_1), X_2(z_2), \dots, X_N(z_N)$  случајни узорак из  $\text{RrNGINAR}_{\max}(\mathcal{M}, \mathcal{A}, \mathcal{P})$  процеса обима  $N$ .

#### 3.3.1 Јул-Вокер оцене

Приступ одређивања Јул-Вокер оцена је заснован на истој идеји као код  $\text{RrNGINAR}(1)$  модела. У жељи да добијемо строго постојане оцене, поделићемо узорак на делове који се могу посматрати као реализације  $\text{NGINAR}(1)$  процеса. У том циљу посматрајмо елемент узорка са редним бројем  $n$ , тако да важи  $P_n = P_{n-P_n} = p_{z_n}$ ,  $z_{n-P_n} = z_{n-P_n-1}$  и  $z_n = z_{n-1} = k \in E_r$  (ове претпоставке обезбеђују то да посматрани елемент, као и они елементи преко којих се он дефинише на основу модела одговарају истом стању  $k$  и истом реду

$p_k$ ). Како је  $z_n = k$ , први услов је еквивалентан  $P_n = P_{n-P_n} = p_k$ . Посматрајмо сада помоћне параметре  $\theta_{l,P_n}^{(z_n)}$ ,  $l \in \{1, 2, \dots, P_n\}$ , дефинисане са  $\theta_{l,P_n}^{(z_n)} = \alpha_{z_n} \phi_{l,P_n}^{(z_n)}$ . Коваријансна структура уведених модела може се представити у матричном облику на следећи начин

$$\begin{bmatrix} \gamma_n^{(1)} \\ \gamma_n^{(2)} \\ \vdots \\ \gamma_n^{(P_n)} \end{bmatrix} = \begin{bmatrix} \gamma_{n-1}^{(0)} & \gamma_{n-1}^{(1)} & \cdots & \gamma_{n-1}^{(P_n-1)} \\ \gamma_{n-1}^{(1)} & \gamma_{n-2}^{(0)} & \cdots & \gamma_{n-2}^{(P_n-2)} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n-1}^{(P_n-1)} & \gamma_{n-2}^{(P_n-2)} & \cdots & \gamma_{n-P_n}^{(0)} \end{bmatrix} \cdot \begin{bmatrix} \theta_{1,P_n}^{(z_n)} \\ \theta_{2,P_n}^{(z_n)} \\ \vdots \\ \theta_{P_n,P_n}^{(z_n)} \end{bmatrix}. \quad (3.3.1)$$

Такође, ово можемо записати на једноставнији начин  $\gamma_n = \Gamma_n \theta_n$ , где смо одговарајуће векторе означили са  $\gamma_n$  и  $\theta_n$ , док  $\Gamma_n$  представља коваријансну матрицу. Вектори  $\gamma_n$  и  $\theta_n$ , као и матрица  $\Gamma_n$  могу бити различитих димензија за различито  $n \in \mathbb{N}$ . Њихове димензије зависе од вредности реда  $P_n$ . Са друге стране, могуће је да су редови  $P_n$  и  $P_m$  међусобно једнаки, за  $n \neq m$ , али да су  $\theta_n$  и  $\theta_m$  различити, пошто могу одговарати различитим стањима  $z_n$  и  $z_m$ . Прецизно, имамо укупно  $p_1 + p_2 + \cdots + p_r$  различитих вектора  $\theta_n$ .

Случај  $\text{RrNGINAR}_1(\mathcal{M}, \mathcal{A}, \mathcal{P})$  модела је сличан. Тада важи  $\gamma_n^{(1)} = \alpha_{z_n} \gamma_{n-1}^{(0)}$ , за  $P_n = 1$ , док за  $P_n = p_{z_n}$  важи (3.3.1). У овом случају имамо  $r + \sum_{i=1}^r I_{\{p_i > 1\}}$  различитих вектора  $\theta_n$ .

Ако користимо (3.3.1) и узмемо у обзир све услове које испуњавају уведени процеси, видимо да се непознати параметри  $\theta_{l,k}^{(k)}$ ,  $l \in \{1, 2, \dots, P_n\}$ ,  $k \in E_r$ , могу оценити на основу следећег матричног система

$$\begin{bmatrix} \hat{\theta}_{1,p_k}^{(k)} \\ \hat{\theta}_{2,p_k}^{(k)} \\ \vdots \\ \hat{\theta}_{p_k,p_k}^{(k)} \end{bmatrix} = \begin{bmatrix} \hat{\gamma}_{0,p_k}^{(k)} & \hat{\gamma}_{1,p_k}^{(k)} & \cdots & \hat{\gamma}_{p_k-1,p_k}^{(k)} \\ \hat{\gamma}_{1,p_k}^{(k)} & \hat{\gamma}_{0,p_k}^{(k)} & \cdots & \hat{\gamma}_{p_k-2,p_k}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\gamma}_{p_k-1,p_k}^{(k)} & \hat{\gamma}_{p_k-2,p_k}^{(k)} & \cdots & \hat{\gamma}_{0,p_k}^{(k)} \end{bmatrix}^{-1} \cdot \begin{bmatrix} \hat{\gamma}_{1,p_k}^{(k)} \\ \hat{\gamma}_{2,p_k}^{(k)} \\ \vdots \\ \hat{\gamma}_{p_k,p_k}^{(k)} \end{bmatrix}, \quad (3.3.2)$$

где се за оцењивање  $\gamma_h^{(p_k)}$  користи  $\hat{\gamma}_{h,p_k}^{(k)}$  дефинисано помоћу

$$\hat{\gamma}_{h,p_k}^{(k)} = \frac{1}{n_{h,p_k}^{(k)}} \sum_{i \in S_{h,p_k}^{(k)}} (X_{i+h}(k) - \hat{\mu}_k)(X_i(k) - \hat{\mu}_k), \quad h \geq 0, \quad k \in E_r.$$

Оцене  $\widehat{\gamma}_{h,p_k}^{(k)}$  су израчунате на основу елемената узорка  $X_1, X_2, \dots, X_N$  који задовољавају услов поменут на почетку ове подсекције, тј. који су одређени скуповима  $S_{0,p_k}^{(k)} = \{i \in \{1, 2, \dots, N\} | z_i = k, P_i = p_k\}$  и  $S_{h,p_k}^{(k)} = \{i \in S_{0,p_k}^{(k)} | i + h \in S_{0,p_k}^{(k)}\}$ , где  $h \geq 1$ . Број елемената ових скупова је  $n_{h,p_k}^{(k)} = |S_{h,p_k}^{(k)}|$  за  $h \geq 0$ . С обзиром на тај услов, овако дефинисане оцене се односе само на оне елементе који одговарају максималним редовима  $p_1, p_2, \dots, p_r$ . Ако желимо да оценимо вероватноће које одговарају мањим редовима, користимо модификоване оцене, о којима ће бити речи на крају ове подсекције.

Како је  $\theta_{l,p_k}^{(k)} = \alpha_k \phi_{l,p_k}^{(k)}$  и  $\sum_{l=1}^{p_k} \phi_{l,p_k}^{(k)} = 1$ , параметри  $\alpha_k$  и  $\phi_{l,p_k}^{(k)}$ ,  $l \in \{1, 2, \dots, p_k\}$ , се могу оценити на следећи начин

$$\widehat{\alpha}_k^{YW} = \sum_{i=1}^{p_k} \widehat{\theta}_{i,p_k}^{(k)}, \quad \widehat{\phi}_{i,p_k}^{YW} = \frac{\widehat{\theta}_{i,p_k}^{(k)}}{\widehat{\alpha}_k^{YW}}, \quad i \in \{1, 2, \dots, p_k\}, \quad k \in E_r.$$

Такође, због  $E(X_i(k)) = \mu_k$  оцена параметра  $\mu_k$  је дата са

$$\widehat{\mu}_k^{YW} = \frac{1}{n_{0,p_k}^{(k)}} \sum_{i \in S_{0,p_k}^{(k)}} X_i(k), \quad k \in E_r.$$

Могућ је и случај да су вредности очекивања за различита стања међусобно једнаке  $\mu_1, \mu_2, \dots, \mu_r$  и једнаке некој константној вредности  $\mu$ . У том случају, ову вредност ћемо оценити помоћу

$$\widehat{\mu}^{YW} = \frac{\sum_{k=1}^r n_{0,p_k}^{(k)} \widehat{\mu}_k^{YW}}{\sum_{k=1}^r n_{0,p_k}^{(k)}}.$$

Ако је  $\alpha_1 = \alpha_2 = \dots = \alpha_r = \alpha$ , параметар  $\alpha$  се оцењује као

$$\widehat{\alpha}^{YW} = \frac{\sum_{k=1}^r n_{0,p_k}^{(k)} \widehat{\alpha}_k^{YW}}{\sum_{k=1}^r n_{0,p_k}^{(k)}},$$

док се за случај  $p_1 = p_2 = \dots = p_r = p$  користи

$$\widehat{\phi}_{i,p}^{YW} = \frac{\sum_{k=1}^r n_{0,p_k}^{(k)} \widehat{\phi}_{i,p_k}^{YW}}{\sum_{k=1}^r n_{0,p_k}^{(k)}}, \quad i \in \{1, 2, \dots, p\},$$

где су у сваком од претходна три случаја тежине одабране као тежине дефинисане заступљеношћу одговарајуће оцене у узорку.

Сличан је приступ и за  $\text{RrNGINAR}_1(\mathcal{M}, \mathcal{A}, \mathcal{P})$  модел, само што у овом случају све вероватноће могу бити оцењене. Оцене за параметар  $\alpha_k$  добијене на основу узорка који одговара реду 1, односно реду  $p_k$ , дате су формулама

$$\hat{\alpha}_{k,1} = \frac{\hat{\gamma}_{1,1}^{(k)}}{\hat{\gamma}_{0,1}^{(k)}}, \quad \hat{\alpha}_{k,p_k} = \sum_{i=1}^{p_k} \hat{\theta}_{i,p_k}^{(k)}, \quad k \in E_r,$$

где  $\theta_{i,p_k}^{(k)}$  задовољава (3.3.2) и оцене коваријанси су дате са

$$\hat{\gamma}_{h,j}^{(k)} = \frac{1}{n_{h,j}^{(k)}} \sum_{i \in S_{h,j}^{(k)}} (X_{i+h}(k) - \hat{\mu}_{k,j})(X_i(k) - \hat{\mu}_{k,j}), \quad h \geq 0, k \in E_r, j \in \{1, p_k\},$$

где су  $\hat{\mu}_{k,1}$  и  $\hat{\mu}_{k,p_k}$  оцене очекивања, тј.  $\hat{\mu}_{k,j} = \frac{1}{n_{0,j}^{(k)}} \sum_{i \in S_{0,j}^{(k)}} X_i(k)$ ,  $j \in \{1, p_k\}$ . Сва израчунавања се обављају над скуповима  $S_{0,1}^{(k)} = \{i \in \{1, 2, \dots, N\} \mid z_i = k, P_i = 1\}$  и  $S_{h,1}^{(k)} = \{i \in S_{0,1}^{(k)} \mid i + h \in S_{0,1}^{(k)}\}$ . Параметри вероватноћа  $\phi_{i,p_k}$  за  $i \in \{1, 2, \dots, p_k\}$  су оцењени помоћу  $\hat{\phi}_{i,p_k} = \frac{\hat{\theta}_{i,p_k}^{(k)}}{\alpha_{p_k}^{(k)}}$ , и  $k \in E_r$ . Коначно,

$$\hat{\alpha}_k^{YW} = \frac{n_{0,1}^{(k)} \hat{\alpha}_{k,1} + n_{0,p_k}^{(k)} \hat{\alpha}_{k,p_k}}{n_{0,1}^{(k)} + n_{0,p_k}^{(k)}}, \quad \hat{\mu}_k^{YW} = \frac{n_{0,1}^{(k)} \hat{\mu}_{k,1} + n_{0,p_k}^{(k)} \hat{\mu}_{k,p_k}}{n_{0,1}^{(k)} + n_{0,p_k}^{(k)}}.$$

Ако неки од скупова  $\mathcal{M}, \mathcal{A}, \mathcal{P}$  садржи само један елемент, тада се оцена за одговарајући параметар добија помоћу тежинске средине

$$\hat{\alpha}^{YW} = \frac{\sum_{k=1}^r (n_{0,1}^{(k)} + n_{0,p_k}^{(k)}) \hat{\alpha}_k^{YW}}{\sum_{k=1}^r (n_{0,1}^{(k)} + n_{0,p_k}^{(k)})}, \quad \hat{\mu}^{YW} = \frac{\sum_{k=1}^r (n_{0,1}^{(k)} + n_{0,p_k}^{(k)}) \hat{\mu}_k^{YW}}{\sum_{k=1}^r (n_{0,1}^{(k)} + n_{0,p_k}^{(k)})},$$

$$\hat{\phi}_{i,p_k}^{YW} = \frac{\sum_{k=1}^r n_{0,p_k}^{(k)} \hat{\phi}_{i,p_k}^{YW}}{\sum_{k=1}^r n_{0,p_k}^{(k)}}, \quad i \in \{1, 2, \dots, p_k\}.$$

Дефинисане оцене су строго постојане, што се може показати на идентичан начин као у претходној глави.

Такође је могуће дефинисати модификоване оцене, замењујући скупове  $S_{h,p}^{(k)}$ , за  $h \geq 1$ , скуповима  $S_{0,p}^{(k)}$ . Овако добијене оцене нису строго постојане, али је примерима разматраним у раду потврђено да дају добре резултате у пракси.

### 3.3.2 Оцене условне максималне веродостојности

Функција логаритамске условне веродостојности дата је следећим изразом

$$\text{Log}L = \sum_{i=2}^N \log P(X_i(z_i) = x_i | A_{i-1}).$$

где је са  $A_{i-1}$  означен догађај  $\{X_{i-1}(z_{i-1}) = x_{i-1} \cap \dots \cap X_{i-P_n}(z_{i-P_i}) = x_{i-P_i}\}$ . На основу теореме потпуне вероватноће је

$$\text{Log}L = \sum_{i=2}^N \sum_{k=1}^{P_{z_i}} \phi_{k,P_i}^{(z_i)} \log P(\alpha_{z_i} * X_{i-k}(z_{i-k}) + \varepsilon_i(z_{i-k}, z_i) = x_i | A_{i-1}).$$

Вероватноћа  $P(\alpha_{z_i} * X_{i-k}(z_{i-k}) + \varepsilon_i(z_{i-k}, z_i) = x_i | A_{i-1})$  се може представити као

$$\begin{aligned} & \sum_{l=0}^{x_{i-k}} P(\alpha_{z_i} * X_{i-k}(z_{i-k}) = l) P(\varepsilon_i(z_{i-k}, z_i) = x_i - l) \cdot I_{\{x_{i-k} \neq 0\}} \\ & + P(\varepsilon_i(z_{i-k}, z_i) = x_i) I_{\{x_{i-k} = 0\}}, \end{aligned}$$

док је, на основу теореме о расподели иновацине случајне променљиве,

$$P(\varepsilon_i(z_{i-k}, z_i) = x) = \left(1 - \frac{\alpha_{z_i} \mu_{z_{i-k}}}{\mu_{z_i} - \alpha_{z_i}}\right) \frac{\mu_{z_i}^x}{(1 + \mu_{z_i})^{x+1}} + \frac{\alpha_{z_i} \mu_{z_{i-k}}}{\mu_{z_i} - \alpha_{z_i}} \cdot \frac{\alpha_{z_i}^x}{(1 + \alpha_{z_i})^{x+1}}.$$

### 3.3.3 Анализа квалитета оцена над симулираним подацима

Како би се потврдило да су Јул-Вокер оцене добре, у смислу да реализације конвергирају тачним вредностима параметара при

повећању обима узорка, симулирани су подаци паралелно помоћу модела  $\text{RrNGINAR}_{max}(\mathcal{M}, \mathcal{A}, \mathcal{P})$  и  $\text{RrNGINAR}_1(\mathcal{M}, \mathcal{A}, \mathcal{P})$ . Симулирано је 100 реализованих серија укупно, свака обима 10000. Ово је такође учињено за различите комбинације параметара. Посматрани су модели са 2 и 3 различита стања, тј. за  $r = 2$  и  $r = 3$ . Најпре је посматран модел са два стања у којем је  $p_1 = 2$ ,  $p_2 = 3$ . Други симулирани модел има три стања и редове  $p_1 = p_2 = 2$ ,  $p_3 = 3$ . Разматрани су баш овакви случајеви, са што мањим бројем стања и мањим редовима, пошто би у супротном матрице параметара биле сувише велике, те табеле не би биле прегледне. Матрица вероватноћа прелаза процеса случајне средине означена је са  $\mathbf{p}_{mat}$ . Матрица  $\phi_k$  садржи вероватноће  $\phi_{i,j}^k$ , за  $k$ -то стање,  $j \in \{1, 2, \dots, p_k\}$ ,  $i \in \{1, 2, \dots, j\}$ , при чему се вероватноће за одговарајући  $\text{RrNGINAR}_1$  модел налазе у последњој колони. Дијагонални елементи матрице  $\mathbf{p}_{mat}$  су заправо вероватноће остајања у истом стању. Као што је то био случај са моделом из претходне главе, мале вредности ових вероватноћа условиле би честу промену стања, што би довело до лоших оцена. Посматраћемо два случаја различитих вредности параметара модела.

1. Први случај односи се на два стања средине, а вредности параметара су

$$\mathbf{p} = (2, 3)', \quad \boldsymbol{\mu} = (1, 2)', \quad \boldsymbol{\alpha} = (0, 3, 0, 6)',$$

$$\phi_1 = \begin{bmatrix} 1 & 0 \\ 0,4 & 0,6 \end{bmatrix}, \quad \phi_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0,5 & 0,5 & 0 \\ 0,2 & 0,3 & 0,5 \end{bmatrix}, \quad \mathbf{p}_{mat} = \begin{bmatrix} 0,7 & 0,3 \\ 0,2 & 0,8 \end{bmatrix}.$$

2. У другом случају се посматра средина са три могућа стања, док су вредности параметара

$$\mathbf{p} = (2, 2, 3)', \quad \boldsymbol{\mu} = (1, 2, 2)', \quad \boldsymbol{\alpha} = (0, 3, 0, 5, 0, 6)'$$

$$\phi_1 = \begin{bmatrix} 1 & 0 \\ 0,3 & 0,7 \end{bmatrix}, \quad \phi_2 = \begin{bmatrix} 1 & 0 \\ 0,4 & 0,6 \end{bmatrix}, \quad \phi_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0,6 & 0,4 & 0 \\ 0,4 & 0,1 & 0,5 \end{bmatrix},$$

$$\mathbf{p}_{mat} = \begin{bmatrix} 0,7 & 0,1 & 0,2 \\ 0,2 & 0,6 & 0,2 \\ 0,1 & 0,1 & 0,8 \end{bmatrix}.$$

Табела 3.2: YW оцене за случај 1. и RrNGINAR<sub>max</sub> модел.

$n$	$\hat{\mu}_1^{YW}$	$\hat{\alpha}_1^{YW}$	$\hat{\phi}_{2,1}^{1,YW}$	$\hat{\phi}_{2,2}^{1,YW}$
500	1,0065	0,3324	0,2991	0,7009
ст. гр.	0,1312	0,2140	0,8953	0,8953
1000	0,9926	0,3321	0,4375	0,5625
ст. гр.	0,0930	0,1463	0,2767	0,2767
5000	0,9988	0,2995	0,4264	0,5736
ст. гр.	0,0399	0,0671	0,0964	0,0964
10000	0,9999	0,2962	0,4194	0,5806
ст. гр.	0,0284	0,0479	0,0576	0,0576

$n$	$\hat{\mu}_2^{YW}$	$\hat{\alpha}_2^{YW}$	$\hat{\phi}_{2,1}^{2,YW}$	$\hat{\phi}_{2,2}^{2,YW}$	$\hat{\phi}_{3,1}^{2,YW}$	$\hat{\phi}_{3,2}^{2,YW}$	$\hat{\phi}_{3,3}^{2,YW}$
500	2,0204	0,6062	0,6067	0,3933	0,1690	0,3325	0,4985
ст. гр.	0,2438	0,1840	1,7684	1,7684	0,1690	0,1361	0,1614
1000	2,0010	0,6128	0,5869	0,4131	0,1830	0,3262	0,4908
ст. гр.	0,1774	0,1372	0,3955	0,3955	0,1062	0,0925	0,1089
5000	1,9935	0,5974	0,5187	0,4813	0,1884	0,3259	0,4857
ст. гр.	0,0864	0,0709	0,1302	0,1302	0,0517	0,0442	0,0524
10000	2,0005	0,5933	0,5106	0,4894	0,1917	0,3250	0,4833
ст. гр.	0,0620	0,0482	0,0819	0,0819	0,0360	0,0339	0,0394

Табела 3.3: YW оцне за случај 1. и RrNGINAR<sub>1</sub> модел.

$n$	$\hat{\mu}_1^{YW}$	$\hat{\alpha}_1^{YW}$	$\hat{\phi}_{2,1}^{1,YW}$	$\hat{\phi}_{2,2}^{1,YW}$
500	1,0065	0,3508	0,5235	0,4765
ст. гр.	0,1312	0,2448	1,2742	1,2742
1000	0,9926	0,3410	0,4455	0,5545
ст. гр.	0,0930	0,1586	0,2377	0,2377
5000	0,9988	0,3137	0,4059	0,5941
ст. гр.	0,0399	0,0780	0,0923	0,0923
10000	0,9999	0,3031	0,3990	0,6010
ст. гр.	0,0284	0,0532	0,0735	0,0735

$n$	$\hat{\mu}_2^{YW}$	$\hat{\alpha}_2^{YW}$	$\hat{\phi}_{3,1}^{2,YW}$	$\hat{\phi}_{3,2}^{2,YW}$	$\hat{\phi}_{3,3}^{2,YW}$
500	2,0204	0,5874	0,1946	0,3268	0,4786
ст. гр.	0,2438	0,1364	0,1957	0,1531	0,1553
1000	2,0010	0,6047	0,1873	0,3459	0,4668
ст. гр.	0,1774	0,0898	0,1471	0,1113	0,1287
5000	1,9935	0,6025	0,1836	0,3399	0,4765
ст. гр.	0,0864	0,0493	0,0632	0,0470	0,0577
10000	2,0005	0,6054	0,1813	0,3394	0,4794
ст. гр.	0,0620	0,0383	0,0410	0,0335	0,0359



Табела 3.4: YW оцене за случај 2. и RrNGINAR<sub>max</sub> модел.

$n$	$\hat{\mu}_1^{YW}$	$\hat{\alpha}_1^{YW}$	$\hat{\phi}_{2,1}^{1,YW}$	$\hat{\phi}_{2,2}^{1,YW}$			
500	0,9874	0,3311	0,3584	0,6416			
ст. гр.	0,1430	0,2395	0,5241	0,5241			
1000	0,9812	0,3142	0,4514	0,5486			
ст. гр.	0,0980	0,1554	0,9597	0,9597			
5000	0,9919	0,3129	0,3202	0,6798			
ст. гр.	0,0468	0,0737	0,1064	0,1064			
10000	0,9955	0,307	0,3213	0,6787			
ст. гр.	0,0328	0,0616	0,0732	0,0732			
$n$	$\hat{\mu}_2^{YW}$	$\hat{\alpha}_2^{YW}$	$\hat{\phi}_{2,1}^{2,YW}$	$\hat{\phi}_{2,2}^{2,YW}$			
500	1,9501	0,5107	0,3753	0,6247			
ст. гр.	0,3365	0,3133	0,6386	0,6386			
1000	1,9782	0,503	0,4194	0,5806			
ст. гр.	0,2310	0,2241	0,2404	0,2404			
5000	2,0042	0,4985	0,4156	0,5844			
ст. гр.	0,1100	0,0903	0,1061	0,1061			
10000	2,0007	0,501	0,4141	0,5859			
ст. гр.	0,0818	0,0693	0,0809	0,0809			
$n$	$\hat{\mu}_3^{YW}$	$\hat{\alpha}_3^{YW}$	$\hat{\phi}_{2,1}^{3,YW}$	$\hat{\phi}_{2,2}^{3,YW}$	$\hat{\phi}_{3,1}^{3,YW}$	$\hat{\phi}_{3,2}^{3,YW}$	$\hat{\phi}_{3,3}^{3,YW}$
500	1,9689	0,5861	0,8083	0,1917	0,3734	0,1340	0,4926
ст. гр.	0,2807	0,1982	1,1059	1,1059	0,3044	0,2200	0,1822
1000	1,9766	0,6023	0,8074	0,1926	0,3866	0,129	0,4843
ст. гр.	0,2147	0,1454	0,9091	0,9091	0,1529	0,1398	0,1095
5000	1,9846	0,5913	0,6285	0,3715	0,3785	0,1456	0,4759
ст. гр.	0,0885	0,0697	0,1884	0,1884	0,0649	0,0601	0,0551
10000	1,9937	0,5902	0,605	0,395	0,3823	0,1376	0,48
ст. гр.	0,0573	0,0489	0,0954	0,0954	0,0438	0,0408	0,0414

Табела 3.5: YW оцене за случај 2. и RrNGINAR<sub>1</sub> модел.

$n$	$\hat{\mu}_1^{YW}$	$\hat{\alpha}_1^{YW}$	$\hat{\phi}_{2,1}^{1,YW}$	$\hat{\phi}_{2,2}^{1,YW}$	
500	0,9874	0,3528	0,3226	0,6774	
ст. гр.	0,1430	0,2757	0,8362	0,8362	
1000	0,9812	0,3134	0,3899	0,6101	
ст. гр.	0,0980	0,198	0,3096	0,3096	
5000	0,9919	0,3092	0,3201	0,6799	
ст. гр.	0,0468	0,0826	0,119	0,119	
10000	0,9955	0,3109	0,327	0,673	
ст. гр.	0,0328	0,0596	0,0789	0,0789	
$n$	$\hat{\mu}_2^{YW}$	$\hat{\alpha}_2^{YW}$	$\hat{\phi}_{2,1}^{2,YW}$	$\hat{\phi}_{2,2}^{2,YW}$	
500	1,9501	0,4821	0,4211	0,5789	
ст. гр.	0,3365	0,2877	1,1938	1,1938	
1000	1,9782	0,4898	0,3847	0,6153	
ст. гр.	0,2310	0,2034	0,3551	0,3551	
5000	2,0042	0,5038	0,4145	0,5855	
ст. гр.	0,1100	0,0983	0,102	0,102	
10000	2,0007	0,5098	0,4124	0,5876	
ст. гр.	0,0818	0,0717	0,0758	0,0758	
$n$	$\hat{\mu}_3^{YW}$	$\hat{\alpha}_3^{YW}$	$\hat{\phi}_{3,1}^{3,YW}$	$\hat{\phi}_{3,2}^{3,YW}$	$\hat{\phi}_{3,3}^{3,YW}$
500	1,9689	0,6058	0,3867	0,0659	0,5475
ст. гр.	0,2807	0,1515	0,2707	0,2733	0,1727
1000	1,9766	0,5923	0,3801	0,1119	0,508
ст. гр.	0,2147	0,1079	0,1581	0,1349	0,1222
5000	1,9846	0,5959	0,3765	0,1509	0,4727
ст. гр.	0,0885	0,0577	0,0583	0,0582	0,0495
10000	1,9937	0,5944	0,3717	0,1539	0,4744
ст. гр.	0,0573	0,0335	0,0437	0,0391	0,0369

У поређењу са  $\text{RrNGINAR}_{max}(p)$  и  $\text{RrNGINAR}_1(p)$  моделима, уопштени модели вишег реда имају већи број параметара, што условљава већу флексибилност у реалним ситуацијама. Параметри који одговарају једном стању могу се оценити само на основу елемената који одговарају том стању. Како би се добили довољно добри резултати без повећања обима узорка, одабрали смо веће вредности параметра тининг оператора, пошто оне одговарају већој корелацији између елемената процеса. Може се приметити, дакле, да су веће вредности  $\alpha_k$  боље оцењене. Слична је ситуација и са параметрима  $\phi_{i,j}^k$ . Што је вредност  $\phi_{i,j}^k$  већа, то је већа корелација реда  $i$  помоћу које се оцењује параметар, што доприноси побољшању резултата. Такође се може приметити да су стандардне грешке за  $\mu_k$  мање за мање вредности ових параметара. Са  $\mu$ ,  $\alpha$  и  $p$  означени су вектори који одговарају скуповима  $M, A, P$ . Стандардне грешке параметара вероватноће су исте у случају два стања, пошто се  $\hat{\phi}_{2,2}^{YW}$  добијају као решења једначине  $\hat{\phi}_{2,2}^{YW} + \hat{\phi}_{2,1}^{YW} = 1$ . Резултати на левој страни табела 3.3 и 3.5 одговарају  $\text{RrNGINAR}_{max}$ , а они на десној страни  $\text{RrNGINAR}_1$  процесу. Приказани резултати показују да се са повећањем обима узорка реализације оцена параметара приближавају стварним вредностима параметара, што је и био циљ ове подсекције.

### 3.4 Примена модела над реалним подацима

Постоје неке ситуације у којима су значајне парцијалне аутокорелационе функције (расф) само одређених редова. Уобичајено је да се онда подаци моделирају неким процесом чији ред је одређен баш тако да расф тог реда буде значајна. Међутим, тиме се губе оне преостале значајне расф. Предност овде разматраних модела је што они могу у себе укључити све значајне расф, пошто је њихов ред променљив и зависи од стања. На пример, ако подаци показују значајне расф реда 2 и 4, онда се може користити модел са скупом редова  $P = \{2, 4\}$ . Дакле, очекивано је да у оваквом случају посматрани модели дају боље резултате. Како бисмо то доказали, упоредићемо их са неким другим моделима који се могу сматрати

адекватним за дате податке. Изабране су серије које представљају број људи ухапшених због препродавања дроге у 53. полицијској станици у Питсбургу, у периоду од јануара 1990. до децембра 2001. године. Ова серија је дужине 144 и преузета је са интернет странице *Forecasting Principles* (<http://www.forecastingprinciples.com>).

Графици посматраних серија, као и њихових узорачких аутокорељационих и парцијалних аутокорељационих функција дати су на сликама 3.1, 3.2 и 3.3. Реализације процеса случајне средине могу се одредити помоћу кластерована података, као што је то био случај код  $RrNGINAR(1)$  модела. Сваки кластер се придружује тачно једном стању и на тај начин се одређује низ реализација  $\{z_n\}$ . График кластерована дат је на слици 3.4. Познајући овај низ, сада се само на основу дефиниције може одредити низ  $\{P_n\}$ .

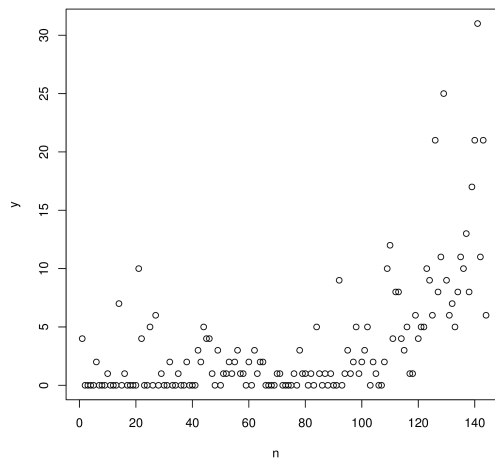
Ради поређења смо посматрали следеће моделе првог реда:

- $PoINAR(1)$  -  $INAR(1)$  модел са Пуасоном маргиналном расподелом дефинисан у Al-Osh и Alzaid (1987),
- $GPQINAR(1)$  - квази-биномни  $INAR(1)$  модел са уопштеном Пуасоном маргиналном расподелом, уведен у Alzaid и Al-Osh (1993),
- $GINAR(1)$  - геометријски  $INAR(1)$  модел дат у Alzaid и Al-Osh (1988).
- $NGINAR(1)$  - нови геометријски  $INAR(1)$  дефинисан у Ristić, Bakouch и Nastić (2009),
- $NBINAR(1)$  - негативни биномни  $INAR(1)$  дат у Zhu и Joe (2006, 2010).

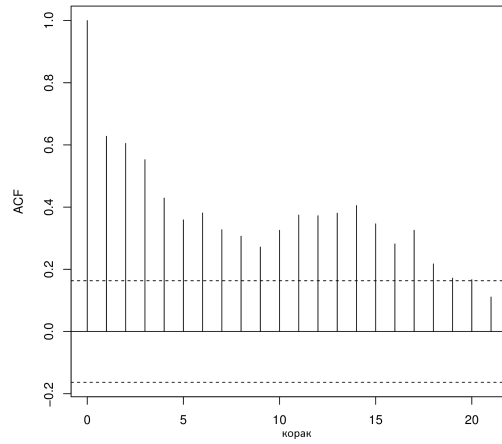
Модели вишег реда укључени у разматрање су  $RrNGINAR_{max}(p)$  и  $RrNGINAR_1(p)$  за  $p \in \{2, 3, 4\}$ .

У циљу одређивања мере у којој модели адекватно описују податке коришћена је стандардна грешка предвиђања (RMS), као и у претходној секцији. У табели 3.6 су приказани подаци за горепоменуте моделе, док табеле 3.6–3.9 садрже резултате за моделе  $RrNGINAR_{max}$  и  $RrNGINAR_1$ . Међу свим посматраним  $R2NGINAR_{max}$  моделима, као најбољи се показао  $R2NGINAR_{max}(2, 4)$ , с обзиром на најмању вредност RMS. Такође, међу  $R2NGINAR_1$  моделима,

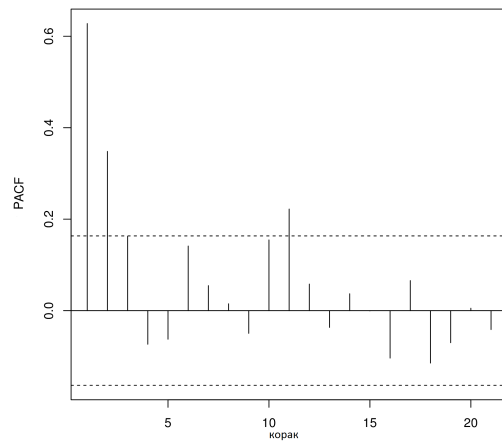
као најбољи се показао  $R2NGINAR_1(4,3)$ . Сви приказани резултати показују да су међу посматраним INAR моделима најбоље резултате дали управо модели који су тема ове секције.



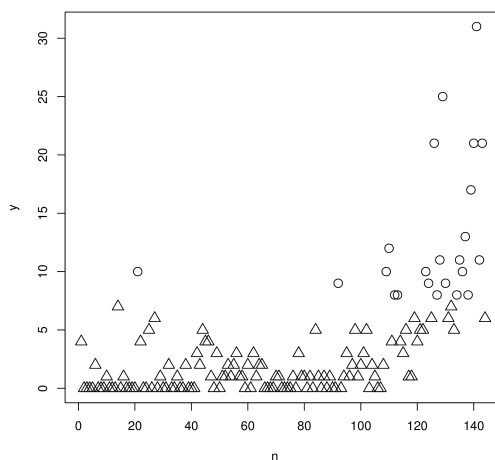
Слика 3.1: График коришћених података.



Слика 3.2: График аутокорељације посматраних података.



Слика 3.3: График парцијалне аутокорељације посматраних података.



Слика 3.4: Кластеровање посматраних података у случају два стања.

Табела 3.6: CML оцене параметара и RMS за различите моделе примењене над посматраним подацима.

Модел	CML	RMS
PoINAR(1)	$\hat{\lambda} = 2,0076$ $\hat{\alpha} = 0,4174$	4,2136
GPQINAR(1)	$\hat{\lambda} = 0,8369$ $\hat{\theta} = 0,5966$ $\hat{\rho} = 0,3470$	4,4148
GINAR(1)	$\hat{q} = 0,7803$ $\hat{\alpha} = 0,3935$	4,4303
NGINAR(1)	$\hat{\mu} = 3,6826$ $\hat{\alpha} = 0,4701$	4,2520
NBINAR(1)	$\hat{q} = 0,2014$ $\hat{\theta} = 0,8742$ $\hat{\alpha} = 0,2882$	4,4386

Табела 3.7: CML оцене и RMS за различите  $RrNGINAR_{max}$  моделе примењене над посматраним подацима.

Модел	CML	RMS
$R2NGINAR_{max}(2)$	$\hat{\boldsymbol{\mu}} = (4,9086, 1,2767)'$ $\hat{\alpha} = 0,2162$ $\hat{\boldsymbol{\phi}} = \begin{pmatrix} 1,0000 & 0,0000 \\ 0,5864 & 0,4136 \end{pmatrix}$	3,9829
$R2NGINAR_{max}(3)$	$\hat{\boldsymbol{\mu}} = (13,0280, 1,1564)'$ $\hat{\alpha} = 0,0425$ $\hat{\boldsymbol{\phi}} = \begin{pmatrix} 1,0000 & 0,0000 & 0,0000 \\ 0,0000 & 1,0000 & 0,0000 \\ 0,5265 & 0,3385 & 0,2280 \end{pmatrix}$	3,0996
$R2NGINAR_{max}(4)$	$\hat{\boldsymbol{\mu}} = (13,0856, 1,5712)'$ $\hat{\alpha} = 0,0597$ $\hat{\boldsymbol{\phi}} = \begin{pmatrix} 1,0000 & 0,0000 & 0,0000 & 0,0000 \\ 0,0000 & 1,0000 & 0,0000 & 0,0000 \\ 0,3314 & 0,3312 & 0,3373 & 0,0000 \\ 0,2620 & 0,1926 & 0,2449 & 0,3005 \end{pmatrix}$	3,0784
$R2NGINAR_{max}(2,2)$	$\hat{\boldsymbol{\mu}} = (13,0869, 1,5947)'$ $\hat{\boldsymbol{\alpha}} = (0,0000, 0,0568)'$ $\hat{\boldsymbol{\phi}}_1 = \begin{pmatrix} 1,0000 & 0,0000 \\ 0,0006 & 0,9994 \end{pmatrix}$ $\hat{\boldsymbol{\phi}}_2 = \begin{pmatrix} 1,0000 & 0,0000 \\ 0,0083 & 0,9917 \end{pmatrix}$	3,0642
$R2NGINAR_{max}(3,3)$	$\hat{\boldsymbol{\mu}} = (13,0869, 1,5941)'$ $\hat{\boldsymbol{\alpha}} = (0,0000, 0,0589)'$ $\hat{\boldsymbol{\phi}}_1 = \begin{pmatrix} 1,0000 & 0,0000 & 0,0000 \\ 0,0008 & 0,9992 & 0,0000 \\ 0,0000 & 0,3303 & 0,6697 \end{pmatrix}$ $\hat{\boldsymbol{\phi}}_2 = \begin{pmatrix} 1,0000 & 0,0000 & 0,0000 \\ 0,0093 & 0,9907 & 0,0000 \\ 0,3250 & 0,3307 & 0,3444 \end{pmatrix}$	3,0737



Табела 3.8: CML оцене и RMS за различите RrNGINAR<sub>max</sub> моделе примењене над посматраним подацима.

Модел	CML	RMS
R2NGINAR <sub>max</sub> (2,3)	$\hat{\boldsymbol{\mu}} = (13,0869, 1,5943)'$ $\hat{\boldsymbol{\alpha}} = (0,0095, 0,0586)'$ $\hat{\boldsymbol{\phi}}_1 = \begin{pmatrix} 1,0000 & 0,0000 \\ 0,0000 & 1,0000 \end{pmatrix}$ $\hat{\boldsymbol{\phi}}_2 = \begin{pmatrix} 1,0000 & 0,0000 & 0,0000 \\ 0,0095 & 0,9905 & 0,0000 \\ 0,3298 & 0,3301 & 0,3401 \end{pmatrix}$	3,0710
R2NGINAR <sub>max</sub> (3,4)	$\hat{\boldsymbol{\mu}} = (13,0869, 1,5932)'$ $\hat{\boldsymbol{\alpha}} = (0,3277, 0,0640)'$ $\hat{\boldsymbol{\phi}}_1 = \begin{pmatrix} 1,0000 & 0,0000 & 0,0000 \\ 0,0097 & 0,9902 & 0,0000 \\ 0,3277 & 0,3305 & 0,3417 \end{pmatrix}$ $\hat{\boldsymbol{\phi}}_2 = \begin{pmatrix} 1,0000 & 0,0000 & 0,0000 & 0,0000 \\ 0,0000 & 1,0000 & 0,0000 & 0,0000 \\ 0,3292 & 0,3300 & 0,3408 & 0,0000 \\ 0,2351 & 0,1994 & 0,2399 & 0,3257 \end{pmatrix}$	3,0748
R2NGINAR <sub>max</sub> (2,4)	$\hat{\boldsymbol{\mu}} = (13,0869, 1,5943)'$ $\hat{\boldsymbol{\alpha}} = (0,0005, 0,0584)'$ $\hat{\boldsymbol{\phi}}_1 = \begin{pmatrix} 1,0000 & 0,0000 \\ 0,0000 & 1,0000 \end{pmatrix}$ $\hat{\boldsymbol{\phi}}_2 = \begin{pmatrix} 1,0000 & 0,0000 & 0,0000 & 0,0000 \\ 0,0005 & 0,9995 & 0,0000 & 0,0000 \\ 0,3296 & 0,3300 & 0,3404 & 0,0000 \\ 0,2461 & 0,2004 & 0,2404 & 0,3131 \end{pmatrix}$	3,08504
R2NGINAR <sub>max</sub> (3,2)	$\hat{\boldsymbol{\mu}} = (13,0870, 1,5947)'$ $\hat{\boldsymbol{\alpha}} = (0,3296, 0,0569)'$ $\hat{\boldsymbol{\phi}}_1 = \begin{pmatrix} 1,0000 & 0,0000 & 0,0000 \\ 0,0099 & 0,9901 & 0,0000 \\ 0,3296 & 0,3302 & 0,3402 \end{pmatrix}$ $\hat{\boldsymbol{\phi}}_2 = \begin{pmatrix} 1,0000 & 0,0000 \\ 0,0000 & 1,0000 \end{pmatrix}$	3,06254

Табела 3.9: ML оцене и RMS за различите RrNGINAR<sub>max</sub> моделе примењене над посматраним подацима.

Модел	CML	RMS
R2NGINAR <sub>max</sub> (4,3)	$\hat{\mu} = (13,0869, 1,5891)'$ $\hat{\alpha} = (0,3278, 0,0748)'$ $\hat{\phi}_1 = \begin{pmatrix} 1,0000 & 0,0000 & 0,0000 & 0,0000 \\ 0,0000 & 1,0000 & 0,0000 & 0,0000 \\ 0,3278 & 0,3300 & 0,3422 & 0,0000 \\ 0,2436 & 0,2019 & 0,2400 & 0,3145 \end{pmatrix}$ $\hat{\phi}_2 = \begin{pmatrix} 1,0000 & 0,0000 & 0,0000 \\ 0,0060 & 0,9939 & 0,0000 \\ 0,2894 & 0,3323 & 0,3782 \end{pmatrix}$	3,0799
R2NGINAR <sub>max</sub> (4,4)	$\hat{\mu} = (13,0869, 1,5932)'$ $\hat{\alpha} = (0,3294, 0,0640)'$ $\hat{\phi}_1 = \begin{pmatrix} 1,0000 & 0,0000 & 0,0000 & 0,0000 \\ 0,0007 & 0,9992 & 0,0000 & 0,0000 \\ 0,3294 & 0,3300 & 0,3406 & 0,0000 \\ 0,2483 & 0,2005 & 0,2400 & 0,3112 \end{pmatrix}$ $\hat{\phi}_2 = \begin{pmatrix} 1,0000 & 0,0000 & 0,0000 & 0,0000 \\ 0,0000 & 1,0000 & 0,0000 & 0,0000 \\ 0,3292 & 0,3300 & 0,3408 & 0,0000 \\ 0,2350 & 0,1994 & 0,2399 & 0,3257 \end{pmatrix}$	3,0863
R2NGINAR <sub>max</sub> (4,2)	$\hat{\mu} = (13,0869, 1,5947)'$ $\hat{\alpha} = (0,3299, 0,0569)'$ $\hat{\phi}_1 = \begin{pmatrix} 1,0000 & 0,0000 & 0,0000 & 0,00000 \\ 0,0009 & 0,9991 & 0,0000 & 0,0000 \\ 0,3299 & 0,3300 & 0,3408 & 0,0000 \\ 0,3401 & 0,2497 & 0,2002 & 0,2400 \end{pmatrix}$ $\hat{\phi}_2 = \begin{pmatrix} 1,0000 & 0,0000 \\ 0,0000 & 1,0000 \end{pmatrix}$	3,08204

Табела 3.10: ML оцене и RMS за различите RrNGINAR<sub>1</sub> моделе примењене над посматраним подацима.

Модел	CML	RMS
R2NGINAR <sub>1</sub> (2)	$\hat{\mu} = (4,9086, 1,2767)'$ $\hat{\alpha} = 0,2161$ $\hat{\phi} = (0,5864, 0,4136)'$	3,9829
R2NGINAR <sub>1</sub> (3)	$\hat{\mu} = (9,3830, 1,66320)'$ $\hat{\alpha} = 0,1602$ $\hat{\phi} = (0,7267, 0,2070, 0,0663)'$	3,2626
R2NGINAR <sub>1</sub> (4)	$\hat{\mu} = (6,9013, 1,4031)'$ $\hat{\alpha} = 0,1776$ $\hat{\phi} = (0,7456, 0,2290, 0,0512, -0,0258)'$	3,6646
R2NGINAR <sub>1</sub> (2,2)	$\hat{\mu} = (13,0868, 1,5867)'$ $\hat{\alpha} = (0,0268, 0,0930)'$ $\hat{\phi}_1 = (0,0000, 1,0000)'$ $\hat{\phi}_2 = (0,0268, 0,9732)'$	3,0624
R2NGINAR <sub>1</sub> (3,3)	$\hat{\mu} = (13,0867, 1,5751)'$ $\hat{\alpha} = (0,4045, 0,1118)'$ $\hat{\phi}_1 = (0,2994, 0,4045, 0,2961)'$ $\hat{\phi}_2 = (0,3078, 0,3947, 0,2975)'$	3,0883
R2NGINAR <sub>1</sub> (4,4)	$\hat{\mu} = (13,0867, 1,5727)'$ $\hat{\alpha} = (0,2581, 0,1116)'$ $\hat{\phi}_1 = (0,2510, 0,2581, 0,2522, 0,2387)'$ $\hat{\phi}_2 = (0,2602, 0,2461, 0,2523, 0,2414)'$	3,0354
R2NGINAR <sub>1</sub> (2,3)	$\hat{\mu} = (13,0869, 1,5894)'$ $\hat{\alpha} = (0,3021, 0,0731)'$ $\hat{\phi}_1 = (0,0000, 1,0000)'$ $\hat{\phi}_2 = (0,3021, 0,3984, 0,2995)'$	3,10146

Табела 3.11: ML оцене и RMS за различите  $RrNGINAR_1$  моделе примењене над посматраним подацима.

Модел	CML	RMS
$R2NGINAR_1(4,2)$	$\hat{\mu} = (13,0867, 1,5811)'$ $\hat{\alpha} = (0,2043, 0,1122)'$ $\hat{\phi}_1 = (0,2509, 0,2043, 0,2413, 0,3035)'$ $\hat{\phi}_2 = (0,0353, 0,9647)'$	3,0302
$R2NGINAR_1(3,4)$	$\hat{\mu} = (13,0866, 1,5731)'$ $\hat{\alpha} = (0,4047, 0,1117)'$ $\hat{\phi}_1 = (0,2996, 0,4047, 0,2957)'$ $\hat{\phi}_2 = (0,2623, 0,1989, 0,2447, 0,2940)'$	3,1009
$R2NGINAR_1(2,4)$	$\hat{\mu} = (13,0869, 1,5890)$ $\hat{\alpha} = (0,2532, 0,0726)'$ $\hat{\phi}_1 = (0,0000, 1,0000)'$ $\hat{\phi}_2 = (0,2532, 0,1996, 0,2411, 0,3062)'$	3,0947
$R2NGINAR_1(3,2)$	$\hat{\mu} = (13,0867, 1,5812)$ $\hat{\alpha} = (0,4025, 0,1122)'$ $\hat{\phi}_1 = (0,3002, 0,4025, 0,2973)'$ $\hat{\phi}_2 = (0,0353, 0,9647)'$	3,1072
$R2NGINAR_1(4,3)$	$\hat{\mu} = (13,0867, 1,5746)'$ $\hat{\alpha} = (0,2080, 0,1118)'$ $\hat{\phi}_1 = (0,2510, 0,2080, 0,2425, 0,2984)'$ $\hat{\phi}_2 = (0,3080, 0,3946, 0,2974)'$	3,0084

## Глава 4

# Укрштени дводимензионални INAR модел првог реда у случајној средини

Ако желимо да посматрамо два процеса која су међусобно зависна, често је практично посматрати дводимензионални процес. Управо то је учињено у раду који су написали Popović, Ristić и Nastić (2016). С обзиром на тему дисертације, у овој глави ће бити конструисан дводимензионални INAR модел у случајној средини. Поред обичне корелације, компоненте тог модела су повезане и преко процеса случајне средине. Притом је ова повезаност “укрштена” у извесном смислу, што ће касније бити објашњено. Ова глава садржи самосталне резултате који су публиковани у раду Laketa (2020). Модел ће најпре бити дефинисан, након чега ће бити истражене његове особине. Дефинисаћемо оцене помоћу метода Јул-Вокера и метода условне максималне веродостојности, чију исправност ћемо тестирати над симулираним подацима. Биће уведен нови метод оцењивања процеса случајне средине, који доприноси побољшању квалитета оцена параметара модела. Коначно, конструисан модел ћемо упоредити са другим моделима у примени над реалним подацима.

## 4.1 Конструкција модела

Као што је то била пракса код претходних модела, најпре ћемо увести један уопштени модел, а затим његов специјални случај који је згоднији у пракси и који ће бити детаљно анализиран.

**ДЕФИНИЦИЈА 4.1.1** *Дводимензионални низ  $\{(X_n(W_n, Q_{n-1}), Y_n(W_{n-1}, Q_n))\}$ ,  $n \in \mathbb{N}$  ненегативних целобројних случајних променљивих је укрупњени дводимензионални целобројни ауторегресивни процес у случајној средини са  $(r_1, r_2)$  стања првог реда (CBRINAR(1,  $r_1, r_2$ )) уколико је*

$$\begin{aligned} X_n(W_n, Q_{n-1}) &= \sum_{i=1}^{X_{n-1}(W_{n-1}, Q_{n-2})} U_i^n + \varepsilon_n(W_{n-1}, W_n, Q_{n-1}), \quad n \in \mathbb{N}, \\ Y_n(W_{n-1}, Q_n) &= \sum_{i=1}^{Y_{n-1}(W_{n-2}, Q_{n-1})} V_i^n + \nu_n(W_{n-1}, Q_{n-1}, Q_n), \quad n \in \mathbb{N}, \end{aligned}$$

где је  $\{(W_n, Q_n)\}$  дводимензионални процес случајне средине са  $(r_1, r_2)$  стања дат дефиницијом 1.0.2,  $\{U_i^n\}$  и  $\{V_i^n\}$ ,  $i \in \mathbb{N}$  за фиксирано  $n \in \mathbb{N}$  су бројачки низови независних и једнако расподељених случајних променљивих које генеришу тининг оператор,  $W_n$  и  $Q_n$  одређују расподелу за  $U_i^n$  и  $V_i^n$  респективно и  $\{\varepsilon_n(i, j, p)\}$ ,  $\{\nu_n(i, p, q)\}$   $n \in \mathbb{N}_0$ ,  $i, j \in E_{r_1}$ ,  $p, q \in E_{r_2}$ , су низови независних и једнако расподељених случајних променљивих који задовољавају следеће услове:

- (A1) низови  $\{\varepsilon_n(i, j, p)\}$ ,  $\{\nu_n(i, p, q)\}$ ,  $n \in \mathbb{N}$   $\{W_n\}$  и  $\{Q_n\}$ ,  $n \in \mathbb{N}_0$  су међусобно независни, за све и све  $i, j \in E_{r_1}$ ,  $p, q \in E_{r_2}$ ,
- (A2)  $X_n(a, b)$  и  $Y_n(c, d)$  не зависе од  $W_m$ ,  $Q_m$ ,  $\varepsilon_m(i, j, p)$  и  $\nu_m(i, p, q)$  за  $n < m$  и све  $i, j, a, c \in E_{r_1}$  и  $p, q, b, d \in E_{r_2}$ .

Ознака  $X_n(W_n, Q_{n-1})$  упућује на то да  $W_n$  одређује расподелу за  $X_n$ , док  $Q_{n-1}$  одређује расподелу низа  $\{U_i^n\}$ . Аналогно,  $Y_n(W_{n-1}, Q_n)$  означава да  $Q_n$  и  $W_{n-1}$  одређују расподелу за  $Y_n$  и  $\{V_i^n\}$ , респективно. Како је расподела случајне променљиве  $X_{n-1}$  одређена са  $W_{n-1}$  и у првој релацији претходне дефиниције се не појављује  $\{U_i^{n-1}\}$ , расподела случајне променљиве  $\varepsilon_n$  зависи од  $W_{n-1}$  и не

зависи од  $Q_{n-2}$ , па можемо записати  $\varepsilon_n(W_{n-1}, W_n, Q_{n-1})$  и слично  $\nu_n(W_{n-1}, Q_{n-1}, Q_n)$ .

Овај модел, с обзиром на његову комплексност, неће бити детаљније анализиран. Међутим, он је у тесној вези са моделом који је дат следећом дефиницијом и који је тема ове главе.

**ДЕФИНИЦИЈА 4.1.2** Дводимензионални случајни процес  $\{(X_n(w_n, q_{n-1}), Y_n(w_{n-1}, q_n))\}$ ,  $n \in \mathbb{N}$ , где је  $(w_n, q_n)$  реализација дводимензионалног процеса случајне средине  $\{(W_n, Q_n)\}$  са  $(r_1, r_2)$  стања у тренутку  $n$ , је укритени INAR(1) модел у случајној средини са  $(r_1, r_2)$  стања, геометријском маргиналном расподелом, заснован на негативном биномном тининг оператору (CBRNGINAR(1,  $r_1, r_2$ )), ако је

$$X_n(w_n, q_{n-1}) = \alpha_{q_{n-1}} * X_{n-1}(w_{n-1}, q_{n-2}) + \varepsilon_n(w_{n-1}, w_n, q_{n-1}), \quad n \in \mathbb{N},$$

$$Y_n(w_{n-1}, q_n) = \beta_{w_{n-1}} * Y_{n-1}(w_{n-2}, q_{n-1}) + \nu_n(w_{n-1}, q_{n-1}, q_n), \quad n \in \mathbb{N},$$

где важе следећи услови

(B1) за  $i \in E_{r_1}$ ,  $\alpha_i \in (0, 1)$ , бројачки низ  $\{U_n^i\}$ ,  $n \in \mathbb{N}$ , укључен у оператор  $\alpha_i *$  је низ независних једнако расподељених случајних променљивих чија је расподела дефинисана са

$$P(U_n^i = u) = \frac{\alpha_i^u}{(1 + \alpha_i)^{u+1}}, \quad u \in \mathbb{N}_0,$$

(B2) за  $j \in E_{r_2}$ ,  $\beta_j \in (0, 1)$ , бројачки низ  $\{V_n^j\}$ ,  $n \in \mathbb{N}$ , укључен у оператор  $\beta_j *$  је низ независних једнако расподељених случајних променљивих чија је расподела дефинисана на следећи начин

$$P(V_n^j = v) = \frac{\beta_j^v}{(1 + \beta_j)^{v+1}}, \quad v \in \mathbb{N}_0,$$

(B3) маргинална расподела је геометријска,  $X_n$  има  $\text{Geom}\left(\frac{\mu_{w_n}}{1 + \mu_{w_n}}\right)$  а  $Y_n$   $\text{Geom}\left(\frac{\eta_{q_n}}{1 + \eta_{q_n}}\right)$  расподелу, тј.

$$P(X_n(w_n, q_{n-1}) = x) = \frac{\mu_{w_n}^x}{(1 + \mu_{w_n})^{x+1}}, \quad x \in \mathbb{N}_0,$$

$$P(Y_n(w_{n-1}, q_n) = y) = \frac{\eta_{q_n}^y}{(1 + \eta_{q_n})^{y+1}}, \quad y \in \mathbb{N}_0,$$

где  $\mu_{w_n} \in \{\mu_1, \mu_2, \dots, \mu_{r_1}\}$ ,  $\eta_{q_n} \in \{\eta_1, \eta_2, \dots, \eta_{r_2}\}$  и  $r_1, r_2 \in \mathbb{N}$ ,

(Б4)  $X_n(a, b)$  и  $Y_n(c, d)$  не зависе од  $\varepsilon_m(i, j, p)$  и  $\nu_m(i, p, q)$  за  $n < m$  произвољне  $i, j, a, c \in E_{r_1}$  и произвољне  $p, q, b, d \in E_{r_2}$ ,

(Б5) Коваријанса између  $X_n(w_n, q_{n-1})$  и  $Y_n(w_{n-1}, q_n)$  је иста као и коваријанса између  $X_m(w_m, q_{m-1})$  и  $Y_m(w_{m-1}, q_m)$  за  $w_n = w_m$ ,  $w_{n-1} = w_{m-1}$ ,  $q_n = q_m$  и  $q_{n-1} = q_{m-1}$ .

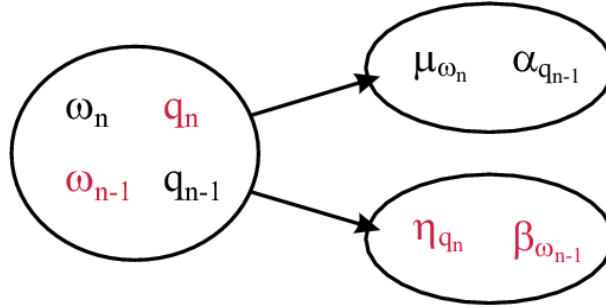
Можемо приметити да су  $\{X_n(W_n, Q_{n-1})\}$  и  $\{Y_n(W_{n-1}, Q_n)\}$  повезани помоћу процеса случајне средине  $\{W_n, Q_n\}$  на такав начин да зависе од обе компоненте (и  $W_n$  и  $Q_n$ ), па је немогуће посматрати их као два независна RrNGINAR(1) процеса. Поред тога, постоји и повезаност између  $X_n(w_n, q_{n-1})$  и  $Y_n(w_{n-1}, q_n)$ , која не потиче из процеса случајне средине, и која ће се користити у дефинисању Јул-Вокер оцена. Дакле, коваријанса између  $X_n$  и  $Y_n$  се може поделити на две различите “компоненте”

- 1) Коваријанса између  $\{X_n(W_n, Q_{n-1})\}$  и  $\{Y_n(W_{n-1}, Q_n)\}$  индукована дводимензионалним процесом случајне средине  $\{W_n, Q_n\}$ ,
- 2) Коваријанса између  $\{X_n(w_n, q_{n-1})\}$  и  $\{Y_n(w_{n-1}, q_n)\}$  за фиксиране вредности  $\{w_n\}$  и  $\{q_n\}$ .

Коваријанса првог типа се не може директно одредити, с обзиром на то да је њено порекло из модела датог дефиницијом 4.1.1, сувише компликованог за анализу. Уместо овог модела, користимо онај из дефиниције 4.1.2, базиран на реализацији процеса  $\{w_n, q_n\}$ . Међутим, то не значи да игноришемо корелацију првог типа. Ми је узимамо у обзир одабиром низа  $\{w_n, q_n\}$ . Овај низ се одређује кластеровањем, а он даље одређује низове параметара маргиналних расподела  $\{\mu_{w_n}\}$  и  $\{\eta_{q_n}\}$  као и низове параметара оператора истањивања  $\{\alpha_{q_{n-1}}\}$  и  $\{\beta_{w_{n-1}}\}$ . Дакле, одређени део корелације садржан је у реализованом процесу случајне средине који се одређује кластеровањем. Оваква особина је нова и даје велику флексибилност, али и једноставност моделу.

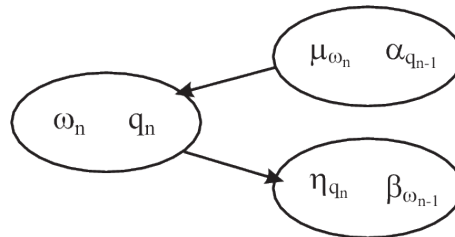
Објаснимо сада мало боље коваријансу првог типа. Уколико су  $\{w_n\}$  и  $\{q_n\}$  познати, онда су и низови параметара  $\{\mu_{w_n}\}$ ,  $\{\eta_{q_n}\}$ ,  $\{\alpha_{q_{n-1}}\}$  и  $\{\beta_{w_{n-1}}\}$  познати, што је илустровано на слици 4.1. Променимо перспективу. Претпоставимо да су нам познате вредности  $\{\mu_{w_n}\}$  и  $\{\alpha_{q_{n-1}}\}$  које одговарају случајној променљивој  $\{X_n(w_n, q_{n-1})\}$ .





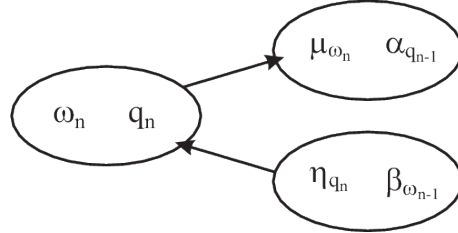
Слика 4.1: Графички приказ утицаја низа  $\{w_n, q_n\}$  на низове  $\{\mu_{w_n}\}$ ,  $\{\eta_{q_n}\}$ ,  $\{\alpha_{q_{n-1}}\}$  и  $\{\beta_{w_{n-1}}\}$ .

Тада можемо одредити  $\{w_n\}$  и  $\{q_n\}$ , а затим и  $\{\eta_{q_n}\}$  и  $\{\beta_{w_{n-1}}\}$  који одговарају  $\{Y_n(w_{n-1}, q_n)\}$ . Ово је илустровано на слици 4.2. Важи и обрат, што је приказано на слици 4.3.



Слика 4.2: Графички приказ утицаја низова  $\{\mu_{w_n}\}$  и  $\{\alpha_{q_{n-1}}\}$  који одговарају  $X_n$  на низ  $\{w_n, q_n\}$ , а затим и низове  $\{\eta_{q_n}\}$  и  $\{\beta_{w_{n-1}}\}$  који одговарају  $Y_n$ .

Истражимо још мало природу самог процеса. Примећујемо да  $W_n$  одређује расподелу за  $X_n$ , а  $Q_n$  расподелу за  $Y_n$ . На овај начин процес случајне средине одређује маргиналне расподеле посматраног процеса. Са друге стране, реч “укрштен” у називу модела има друго порекло. Наиме,  $W_{n-1}$  такође одређује параметар тининг оператора који фигурише у релацији којом се описује зависност између два суседна елемента  $Y_n$  и  $Y_{n-1}$ , док  $Q_{n-1}$  утиче на зависност између  $X_n$  и  $X_{n-1}$ , а важи аналогно и обратни случај.



Слика 4.3: Графички приказ утицаја низова  $\{\eta_{q_n}\}$  и  $\{\beta_{w_{n-1}}\}$  који одговарају  $Y_n$  на низ  $\{w_n, q_n\}$ , а затим и низове  $\{\mu_{w_n}\}$  и  $\{\alpha_{q_{n-1}}\}$  који одговарају  $X_n$ .

## 4.2 Особине модела

У овој секцији приказаћемо неке особине уведеног модела: расподелу иновационе случајне променљиве, матричну репрезентацију, коваријансну и корелациону структуру, као и условна очекивања и дисперзије. Неке од њих показују велику сличност са особинама модела из претходне главе.

### Расподела иновационе случајне променљиве

Иновационе случајне променљиве су расподељене слично као у претходној глави, посматрајући специјални случај када је  $\mathcal{P}_1 = \mathcal{P}_2 = \{1\}$ . Међутим, како су параметри тининг оператора одређени другачије, потребно је модификовати овај резултат на одговарајући начин.

**Теорема 4.2.1** *Нека је  $\{(X_n(w_n, q_{n-1}), Y_n(w_{n-1}, q_n))\}$ ,  $n \in \mathbb{N}$ , CBRNGI-NAR( $1, r_1, r_2$ ) модел дат дефиницијом 4.1.2 за  $r_1, r_2 \in \mathbb{N}$  и нека су  $\mu_1, \mu_2, \dots, \mu_{r_1}, \eta_1, \eta_2, \dots, \eta_{r_2}$  позитивни реални бројеви. Ако  $0 \leq \alpha_i \leq \min \left\{ \frac{\mu_l}{1+\mu_k}; k, l \in E_{r_1} \right\}$  и  $0 \leq \beta_j \leq \min \left\{ \frac{\eta_l}{1+\eta_k}; k, l \in E_{r_2} \right\}$ , за све  $i \in E_{r_1}$  и  $j \in E_{r_2}$  и ако  $w_{n-1} = i$ ,  $w_n = j$ ,  $q_{n-1} = p$ ,  $q_n = q$ ,  $i, j \in E_{r_1}$ ,  $p, q \in E_{r_2}$ , онда се расподела случајних променљивих  $\varepsilon_n(i, j, p)$  и  $\nu_n(i, p, q)$  може записати као мешавина геометријских расподела*

$$\varepsilon_n(i, j, p) \stackrel{d}{=} \begin{cases} \text{Geom} \left( \frac{\mu_j}{1+\mu_j} \right), & \text{с.в. } 1 - \frac{\alpha_p \mu_i}{\mu_j - \alpha_p}, \\ \text{Geom} \left( \frac{\alpha_p}{1+\alpha_p} \right), & \text{с.в. } \frac{\alpha_p \mu_i}{\mu_j - \alpha_p}, \end{cases} \quad (4.2.1)$$

$$\nu_n(i, p, q) \stackrel{d}{=} \begin{cases} \text{Geom} \left( \frac{\eta_q}{1+\eta_q} \right), & \text{с.в. } 1 - \frac{\beta_i \eta_p}{\eta_q - \beta_i}, \\ \text{Geom} \left( \frac{\beta_i}{1+\beta_i} \right), & \text{с.в. } \frac{\beta_i \eta_p}{\eta_q - \beta_i}. \end{cases} \quad (4.2.2)$$

*Доказ.* Први резултат се може добити применом теореме за расподелу иновационе случајне променљиве 3.2.1 из претходне главе, при чему је уместо  $\alpha_j$  потребно ставити  $\alpha_p$ . Други резултат се добија аналогно.  $\square$

Ако користимо исто означавање као у претходној теореме, онда су као њена директна последица очекивања случајних променљивих  $\varepsilon_n(i, j, p)$  и  $\nu_n(i, p, q)$  редом  $\mu_j - \alpha_p \mu_i$  и  $\eta_q - \beta_i \eta_p$ , а дисперзије  $\mu_j(1 + \mu_j) - \alpha_p \mu_i(1 + 2\alpha_p + \alpha_p \mu_i)$  и  $\eta_q(1 + \eta_q) - \beta_i \eta_p(1 + 2\beta_i + \beta_i \eta_p)$ .

### Матрична репрезентација

Сада ћемо модел приказати помоћу матрица, како бисмо радили са што једноставнијим формулама. Пре свега, уведемо ознаке као у претходним главама, означивши  $\mathbf{X}_n(\mathbf{w}_n, \mathbf{q}_{n-1})$  са  $\mathbf{X}_n$ ,  $\mathbf{Y}_n(\mathbf{w}_{n-1}, \mathbf{q}_n)$  са  $\mathbf{Y}_n$ ,  $\varepsilon_n(\mathbf{w}_{n-1}, \mathbf{w}_n, \mathbf{q}_{n-1})$  са  $\varepsilon_n$  и  $\nu_n(\mathbf{w}_{n-1}, \mathbf{q}_{n-1}, \mathbf{q}_n)$  са  $\nu_n$ .

Означимо даље,

$$\mathbf{X}_n = \begin{bmatrix} X_n \\ Y_n \end{bmatrix}, \quad \varepsilon_n = \begin{bmatrix} \varepsilon_n \\ \nu_n \end{bmatrix}, \quad \mathbf{A}_n = \begin{bmatrix} \alpha_{q_n} & 0 \\ 0 & \beta_{w_n} \end{bmatrix}$$

и дефинишимо оператор  $*$  између матрица на следећи начин

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} a_{11} * X + a_{12} * Y \\ a_{21} * X + a_{22} * Y \end{bmatrix}.$$

Сада модел можемо представити у матричном облику

$$\mathbf{X}_n = \mathbf{A}_{n-1} * \mathbf{X}_{n-1} + \varepsilon_n.$$

### Коваријансна и корелациона структура

Следећа теорема садржи изразе којима се одређују коваријансна и корелациона структура посматраног процеса.

**Теорема 4.2.2** *Нека је  $\{\mathbf{X}_n\} = \{(X_n, Y_n)\}$ ,  $n \in \mathbb{N}$ , CBRNGINAR(1,  $r_1, r_2$ ) модел дат дефиницијом 4.1.2 и нека су  $\mu_1, \mu_2, \dots, \mu_{r_1}, \eta_1, \eta_2, \dots, \eta_{r_2}$  позитивни реални бројеви и  $k$  цео број. Тада је*

(a) коваријансна матрица случајних променљивих  $\mathbf{X}_k$  и  $\mathbf{X}_0$ , дата помоћу

$$\text{Cov}(\mathbf{X}_k, \mathbf{X}_0) = \mathbf{A}_0 \mathbf{A}_1 \dots \mathbf{A}_{k-1} \text{Cov}(\mathbf{X}_0, \mathbf{X}_0), \quad (4.2.3)$$

(b) корелациона матрица случајних променљивих  $\mathbf{X}_k$  и  $\mathbf{X}_0$ , дата помоћу

$$\text{Corr}(\mathbf{X}_k, \mathbf{X}_0) = \begin{bmatrix} \sqrt{\frac{\text{Var}(X_0)}{\text{Var}(X_k)}} & 0 \\ 0 & \sqrt{\frac{\text{Var}(Y_0)}{\text{Var}(Y_k)}} \end{bmatrix} \mathbf{A}_0 \mathbf{A}_1 \dots \mathbf{A}_{k-1} \text{Corr}(\mathbf{X}_0, \mathbf{X}_0).$$

где је  $D(X_i) = \mu_{w_i}(1 + \mu_{w_i})$  и  $\text{Var}(Y_i) = \eta_{q_i}(1 + \eta_{q_i})$  за произвољан цео број  $i$ .

*Доказ.*

а) Како негативни биномни тининг оператор не зависи од  $\varepsilon_k$  и  $Y_0$ , имамо

$$\begin{aligned} \text{Cov}(X_k, Y_0) &= \text{Cov}(\alpha_{q_{k-1}} * X_{k-1} + \varepsilon_k, Y_0) = \alpha_{q_{k-1}} \text{Cov}(X_{k-1}, Y_0) \\ &= \alpha_{q_0} \alpha_{q_1} \dots \alpha_{q_{k-1}} \text{Cov}(X_0, Y_0), \end{aligned}$$

и аналогно

$$\begin{aligned} \text{Cov}(X_k, X_0) &= \alpha_{q_0} \alpha_{q_1} \dots \alpha_{q_{k-1}} \text{Cov}(X_0, X_0), \\ \text{Cov}(X_0, Y_k) &= \beta_{w_0} \beta_{w_1} \dots \beta_{w_{k-1}} \text{Cov}(X_0, Y_0), \\ \text{Cov}(Y_k, Y_0) &= \beta_{w_0} \beta_{w_1} \dots \beta_{w_{k-1}} \text{Cov}(Y_0, Y_0). \end{aligned}$$

Како је

$$\mathbf{A}_0 \mathbf{A}_1 \dots \mathbf{A}_{k-1} = \begin{bmatrix} \alpha_{q_0} \alpha_{q_1} \dots \alpha_{q_{k-1}} & 0 \\ 0 & \beta_{w_0} \beta_{w_1} \dots \beta_{w_{k-1}} \end{bmatrix},$$

то је очигледно

$$\begin{aligned} \text{Cov}(\mathbf{X}_k, \mathbf{X}_0) &= \begin{bmatrix} \text{Cov}(X_k, X_0) & \text{Cov}(X_k, Y_0) \\ \text{Cov}(Y_k, X_0) & \text{Cov}(Y_k, Y_0) \end{bmatrix} \\ &= \begin{bmatrix} \alpha_{q_0} \alpha_{q_1} \dots \alpha_{q_{k-1}} & 0 \\ 0 & \beta_{w_0} \beta_{w_1} \dots \beta_{w_{k-1}} \end{bmatrix} \text{Cov}(\mathbf{X}_0, \mathbf{X}_0) \\ &= \mathbf{A}_0 \mathbf{A}_1 \dots \mathbf{A}_{k-1} \text{Cov}(\mathbf{X}_0, \mathbf{X}_0). \end{aligned}$$

б) Што се корелације тиче, имамо

$$\begin{aligned}
\text{Corr}(\mathbf{X}_k, \mathbf{X}_0) &= \begin{bmatrix} \frac{\text{Cov}(X_k, X_0)}{\sqrt{\text{Var}(X_k)\text{Var}(X_0)}} & \frac{\text{Cov}(X_k, Y_0)}{\sqrt{\text{Var}(X_k)\text{Var}(Y_0)}} \\ \frac{\text{Cov}(Y_k, X_0)}{\sqrt{\text{Var}(Y_k)\text{Var}(X_0)}} & \frac{\text{Cov}(Y_k, Y_0)}{\sqrt{\text{Var}(Y_k)\text{Var}(Y_0)}} \end{bmatrix} \\
&= \mathbf{A}_0 \mathbf{A}_1 \dots \mathbf{A}_{k-1} \begin{bmatrix} \frac{\text{Cov}(X_0, X_0)}{\sqrt{\text{Var}(X_k)\text{Var}(X_0)}} & \frac{\text{Cov}(X_0, Y_0)}{\sqrt{\text{Var}(X_k)\text{Var}(Y_0)}} \\ \frac{\text{Cov}(Y_0, X_0)}{\sqrt{\text{Var}(Y_k)\text{Var}(X_0)}} & \frac{\text{Cov}(Y_0, Y_0)}{\sqrt{\text{Var}(Y_k)\text{Var}(Y_0)}} \end{bmatrix} \\
&= \mathbf{A}_0 \mathbf{A}_1 \dots \mathbf{A}_{k-1} \begin{bmatrix} \sqrt{\frac{\text{Var}(X_0)}{\text{Var}(X_k)}} & 0 \\ 0 & \sqrt{\frac{\text{Var}(Y_0)}{\text{Var}(Y_k)}} \end{bmatrix} \cdot \\
&\quad \text{Corr}(\mathbf{X}_0, \mathbf{X}_0) \\
&= \begin{bmatrix} \sqrt{\frac{\text{Var}(X_0)}{\text{Var}(X_k)}} & 0 \\ 0 & \sqrt{\frac{\text{Var}(Y_0)}{\text{Var}(Y_k)}} \end{bmatrix} \mathbf{A}_0 \mathbf{A}_1 \dots \mathbf{A}_{k-1} \cdot \\
&\quad \text{Corr}(\mathbf{X}_0, \mathbf{X}_0),
\end{aligned}$$

где последња једнакост следи из чињенице да дијагоналне матрице комутирају.  $\square$

### Условно очекивање и дисперзија

Следећа теорема даје формуле за условне величине.

**Теорема 4.2.3** Нека је  $\{\mathbf{X}_n\} = \{(X_n, Y_n)\}$ ,  $n \in \mathbb{N}$ , CBRNGINAR(1,  $r_1, r_2$ ) модел дат дефиницијом 4.1.2 и нека је  $k \in \mathbb{N}$ . Тада је

а) условно очекивање случајне променљиве  $\mathbf{X}_{n+k}$  за дато  $\mathbf{X}_n$  дато са

$$E(\mathbf{X}_{n+k} | \mathbf{X}_n) = \mathbf{A}_n \mathbf{A}_{n+1} \dots \mathbf{A}_{n+k-1} [\mathbf{X}_n - \boldsymbol{\mu}_n] + \boldsymbol{\mu}_{n+k}, \quad k \in \mathbb{N}_0, \quad (4.2.4)$$

$$\text{где } \boldsymbol{\mu}_n = \begin{bmatrix} \mu_{w_n} \\ \eta_{q_n} \end{bmatrix},$$

б) условне дисперзије случајних променљивих  $X_{n+k}$  и  $Y_{n+k}$  за дате

$X_n$  и  $Y_n$  су дате помоћу

$$\begin{aligned}
\text{Var}(X_{n+k}|X_n) &= \left\{ \alpha_{q_n}(1 + \alpha_{q_n}) \left( \prod_{s=1}^{k-1} \alpha_{q_{n+s}}^2 I\{k > 1\} + I\{k = 1\} \right) \right. \\
&\quad + \sum_{i=2}^{k-1} \left( \prod_{s=0}^{i-2} \alpha_{q_{n+s}} \right) \alpha_{q_{n+i-1}}(1 + \alpha_{q_{n+i-1}}) \left( \prod_{s=i}^{k-1} \alpha_{q_{n+s}}^2 \right) \\
&\quad \cdot I\{k > 2\} \\
&\quad + \left. \left( \prod_{s=0}^{k-1} \alpha_{q_{n+s}} \right) \alpha_{q_{n+k-1}}(1 + \alpha_{q_{n+k-1}}) I\{k > 1\} \right\} \\
&\quad \cdot (X_n - \mu_{w_n}) \\
&\quad \cdot \mu_{w_{n+k}}(1 + \mu_{w_{n+k}}) - \left( \prod_{s=0}^{k-1} \alpha_{q_{n+s}}^2 \right) \mu_{w_n}(1 + \mu_{w_n}), \\
\text{Var}(Y_{n+k}|Y_n) &= \left\{ \beta_{w_n}(1 + \beta_{w_n}) \left( \prod_{s=1}^{k-1} \beta_{q_{n+s}}^2 I\{k > 1\} + I\{k = 1\} \right) \right. \\
&\quad + \sum_{i=2}^{k-1} \left( \prod_{s=0}^{i-2} \beta_{w_{n+s}} \right) \beta_{w_{n+i-1}}(1 + \beta_{w_{n+i-1}}) \left( \prod_{s=i}^{k-1} \beta_{w_{n+s}}^2 \right) \\
&\quad \cdot I\{k > 2\} \\
&\quad + \left. \left( \prod_{s=0}^{k-1} \beta_{w_{n+s}} \right) \beta_{w_{n+k-1}}(1 + \beta_{w_{n+k-1}}) I\{k > 1\} \right\} \\
&\quad \cdot (X_n - \eta_{q_n}) \\
&\quad + \eta_{q_{n+k}}(1 + \eta_{q_{n+k}}) - \left( \prod_{s=0}^{k-1} \beta_{w_{n+s}}^2 \right) \eta_{q_n}(1 + \eta_{q_n}).
\end{aligned}$$

*Доказ.*

- а) На основу особина негативног биномног тининг оператора имамо

$$\begin{aligned}
E(X_{n+k}|X_n) &= \alpha_{q_{n+k-1}} E(X_{n+k-1}|X_n) + E(\varepsilon_{n+k}) \\
&= \alpha_{q_{n+k-1}} [\alpha_{q_{n+k-2}} E(X_{n+k-2}|X_n) + E(\varepsilon_{n+k-1})] \\
&\quad + E(\varepsilon_{n+k})
\end{aligned}$$

$$= \left( \prod_{s=0}^{k-1} \alpha_{q_{n+s}} \right) X_n + \sum_{l=1}^{k-1} \left( \prod_{s=l}^{k-1} \alpha_{q_{n+s}} \right) E(\varepsilon_{n+l}) \\ + E(\varepsilon_{n+k}).$$

Ако сада искористимо то што нам је расподела резидуала позната и формулу за очекивање дату након Теореме 4.2.1, имамо да је  $E(\varepsilon_{n+l}) = \mu_{w_{n+l}} - \alpha_{q_{n+l-1}} \mu_{w_{n+l-1}}$ , па је

$$E(X_{n+k}|X_n) = \left( \prod_{s=0}^{k-1} \alpha_{q_{n+s}} \right) X_n \\ + \sum_{l=1}^{k-1} \left( \prod_{s=l}^{k-1} \alpha_{q_{n+s}} \right) (\mu_{w_{n+l}} - \alpha_{q_{n+l-1}} \mu_{w_{n+l-1}}) \\ + \mu_{w_{n+l}} - \alpha_{q_{n+l-1}} \mu_{w_{n+k-1}} \\ = \left( \prod_{s=0}^{k-1} \alpha_{q_{n+s}} \right) X_n + \sum_{l=1}^{k-1} \left( \prod_{s=l}^{k-1} \alpha_{q_{n+s}} \right) \mu_{w_{n+l}} \\ + \sum_{l=1}^{k-1} \left( \prod_{s=l}^{k-1} \alpha_{q_{n+s}} \right) \alpha_{q_{n+l-1}} \mu_{w_{n+l-1}} \\ + \mu_{w_{n+k}} - \alpha_{q_{n+k-1}} \mu_{w_{n+k-1}}.$$

Сада можемо пренумерисати последњу суму, тако да коначно добијемо

$$E(X_{n+k}|X_n) = \left( \prod_{s=0}^{k-1} \alpha_{q_{n+s}} \right) X_n + \sum_{l=1}^{k-1} \left( \prod_{s=l}^{k-1} \alpha_{q_{n+s}} \right) \mu_{w_{n+l}} \\ + \sum_{l=0}^{k-2} \left( \prod_{s=l}^{k-1} \alpha_{q_{n+s}} \right) \mu_{w_{n+l}} \\ + \mu_{w_{n+k}} - \alpha_{q_{n+k-1}} \mu_{w_{n+k-1}} \\ = \left( \prod_{s=0}^{k-1} \alpha_{q_{n+s}} \right) X_n + \alpha_{q_{n+k-1}} \mu_{w_{n+k-1}} \\ + \left( \prod_{s=0}^{k-1} \alpha_{q_{n+s}} \right) \mu_{w_n} \\ + \mu_{w_{n+k}} - \alpha_{q_{n+k-1}} \mu_{w_{n+k-1}}$$

$$= \left( \prod_{s=0}^{k-1} \alpha_{q_{n+s}} \right) (X_n - \mu_{w_n}) + \mu_{w_{n+k}}.$$

Сличним закључивањем имамо

$$E(Y_{n+k}|Y_n) = \left( \prod_{s=0}^{k-1} \beta_{q_{n+s}} \right) (Y_n - \eta_{q_n}) + \eta_{q_{n+k}}.$$

Из претходних једначина следи тражени резултат.

- б) На основу особина негативног биномног тининг оператора добијамо следећу рекурентну релацију

$$D(X_{n+k}|X_n) = \alpha_{q_{n+k-1}}^2 D(X_{n+k-1}|X_n) + \alpha_{q_{n+k-1}}(1 + \alpha_{q_{n+k-1}})E(X_{n+k-1}|X_n) + D(\varepsilon_{n+k}).$$

Уведимо следеће ознаке

$$a_k = \alpha_{q_{n+k-1}}^2, \quad b_k = \alpha_{q_{n+k-1}}(1 + \alpha_{q_{n+k-1}})E(X_{n+k-1}|X_n) + D(\varepsilon_{n+k}).$$

Сада можемо записати

$$D(X_{n+k}|X_n) = a_k D(X_{n+k-1}|X_n) + b_k.$$

Како је  $D(X_n|X_n) = 0$ , то је општа формула

$$D(X_{n+k}|X_n) = \sum_{i=1}^{k-1} \left( \prod_{j=i+1}^k a_j \right) b_i I\{k \geq 2\} + b_k.$$

Замењујући  $E(X_{n+k-1}|X_n)$  и  $D(\varepsilon_{n+k})$ , који су раније израчунати, у  $b_i$  за  $i \in \{1, 2, \dots, k\}$ , добијамо

$$b_i = \alpha_{q_{n+i-1}}(1 + \alpha_{q_{n+i-1}}) \left( \prod_{s=0}^{i-2} \alpha_{q_{n+s}} I\{i \geq 2\} + I\{i = 1\} \right) (X_n - \mu_{w_n}) + \mu_{w_{n+i}}(1 + \mu_{w_{n+i}}) - \alpha_{q_{n+i-1}}^2 \mu_{w_{n+i-1}}(1 + \mu_{w_{n+i-1}}).$$

Затим, замењујући  $b_i$  за  $i \in \{1, 2, \dots, k\}$  у изразе за  $D(X_{n+k}|X_n)$ , добијамо

$$D(X_{n+k}|X_n) = I\{k \geq 2\} \sum_{i=1}^{k-1} \left[ \left( \prod_{j=i+1}^k a_j \right) \alpha_{q_{n+i-1}}(1 + \alpha_{q_{n+i-1}}) \right]$$



$$\begin{aligned}
& \cdot \left( \prod_{s=0}^{i-2} \alpha_{q_{n+s}} I\{i \geq 2\} + I\{i = 1\} \right) (X_n - \mu_{w_n}) \\
& + \mu_{w_{n+i}} (1 + \mu_{w_{n+i}}) - \alpha_{q_{n+i-1}}^2 \mu_{w_{n+i-1}} (1 + \mu_{w_{n+i-1}}) \Big] \\
& \alpha_{q_{n+k-1}} (1 + \alpha_{q_{n+k-1}}) (X_n - \mu_{w_n}) \\
& \cdot \left( \prod_{s=0}^{k-2} \alpha_{q_{n+s}} I\{k \geq 2\} + I\{i = 1\} \right) \\
& + \mu_{w_{n+k}} (1 + \mu_{w_{n+k}}) - \alpha_{q_{n+k-1}}^2 \mu_{w_{n+k-1}} (1 + \mu_{w_{n+k-1}}).
\end{aligned}$$

Сличан поступак даје одговарајући резултат за  $D(Y_{n+k}|X_n)$ .  
□

С обзиром на претходну теорему, при довољно великој вредности  $k$ , можемо сматрати да је  $E(\mathbf{X}_{n+k}|\mathbf{X}_n)$  приближно једнако  $\mu_{w_{n+k}}$ , тј.  $E(\mathbf{X}_{n+k}|\mathbf{X}_n) \approx \mu_{w_{n+k}} = E(\mathbf{X}_{n+k})$ . Такође,  $Var(X_{n+k}|X_n) \approx \mu_{w_{n+k}}(1 + \mu_{w_{n+k}}) = D(X_{n+k})$  и  $Var(Y_{n+k}|Y_n) \approx \eta_{q_{n+k}}(1 + \eta_{q_{n+k}}) = D(Y_{n+k})$ , за велико  $k$ . Овде не можемо говорити о конвергенцији, с обзиром на то да имамо различите вредности за  $w_n$ , међутим, то можемо тврдити за поднизовете који садрже елементе који одговарају истом стању (као што је то био случај у другој глави). Наиме, ако  $k \rightarrow \infty$ , низ  $\alpha_{w_0}, \alpha_{w_1}, \dots, \alpha_{w_k}$  садржи коначан број стања, па мора постојати стање које се појављује бесконачан број пута, тј. за бесконачно много  $i$  важи  $w_i = s$ , па имамо  $\alpha_s$  бесконачно много пута у низу, а како је  $\alpha_s \in (0, 1)$  производ конвергира ка нули.

### 4.3 Оцењивање параметара

Као и у претходним главама, даћемо оцене базиране на два метода: метод момената и метод условне максималне веродостојности. У том циљу посматрајмо узорак  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$  обима  $N$  из CBRNGINAR( $1, r_1, r_2$ ) модела.

#### 4.3.1 Јул-Вокер оцењивање

Као што је то било са до сада описаним моделима у овој дисертацији, основна идеја приликом дефинисања Јул-Вокер оцена

јесте да се узорак подели на делове који се могу посматрати као узорци неког стационарног процеса. У оценоу ће се укључити и корелација између компонената модела  $\{X_n(w_n, q_{n-1})\}$  и  $\{Y_n(w_{n-1}, q_n)\}$ , што наравно нисмо могли учинити у претходним једнодимензионалним моделима.

Како прва компонента има  $r_1$  стања околине, а друга  $r_2$ , поделићемо узорак на  $r_1 r_2$  делова, тако да сваком делу одговара тачно једно стање за прву и тачно једно стање за другу компоненту. Нека је  $I_{k,l} = \{i \in \{1, 2, \dots, N\} | w_{i+1} = w_i = w_{i-1} = k, q_{i+1} = q_i = q_{i-1} = l\}$  за  $k \in \{1, 2, \dots, r_1\}, l \in \{1, 2, \dots, r_2\}$  скуп индекса  $i$  за које  $X_i$  одговара параметрима  $\mu_k$  и  $\alpha_l$ , а  $Y_i$  параметрима  $\eta_l$  и  $\beta_k$ . Одговарајући подзорак означен је са  $S_{k,l} = \{\mathbf{X}_i | i \in I_{k,l}\}$  и његов обим је  $n_{k,l} = |S_{k,l}|$ . Испод су дате оценоу за очекивање, коваријансу и параметар тининг оператора дефинисане на подзорку  $S_{k,l}$ . Индекси 1, 0 у ознаци узорачке коваријансе  $\hat{\gamma}_{1,0}$  означавају да се ради о оценоу коваријансе  $Cov(X_{n+1}, Y_n)$ , 0, 1 у  $\hat{\gamma}_{0,1}$  указује на  $Cov(X_n, Y_{n+1})$  и 0, 0 се тиче  $Cov(X_n, Y_n)$ . Индекси у коваријансама, као и обично, означавају њихов ред.

$$\begin{aligned}\hat{\mu}_{k,l} &= \frac{1}{n_{k,l}} \sum_{i \in I_{k,l}} X_i, & \hat{\eta}_{k,l} &= \frac{1}{n_{k,l}} \sum_{i \in I_{k,l}} Y_i, \\ \hat{\gamma}_0^{X,k,l} &= \frac{1}{n_{k,l}} \sum_{i \in I_{k,l}} (X_i - \hat{\mu}_{k,l})^2, & \hat{\gamma}_0^{Y,k,l} &= \frac{1}{n_{k,l}} \sum_{i \in I_{k,l}} (Y_i - \hat{\eta}_{k,l})^2, \\ \hat{\gamma}_1^{X,k,l} &= \frac{1}{n_{k,l}} \sum_{\{i,i+1\} \subseteq I_{k,l}} (X_{i+1} - \hat{\mu}_{k,l})(X_i - \hat{\mu}_{k,l}), \\ \hat{\gamma}_1^{Y,k,l} &= \frac{1}{n_{k,l}} \sum_{\{i,i+1\} \subseteq I_{k,l}} (Y_{i+1} - \hat{\eta}_{k,l})(Y_i - \hat{\eta}_{k,l}), \\ \hat{\gamma}_{1,0}^{X,Y,k,l} &= \frac{1}{n_{k,l}} \sum_{\{i,i+1\} \subseteq I_{k,l}} (X_{i+1} - \hat{\mu}_{k,l})(Y_i - \hat{\eta}_{k,l}), \\ \hat{\gamma}_{0,1}^{X,Y,k,l} &= \frac{1}{n_{k,l}} \sum_{\{i,i+1\} \subseteq I_{k,l}} (X_i - \hat{\mu}_{k,l})(Y_{i+1} - \hat{\eta}_{k,l}), \\ \hat{\gamma}_{0,0}^{X,Y,k,l} &= \frac{1}{n_{k,l}} \sum_{i \in I_{k,l}} (X_i - \hat{\mu}_{k,l})(Y_i - \hat{\eta}_{k,l}).\end{aligned}$$

Ове оцене су строго постојане, пошто се дефинишу на делу узорка који одговара стационарном процесу.

Нека је  $n \in \{1, 2, \dots, N\}$  такав број за који важи  $w_{n-1} = w_n = w_{n+1} = k$  и  $q_{n-1} = q_n = q_{n+1} = l$ . На основу особина коваријансе разматраних у претходној секцији, имамо

$$\text{Cov}(X_{n+1}, Y_n) = \alpha_l \text{Cov}(X_n, Y_n), \quad \text{Cov}(X_{n+1}, X_n) = \alpha_l D(X_n).$$

Стога можемо записати

$$\begin{aligned} \alpha_l &= \frac{\text{Cov}(X_{n+1}, Y_n)}{\text{Cov}(X_n, Y_n)} = \frac{\text{Cov}(X_{n+1}, X_n)}{D(X_n)} \\ &= \frac{1}{2} \left[ \frac{\text{Cov}(X_{n+1}, Y_n)}{\text{Cov}(X_n, Y_n)} + \frac{\text{Cov}(X_{n+1}, X_n)}{D(X_n)} \right], \end{aligned}$$

и аналогно, из

$$\text{Cov}(X_n, Y_{n+1}) = \beta_k \text{Cov}(X_n, Y_n), \quad \text{Cov}(Y_n, Y_{n+1}) = \beta_k D(Y_n),$$

имамо

$$\begin{aligned} \beta_k &= \frac{\text{Cov}(X_n, Y_{n+1})}{\text{Cov}(X_n, Y_n)} = \frac{\text{Cov}(Y_{n+1}, Y_n)}{D(Y_n)} \\ &= \frac{1}{2} \left[ \frac{\text{Cov}(X_n, Y_{n+1})}{\text{Cov}(X_n, Y_n)} + \frac{\text{Cov}(Y_{n+1}, Y_n)}{D(Y_n)} \right]. \end{aligned}$$

У складу са претходним разматрањем, Јул-Вокер оцене су дате са

$$\begin{aligned} \hat{\mu}_k &= \frac{\sum_{l=1}^{r_2} n_{k,l} \hat{\mu}_{k,l}}{\sum_{l=1}^{r_2} n_{k,l}}, & \hat{\eta}_l &= \frac{\sum_{k=1}^{r_1} n_{k,l} \hat{\eta}_{k,l}}{\sum_{k=1}^{r_1} n_{k,l}}, \\ \hat{\alpha}_k &= \frac{\sum_{l=1}^{r_2} n_{k,l} \hat{\alpha}_{k,l}}{\sum_{l=1}^{r_2} n_{k,l}}, & \hat{\beta}_l &= \frac{\sum_{k=1}^{r_1} n_{k,l} \hat{\beta}_{k,l}}{\sum_{k=1}^{r_1} n_{k,l}}, \end{aligned}$$

за  $k \in E_{r_1}$ ,  $l \in E_{r_2}$ , где

$$\hat{\alpha}_{k,l} = \frac{1}{2} \left( \frac{\hat{\gamma}_{1,0}^{X,Y,k,l}}{\hat{\gamma}_{0,0}^{X,Y,k,l}} + \frac{\hat{\gamma}_1^{X,k,l}}{\hat{\gamma}_0^{X,k,l}} \right), \quad \hat{\beta}_{k,l} = \frac{1}{2} \left( \frac{\hat{\gamma}_{0,1}^{X,Y,k,l}}{\hat{\gamma}_{0,0}^{X,Y,k,l}} + \frac{\hat{\gamma}_1^{Y,k,l}}{\hat{\gamma}_0^{Y,k,l}} \right).$$

Све дефинисане оцене су такође строго постојане, с обзиром на то да представљају линеарне комбинације строго постојаних оцена.

Ако погледамо сада последњи израз, можемо приметити у првом члану оцене за параметар тининг оператора како се појављује коваријанса између компонената процеса, о чему је и било речи раније.

### 4.3.2 Оцене условне максималне веродостојности

Као и раније у раду због једноставнијег рачуна, користимо метод условне максималне веродостојности. Логаритамска функција условне максималне веродостојности може се представити као следећи производ

$$\begin{aligned}
 \log L &= \sum_{i=2}^N \log P((X_i, Y_i) = (x_i, y_i) | (X_{i-1}, Y_{i-1}) = (x_{i-1}, y_{i-1})) \\
 &= \sum_{i=2}^N \log [P(X_i = x_i | A_i) P(Y_i = y_i | A_i)] \\
 &= \sum_{i=2}^N \log P(X_i = x_i | A_i) + \sum_{i=2}^N \log P(Y_i = y_i | A_i),
 \end{aligned}$$

где смо са  $A_i$  означили догађај  $\{(X_{i-1}, Y_{i-1}) = (x_{i-1}, y_{i-1})\}$ . Извођење, потпуно аналогно оном у другој глави, даје

$$\begin{aligned}
 \log P(X_i = x_i | A_i) &= \log \left\{ \sum_{k=0}^{x_i} \binom{x_{i-1} + k - 1}{x_{i-1} - 1} \frac{\alpha_{q_{i-1}}^k}{(1 + \alpha_{q_{i-1}})^{k+x_{i-1}}} \right. \\
 &\quad \cdot \left[ \left( 1 - \frac{\alpha_{q_{i-1}} \mu_{w_{i-1}}}{\mu_{w_i} - \alpha_{q_{i-1}}} \right) \frac{\mu_{w_i}^{x_i - k}}{(1 + \mu_{w_i})^{x_i - k + 1}} \right. \\
 &\quad \left. \left. + \frac{\alpha_{q_{i-1}} \mu_{w_{i-1}}}{\mu_{w_i} - \alpha_{q_{i-1}}} \cdot \frac{\alpha_{q_{i-1}}^{x_i - k}}{(1 + \alpha_{q_{i-1}})^{x_i - k + 1}} \right] I_{\{x_{i-1} \neq 0\}} \right. \\
 &\quad \left. + \left[ \left( 1 - \frac{\alpha_{q_{i-1}} \mu_{w_{i-1}}}{\mu_{w_i} - \alpha_{q_{i-1}}} \right) \frac{\mu_{w_i}^{x_i}}{(1 + \mu_{w_i})^{x_i + 1}} \right. \right. \\
 &\quad \left. \left. + \frac{\alpha_{q_{i-1}} \mu_{w_{i-1}}}{\mu_{w_i} - \alpha_{q_{i-1}}} \cdot \frac{\alpha_{q_{i-1}}^{x_i}}{(1 + \alpha_{q_{i-1}})^{x_i + 1}} \right] I_{\{x_{i-1} = 0\}} \right\}.
 \end{aligned}$$

Исто тако је и

$$\begin{aligned}
 \log P(Y_i = y_i | A_i) &= \log \left\{ \sum_{k=0}^{y_i} \binom{y_{i-1} + k - 1}{y_{i-1} - 1} \frac{\beta_{w_{i-1}}^k}{(1 + \beta_{w_{i-1}})^{k+y_{i-1}}} \right. \\
 &\quad \cdot \left[ \left( 1 - \frac{\beta_{w_{i-1}} \eta_{q_{i-1}}}{\eta_{q_i} - \beta_{w_{i-1}}} \right) \frac{\eta_{q_i}^{y_i - k}}{(1 + \eta_{q_i})^{y_i - k + 1}} \right.
 \end{aligned}$$

$$\begin{aligned}
& + \frac{\beta_{w_{i-1}} \eta_{q_{i-1}}}{\eta_{q_i} - \beta_{w_{i-1}}} \cdot \frac{\beta_{w_{i-1}}^{y_i - k}}{(1 + \beta_{w_{i-1}})^{y_i - k + 1}} \Big] I_{\{y_{i-1} \neq 0\}} \\
& + \left[ \left( 1 - \frac{\beta_{w_{i-1}} \eta_{q_{i-1}}}{\eta_{q_i} - \beta_{w_{i-1}}} \right) \frac{\eta_{q_i}^{y_i}}{(1 + \eta_{q_i})^{y_i + 1}} \right. \\
& \left. + \frac{\beta_{w_{i-1}} \eta_{q_{i-1}}}{\eta_{q_i} - \beta_{w_{i-1}}} \cdot \frac{\beta_{w_{i-1}}^{y_i}}{(1 + \beta_{w_{i-1}})^{y_i + 1}} \right] I_{\{y_{i-1} = 0\}} \Big\}.
\end{aligned}$$

Сада се оцене непознатих параметара могу добити максимизирањем ове функције, што се у пракси као и до сада ради ну- меричким поступком.

#### 4.4 Анализа квалитета оцена над симулираним подацима

Квалитет претходно дефинисаних оцена је испитан над симулираним подацима који одговарају посматраном моделу. Симулирали смо 100 реализација обима 10000 процеса CBRNGINAR(1,  $r_1, r_2$ ) за различите вредности параметара  $r_1$  и  $r_2$ . Први корак је, наравно, симулација дводимензионалног низа стања  $\{w_n, q_n\}$ . Матрице прелаза које одговарају низовима  $\{W_n\}$  и  $\{Q_n\}$  означене су редом са  $\mathbf{p}_{mat}^{(1)}$  и  $\mathbf{p}_{mat}^{(2)}$ , док су вектори иницијалних вероватноћа означени са  $\mathbf{p}_{vec}^{(1)}$  и  $\mathbf{p}_{vec}^{(2)}$ . Као и код претходних модела ове дисертације, би- рамо такве матрице вероватноћа прелаза које имају велике ди- јагоналне елементе, како би се избегла ситуација честе промене стања. Такође, потребно је водити рачуна о томе да вредности параметара  $\alpha_1, \alpha_2, \dots, \alpha_{r_1}$  и  $\beta_1, \beta_2, \dots, \beta_{r_2}$  задовоље ограничења која су наметнута теоремом 4.2.1, при чему је пожељно да оне буду што веће могуће, како би квалитет оцена био бољи, с обзиом на то да већа вредност ових параметара имплицира већу корелацију. Сви параметри су приказани у следећим табелама. Број  $n$  представља обим реализације и узима вредности 100, 500, 1000, 5000 и 10000. Резултати при различитом избору вредности  $r_1$  и  $r_2$  су представ- љени одвојено.

1. Најпре посматрајмо случај у ком прва компонента процеса случајне средине има једно, а друга два стања, тј.  $r_1 = 1$  и

$r_2 = 2$ . Како  $\{w_n\}$  има само једно стање,  $\mathbf{p}_{vec}^{(1)}$  и  $\mathbf{p}_{mat}^{(1)}$  су једнозначно одређени, па их није потребно задавати. Низ  $\{q_n\}$  има 2 стања, па  $\mathbf{p}_{vec}^{(2)}$  и  $\mathbf{p}_{mat}^{(2)}$  имају димензије 2 тј.  $2 \times 2$ . Користићемо  $\mathbf{p}_{vec}^{(2)} = (0,5,0,5)'$  и  $\mathbf{p}_{mat}^{(2)} = \begin{bmatrix} 0,6 & 0,4 \\ 0,4 & 0,6 \end{bmatrix}$  са незнатно већим дијагоналним елементима, како бисмо обезбедили довољно дугачке подзорке из истог стања. Вредности параметара очекивања су  $\mu_1 = 2$ ,  $\eta_1 = 1$  и  $\eta_2 = 2$ . Коначно, поштујући ограничења наметнута избором претходних параметара, али тежећи избору што већих вредности параметара тининг оператора, бирамо  $\alpha_1 = 0,5$ ,  $\alpha_2 = 0,6$  и  $\beta_1 = 0,3$ . Над симулираним узорцима су израчунате Јул-Вокер оцене и резултати су приказани у табели 4.1.

Како бисмо испитали на који начин вредности параметара утичу на квалитет оцена, посматраћемо три различита случаја у оквиру овог, са  $r_1 = 1$ ,  $r_2 = 2$ . У даљем тексту ћемо претходно задате вредности параметара означавати као “основни случај”, који ћемо затим модификовати и испитати промене у резултатима.

- 1.1. У овом случају ћемо одабрати параметре као у основном случају, с тим што је  $\alpha_1 = 0,1$ ,  $\alpha_2 = 0,2$  и  $\beta_1 = 0,05$ , тј. бирамо мање вредности параметара тининг оператора. На основу табеле 4.2 можемо приметити да су оцене ових параметара сада лошије него у основном случају.
- 1.2. Овај случај се од основног разликује само по параметрима  $\mu_1 = 4$ ,  $\eta_1 = 2$ ,  $\eta_2 = 4$ , који су већи. У табели 4.3 је очигледно да су оцене очекивања горе него у основном случају.
- 1.3. Последња модификација основног случаја се односи на матрицу вероватноћа прелаза, тј.  $\mathbf{p}_{mat}^{(2)} = \begin{bmatrix} 0,1 & 0,9 \\ 0,9 & 0,1 \end{bmatrix}$ . У табели 4.4 се види да за  $n = 100$  нема оцена параметара  $\alpha_1$  и  $\beta_1$ , што је последица тога да нема довољно дугих подзорака, пошто су вероватноће остајања у истом стању сувише мале (једнаке су 0.1). Такође, све оцене су генерално горе.

На основу претходних разматрања, може се закључити да је за  $r_1 = 1$ ,  $r_2 = 2$  основни случај најбољи од свих разматраних, те ћемо и за другачије изборе  $r_1$  и  $r_2$  преостале параметре бирати на сличан начин, како бисмо добили што боље резултате, без превеликог обима узорка.

2. Сада посматрајмо ситуацију у којој обе компоненте процеса имају два стања, тј.  $r_1 = r_2 = 2$ . Иницијалне вероватноће стања су  $\mathbf{p}_{vec}^{(1)} = \mathbf{p}_{vec}^{(2)} = (0,5, 0,5)'$ , тј. у почетку су стања подједнако вероватна. Матрице  $\mathbf{p}_{mat}^{(1)}$  и  $\mathbf{p}_{mat}^{(2)}$  су једнаке  $\begin{bmatrix} 0,7 & 0,3 \\ 0,3 & 0,7 \end{bmatrix}$ , како би дијагонални елементи били довољно велики. Вредности очекивања су  $\mu_1 = 3$ ,  $\mu_2 = 2$ ,  $\eta_1 = 4$  и  $\eta_2 = 2,5$ , док су вредности параметара тининг оператора  $\alpha_1 = 0,4$ ,  $\alpha_2 = 0,35$  и  $\beta_1 = 0,45$ ,  $\beta_2 = 0,3$ . Одговарајуће Јул-Вокер оцене приказане су у табели 4.5.
3. У трећем случају је  $r_1 = 2$  и  $r_2 = 3$ . Параметри се бирају аналогним закључивањем. Вероватноће које одговарају процесу случајне средине задате су помоћу  $\mathbf{p}_{vec}^{(1)} = (0,5, 0,5)'$ ,  $\mathbf{p}_{vec}^{(1)} = (0,4, 0,3, 0,3)'$ ,  $\mathbf{p}_{mat}^{(1)} = \begin{bmatrix} 0,6 & 0,4 \\ 0,4 & 0,6 \end{bmatrix}$  и  $\mathbf{p}_{mat}^{(2)} = \begin{bmatrix} 0,6 & 0,2 & 0,2 \\ 0,2 & 0,6 & 0,2 \\ 0,2 & 0,2 & 0,6 \end{bmatrix}$ , очекивања су  $\mu_1 = 3$ ,  $\mu_2 = 3$ ,  $\eta_1 = 3$ ,  $\eta_2 = 2,5$ ,  $\eta_3 = 2$ , а параметри тининг оператора  $\alpha_1 = 0,45$ ,  $\alpha_2 = 0,25$ ,  $\alpha_3 = 0,35$  и  $\beta_1 = 0,4$ ,  $\beta_2 = 0,3$ . Сви резултати се налазе у табели 4.6. За  $n = 100$  нема оцена параметара тининг оператора, због недовољно дугих подузорака, што је узроковано превеликим бројем стања, тј. великим производом  $r_1 r_2$ .

Табела 4.1: YW оцене за случај 1.

$n$	$\hat{\mu}_1^{YW}$	$\hat{\alpha}_1^{YW}$	$\hat{\alpha}_2^{YW}$	$\hat{\eta}_1^{YW}$	$\hat{\eta}_2^{YW}$	$\hat{\beta}_1^{YW}$
100	2,0505	0,4284	0,5482	1,0592	2,0582	0,3102
ст. гр.	0,8475	0,219	0,2122	0,3920	0,7348	0,2144
500	1,9925	0,4922	0,6032	1,0263	2,0132	0,3026
ст. гр.	0,1749	0,0468	0,0537	0,0736	0,1496	0,0441
1000	1,9976	0,5058	0,6045	0,9978	2,0087	0,3063
ст. гр.	0,0887	0,0259	0,0285	0,0413	0,0755	0,0240
5000	1,9998	0,5013	0,5991	1,0083	2,0024	0,2989
ст. гр.	0,0177	0,0049	0,0053	0,008	0,0144	0,0046
10000	2,0013	0,5023	0,6014	1,0054	2,0002	0,2985
ст. гр.	0,0079	0,0024	0,0027	0,0042	0,0073	0,0022

Табела 4.2: YW оцене за случај 1.1.

$n$	$\hat{\mu}_1^{YW}$	$\hat{\alpha}_1^{YW}$	$\hat{\alpha}_2^{YW}$	$\hat{\eta}_1^{YW}$	$\hat{\eta}_2^{YW}$	$\hat{\beta}_1^{YW}$
100	2,0040	0,0965	0,2197	0,9560	1,9456	0,0491
ст. гр.	0,6843	0,1833	0,1951	0,3845	0,5618	0,1862
500	2,0090	0,0867	0,2045	0,9892	2,0098	0,0501
ст. гр.	0,1492	0,0407	0,0399	0,0691	0,1209	0,0409
1000	2,0109	0,0982	0,2015	0,9957	2,0125	0,0506
ст. гр.	0,0738	0,0207	0,0191	0,0324	0,0616	0,0201
5000	2,0037	0,0966	0,1978	0,9981	1,9907	0,0510
ст. гр.	0,0120	0,0039	0,0040	0,0069	0,0109	0,0043
10000	2,0024	0,0978	0,1980	1,0032	1,9918	0,0508
ст. гр.	0,0063	0,0019	0,0020	0,0036	0,0054	0,0019



Табела 4.3: YW оцене за случај 1.2.

$n$	$\hat{\mu}_1^{YW}$	$\hat{\alpha}_1^{YW}$	$\hat{\alpha}_2^{YW}$	$\hat{\eta}_1^{YW}$	$\hat{\eta}_2^{YW}$	$\hat{\beta}_1^{YW}$
100	3,7576	0,4303	0,5559	1,9511	4,0421	0,2624
ст. гр.	1,3534	0,1892	0,1925	0,6772	1,4425	0,1859
500	3,9362	0,4791	0,5860	1,9979	3,9681	0,2871
ст. гр.	0,2761	0,0420	0,0513	0,1414	0,2344	0,0406
1000	3,9847	0,4908	0,5922	1,9952	3,9845	0,2995
ст. гр.	0,1451	0,0234	0,0271	0,0687	0,1201	0,0220
5000	4,0235	0,5020	0,6037	1,9909	4,0040	0,2974
ст. гр.	0,0311	0,0054	0,0055	0,0116	0,0202	0,0045
10000	4,0117	0,4972	0,6008	1,9902	3,9959	0,2987
ст. гр.	0,0171	0,0026	0,0029	0,0070	0,0112	0,0020

Табела 4.4: YW оцене за случај 1.3.

$n$	$\hat{\mu}_1^{YW}$	$\hat{\alpha}_1^{YW}$	$\hat{\alpha}_2^{YW}$	$\hat{\eta}_1^{YW}$	$\hat{\eta}_2^{YW}$	$\hat{\beta}_1^{YW}$
100	3,1501		0,6864	1,1961	3,4444	
ст. гр.	3,1217		0,6503	1,1226	3,4719	
500	2,6836	0,5139	0,6773	1,2810	2,1347	0,2954
ст. гр.	1,0828	0,1084	0,1362	0,4996	0,7344	0,1297
1000	2,2024	0,4923	0,6081	1,0052	2,0022	0,2816
ст. гр.	0,4789	0,0591	0,0805	0,2332	0,4348	0,0652
5000	1,9561	0,4978	0,5944	0,9968	1,9402	0,2991
ст. гр.	0,0662	0,0142	0,0156	0,0419	0,0688	0,0137
10000	1,9647	0,4842	0,6060	0,9807	1,9706	0,3019
ст. гр.	0,0327	0,0074	0,0085	0,0229	0,0350	0,0070

Табела 4.5: YW оцене за случај 2.

$n$	$\hat{\mu}_1^{YW}$	$\hat{\mu}_2^{YW}$	$\hat{\alpha}_1^{YW}$	$\hat{\alpha}_2^{YW}$	$\hat{\eta}_1^{YW}$	$\hat{\eta}_2^{YW}$	$\hat{\beta}_1^{YW}$	$\hat{\beta}_2^{YW}$
100	3,0393	2,0638	0,3624	0,3349	3,9433	2,8160	0,4243	0,2976
ст. гр.	1,9701	1,3291	0,3076	0,3247	2,1742	1,9360	0,3231	0,2775
500	2,9438	1,9624	0,3717	0,3378	4,0585	2,5449	0,4332	0,2898
ст. гр.	0,3578	0,2361	0,0699	0,0706	0,4501	0,3302	0,0705	0,0641
1000	3,0125	1,9721	0,3796	0,3435	4,0775	2,4946	0,4468	0,2842
ст. гр.	0,1609	0,1289	0,0360	0,0356	0,2411	0,1577	0,0369	0,0315
5000	3,0207	1,9986	0,3935	0,3504	4,0368	2,5018	0,4513	0,2937
ст. гр.	0,0369	0,0245	0,0076	0,0075	0,0482	0,0311	0,0092	0,0074
10000	3,0195	1,9904	0,3996	0,3484	4,0192	2,5038	0,4511	0,2980
ст. гр.	0,0179	0,0120	0,0040	0,0040	0,0250	0,0158	0,0046	0,0036

## 4.5 Нови метод оцењивања процеса случајне средине

Нека је  $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$  узорак обима  $N$  који желимо да моделирамо CBRNGINAR( $1, r_1, r_2$ ) процесом. Важан корак у примени овог модела је одређивање помоћног низа стања  $\{(w_n, q_n)\}$ ,  $n \in \{1, 2, \dots, N\}$ , што се постиже кластеровањем података. Када се једном овај низ одреди, могуће је оценити матрице вероватноћа прелаза процеса случајне средине. Наиме, вредности  $(p_{mat})_{ls}^{(i)}$  за  $i = 1, 2$ ,  $l, s \in \{1, 2, \dots, r_i\}$  се могу једноставно оценити као количник броја преласка из стања  $l$  у стање  $s$  и броја преласка из стања  $l$  у сва стања.

Како је посматрани модел димензионалан, користићемо димензионално кластеровање. За сваку тачку  $(x_n, y_n)$ ,  $n \in \{1, \dots, N\}$  потребно је одредити  $(w_n, q_n)$ . Природно решење је кластеровање података  $\{(x_n, y_n)\}$ , међутим, ово није најбоље могуће решење, с обзиром на природу модела. Наиме, ово би било најбоље решење ако би низ  $(w_n, q_n)$  одређивао једино параметре очекивања компоненти  $X_n$  и  $Y_n$ , што овде није случај. Из конструкције модела

Табела 4.6: YW оцене за случај 3.

$n$	$\hat{\mu}_1^{YW}$	$\hat{\mu}_2^{YW}$	$\hat{\alpha}_1^{YW}$	$\hat{\alpha}_2^{YW}$	$\hat{\alpha}_3^{YW}$
100	3,3466	2,5452			0,3345
ст. гр.	3,2301	2,5198			0,4892
500	3,0311	2,2167		0,2188	0,3382
ст. гр.	0,5927	0,9411		0,0795	0,0902
1000	3,0404	2,0156	0,4320	0,2372	0,3532
ст. гр.	0,2690	0,4853	0,1056	0,0409	0,0504
5000	3,0004	2,0355	0,4627	0,2513	0,3440
ст. гр.	0,0504	0,0896	0,0231	0,0093	0,0111
10000	2,9937	2,0028	0,4668	0,2507	0,3483
ст. гр.	0,0266	0,0457	0,0136	0,0044	0,0057

$n$	$\hat{\eta}_1^{YW}$	$\hat{\eta}_2^{YW}$	$\hat{\eta}_3^{YW}$	$\hat{\beta}_1^{YW}$	$\hat{\beta}_2^{YW}$
100	3,4921	2,3965	2,1176		
ст. гр.	3,7701	2,1981	1,8155		
500	3,1193	2,3678	1,9860	0,3862	0,2922
ст. гр.	1,1793	0,4555	0,3657	0,0968	0,1543
1000	3,1691	2,4309	1,9874	0,3988	0,2998
ст. гр.	0,7628	0,2081	0,1757	0,0574	0,0977
5000	3,0101	2,4903	1,9716	0,3938	0,2936
ст. гр.	0,1120	0,0443	0,0392	0,0101	0,0171
10000	2,9673	2,4989	1,9921	0,3940	0,2901
ст. гр.	0,0543	0,0198	0,0190	0,0055	0,0097

видимо да он такође одређује параметре тининг оператора. Тачније,  $w_n$  одређује очекивање случајне променљиве  $X_n$  и параметар тининг оператора  $\beta_{w_n}$ . Зато би било добро  $w_n$  одредити узимајући у обзир и  $\mu_{w_n}$  и  $\beta_{w_n}$ . Из релације  $Y_{n+1} = \beta_{w_n} * Y_n + \eta_{n+1}$  следи да је  $E(Y_{n+1}) = \beta_{w_n} E(Y_n)$ . Зато ћемо кластеровати  $\{(x_n, \tilde{y}_n)\}$ , где је  $\tilde{y}_n = y_{n+1}/y_n$ , међутим морамо водити рачуна о томе да  $y_n$  буде различито од нуле. У случају да ово није испуњено, тј да је  $y_n = 0$ , користићемо количник максималне и минималне вредности. Дакле, дефинисаћемо  $\tilde{y}_n$  на следећи начин

$$\tilde{y}_n = \begin{cases} y_{n+1}/y_n, & \text{када } y_n \neq 0, \\ \max_{l \in \{1, \dots, N\}}(y_l) / \min_{l \in \{1, \dots, N\}}(y_l), & \text{када } y_n = 0, \end{cases}$$

за  $n \in \{1, 2, 3, \dots, N-1\}$ . За  $n = N$  можемо дефинисати  $\tilde{y}_N = y_N$ . Ако желимо да обе компоненте  $\{\tilde{y}_n\}$  и  $\{x_n\}$  имају једнак утицај на кластеровање, нормализоваћемо  $\{\tilde{y}_n\}$  тако да има исту средину као  $\{x_n\}$ . Ако, затим, желимо да контролишемо која компонента ће имати колики утицај, можемо увести коефицијент  $k \in \mathbb{R}$  тако да је

$$\tilde{y}_n = k \tilde{y}_n \frac{\sum_{m=1}^N x_m}{\sum_{m=1}^N \tilde{y}_m},$$

за  $n \in \{1, 2, \dots, N\}$ . Дакле, коефицијент  $k$  описује утицај компоненте  $\{\tilde{y}_n\}$  на кластеровање у односу на утицај компоненте  $\{x_n\}$ . Што је  $k$  веће, то је јачи утицај компоненте  $\{\tilde{y}_n\}$ , а за  $k = 1$  обе компоненте подједнако утичу. Слично, за  $n \in \{1, 2, 3, \dots, N-1\}$  дефинишимо

$$\tilde{x}_n = \begin{cases} x_{n+1}/x_n, & \text{када } x_n \neq 0, \\ \max_{l \in \{1, \dots, N\}}(x_l) / \min_{l \in \{1, \dots, N\}}(x_l), & \text{када } x_n = 0, \end{cases}$$

$\tilde{x}_N = x_N$  и за  $n \in \{1, 2, \dots, N\}$

$$\tilde{x}_n = k \tilde{x}_n \frac{\sum_{m=1}^N y_m}{\sum_{m=1}^N \tilde{y}_m}.$$

Коначно, кластеровањем  $\{(x_n, \tilde{y}_n)\}$  добијамо  $r_1$  кластера, где се сваки кластер придружује једном од  $r_1$  стања, тако да ако  $(x_n, \tilde{y}_n)$  припада  $i$ -том кластеру, бирамо  $w_n = i$ . Аналогно, кластеровањем  $\{(y_n, \tilde{x}_n)\}$  у  $r_2$  различитих кластера, одређујемо низ  $\{q_n\}$ .

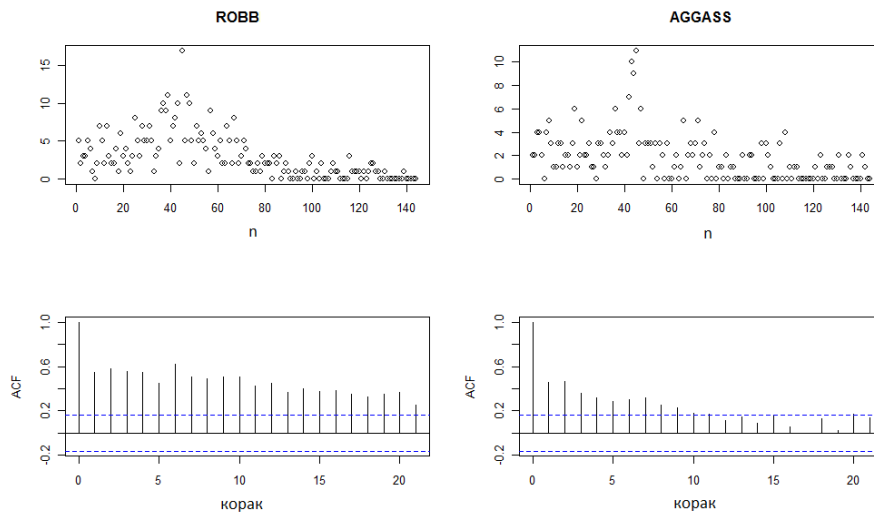
Више о избору коефицијента  $k$  у пракси биће речено у наредној секцији.

## 4.6 Примена модела над реалним подацима

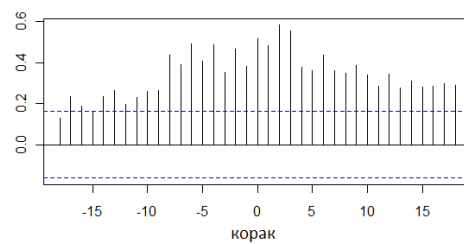
Циљ ове секције је да оправда увођење модела поређењем резултата над реалним подацима са другим дводимензионалним моделима. Коришћени су подаци као у раду Поповић, Ристић и Настич (2016), а тичу се криминалних активности у Питсбургу у периоду од јануара 1990. до децембра 2001. Посматрана је серија која представља број пљачки (ROBB, robberies) као и она која представља број тежих напада (AGGAS, aggravated assaults) из 510. полицијске станице. Графици временских серија и узорачке аутокорељације приказани су на слици 4.4, док су корелације између компонената приказане на слици 4.5. Један од модела са којим ћемо поредити уведени модел је геометријски дводимензионални INAR(1) модел са геометријском маргиналном расподелом (BVGGINAR(1)) из поменутог рада, као и његов специјални случај, модел VVEGINAR(1) који претпоставља једнакост маргиналних расподела компонената. Поред њих, у разматрање ће бити укључена два дводимензионална модела из Pedeli и Karlis (2011), која су заснована на биномном тининг оператору: BVPOIBINAR(1) који претпоставља Пуасонову расподелу иновационог процеса и BVNBIBINAR(1) који користи негативну биномну расподелу. За мерење адекватности модела у описивању одабраних података коришћена је стандардна грешка предвиђања (RMS), као у претходним главама. Оцењени параметри посматраних модела, као и одговарајуће вредности стандардне грешке предвиђања приказани су у табели 4.7.

### 4.6.1 Дискусија резултата у односу на коефицијент $k$ метода оцењивања из претходне секције

Овде ћемо применити претходно описан метод за оцењивање процеса случајне средине. Испитаћемо на који начин вредност



Слика 4.4: Графици серија и аутокорељација за податке ROBB и AGGAS.



Слика 4.5: Међусобна корелација за податке ROBB и AGGAS.

Табела 4.7: CML оцене параметара и вредност RMS за различите дводимензионалне моделе.

Модел	CML	RMS ROBB	RMS AGGASS
BVGGINAR(1)	$\hat{\alpha} = 0,499$ $\hat{\beta} = 0,281$ $\hat{p} = 0,887$ $\hat{q} = 0,805$ $\hat{a} = 2,877$ $\hat{b} = 1,765$	2496	1827
BVPOIBINAR(1)	$\hat{\alpha}_1 = 0,413$ $\hat{\lambda}_1 = 1,664$ $\hat{\alpha}_2 = 0,210$ $\hat{\lambda}_2 = 1,389$ $\hat{\phi} = 0,443$	2541	1857
BVNBIBINAR(1)	$\hat{\alpha}_1 = 0,413$ $\hat{\lambda}_1 = 1,665$ $\hat{\alpha}_2 = 0,169$ $\hat{\lambda}_2 = 1,461$ $\hat{\beta} = 0,883$	2541	1857
BVEGINAR(1)	$\hat{\alpha} = 0,468$ $\hat{\beta} = 0,289$ $\hat{p} = 0,980$ $\hat{q} = 0,625$ $\hat{a} = \hat{b} = 2,650$	2541	1880

коэффициента  $k$  утиче на RMS. У том циљу су вредности стандардне грешке предвиђања за примену CBRNGINAR(1, 3, 3) модела при различитим изборима овог коэффицијента приказане у табели 4.8. Када је вредност  $k$  јако велика, серије  $\{\tilde{x}_n\}$  и  $\{\tilde{y}_n\}$  не утичу значајно на кластероваче, па RMS за  $k \geq 10$  има константну вредност, тј. даље повећавање коэффицијента  $k$  не мења кластере. Са друге стране, за мале вредности  $k$ , утицај серија  $\{x_n\}$  и  $\{y_n\}$  постаје занемарљив и RMS има константну вредност за  $k \leq 0,01$ . Видимо да је најмања вредност стандардне грешке предвиђања негде између ова два гранична случаја, за  $k = 0,1$ .

Табела 4.8: RMS за CBRNGINAR(1, 3, 3) модел са различитим коэффицијентима  $k$ .

$k$	RMS ROBB	RMS AGGASS
100	2,527	1,746
10	2,527	1,746
1	1,929	1,321
0,5	1,730	1,269
0,1	1,279	0,807
0,01	1,399	0,808
0,001	1,399	0,808

## 4.6.2 Резултати примене

Ако погледамо графике, можемо приметити да су серије нестационарне и да изгледа као да имају најмање по три стања. Због тога смо користили CBRNGINAR(1, 3, 3) модел. Могуће је применити моделе са мањим  $r_1$  (CBRNGINAR(1, 2, 3)) или мањим  $r_2$  (CBRNGINAR(1, 3, 2)) и добити и даље добре резултате. Међутим, ако покушамо са CBRNGINAR(1, 2, 2) моделом, RMS серије ROBB ће бити 23709,295, па овај модел није добар избор. Оцене добијене методом условне максималне веродостојности и RMS вредности за коэффициент  $k = 0,1$  и моделе CBRNGINAR(1, 3, 3), CBRNGINAR(1, 2, 3) и CBRNGINAR(1, 3, 2) дати су у табелама 4.9, 4.10 и 4.11, рес-



пективно. Коришћена је следећа нотација

$$\begin{aligned}\widehat{\boldsymbol{\mu}} &= (\widehat{\mu}_1, \dots, \widehat{\mu}_{r_1})', & \widehat{\boldsymbol{\alpha}} &= (\widehat{\alpha}_1, \dots, \widehat{\alpha}_{r_2})', \\ \widehat{\boldsymbol{\eta}} &= (\widehat{\eta}_1, \dots, \widehat{\eta}_{r_2})', & \widehat{\boldsymbol{\beta}} &= (\widehat{\beta}_1, \dots, \widehat{\beta}_{r_1})'.\end{aligned}\quad (4.6.1)$$

Табела 4.9: Оцене параметара методом условне максималне веродостојности и стандардна грешка предвиђања за коефицијент  $k = 0,1$ .

CBRNGINAR(1, 3, 3)	
$\widehat{\boldsymbol{\mu}} = (0,431, 8,499, 3,223)'$	$\widehat{\boldsymbol{\eta}} = (6,818, 2,683, 0,339)'$
$\widehat{\boldsymbol{\alpha}} = (0,072, 0,018, 0,097)'$	$\widehat{\boldsymbol{\beta}} = (0,024, 0,075, 0,053)'$
$\widehat{\mathbf{p}}_{mat}^{(1)} = \begin{bmatrix} 0,763 & 0,034 & 0,203 \\ 0,000 & 0,318 & 0,682 \\ 0,242 & 0,210 & 0,548 \end{bmatrix}$	$\widehat{\mathbf{p}}_{mat}^{(2)} = \begin{bmatrix} 0,273 & 0,364 & 0,364 \\ 0,123 & 0,439 & 0,439 \\ 0,013 & 0,360 & 0,627 \end{bmatrix}$
<b>RMS ROBB</b>	<b>RMS AGGASS</b>
1,279	0,807

Табела 4.10: Оцене параметара методом условне максималне веродостојности и стандардна грешка предвиђања за коефицијент  $k = 0,1$ .

CBRNGINAR(1, 2, 3)	
$\widehat{\boldsymbol{\mu}} = (8,492, 1,861)'$	$\widehat{\boldsymbol{\eta}} = (0,588, 4,238, 0,790)'$
$\widehat{\boldsymbol{\alpha}} = (0,065, 0,270, 0,371)'$	$\widehat{\boldsymbol{\beta}} = (0,127, 0,070)'$
$\widehat{\mathbf{p}}_{mat}^{(1)} = \begin{bmatrix} 0,318 & 0,682 \\ 0,124 & 0,876 \end{bmatrix}$	$\widehat{\mathbf{p}}_{mat}^{(2)} = \begin{bmatrix} 0,000 & 0,083 & 0,917 \\ 0,049 & 0,439 & 0,512 \\ 0,111 & 0,244 & 0,644 \end{bmatrix}$
<b>RMS ROBB</b>	<b>RMS AGGASS</b>
1,727	1,218

На основу вредности стандардних грешака предвиђања примењених модела, закључује се да модел CBRNGINAR(1,  $r_1$ ,  $r_2$ ), при

Табела 4.11: Оцене параметара методом условне максималне веродостојности и стандардна грешка предвиђања за коефицијент  $k = 0,1$ .

CBRNGINAR(1, 3, 2)	
$\hat{\boldsymbol{\mu}} = (4,204, 0,837, 8,890)'$	$\hat{\boldsymbol{\eta}} = (3,350, 0,343)'$
$\hat{\boldsymbol{\alpha}} = (0,206, 0,142)'$	$\hat{\boldsymbol{\beta}} = (0,145, 0,157, 0,243)'$
$\hat{\mathbf{P}}_{mat}^{(1)} = \begin{bmatrix} 0,400 & 0,450 & 0,150 \\ 0,131 & 0,786 & 0,083 \\ 0,632 & 0,053 & 0,316 \end{bmatrix}$	$\hat{\mathbf{P}}_{mat}^{(2)} = \begin{bmatrix} 0,574 & 0,426 \\ 0,373 & 0,627 \end{bmatrix}$
<b>RMS ROBB</b>	<b>RMS AGGASS</b>
1,188	1,242

различитим одабирима вредности  $r_1$  и  $r_2$  даје најбоље резултате, у смислу најмање грешке, за посматране податке. На основу графика података, може се видети да је AGGASS серија више променљива од ROBB серије, због чега је вредност RMS мања за AGGASS серију. Ако желимо да добијемо мању вредност RMS за једну компоненту, можемо повећати одговарајући број стања. Ово се види у табели 4.12 у којој су посматрани модели CBRNGINAR(1, 4, 3) и CBRNGINAR(1, 3, 4) са коефицијентом  $k = 0,1$ .

Табела 4.12: RMS за CBRNGINAR(1, 4, 3) и CBRNGINAR(1, 3, 4) моделе са коефицијентом  $k = 0,1$ .

модел	RMS ROBB	RMS AGGASS
CBRNGINAR(1, 4, 3)	0,988	0,809
CBRNGINAR(1, 3, 4)	1,187	0,649

# Закључак

У дисертацији је дат приказ неких INAR модела у случајној средини. Због тога је најпре било речи о самим INAR моделима. Затим је приказано у другој глави на који начин увођење случајне средине може решити проблем примене у случају података са нестационарним карактером. Притом је дефинисан процес случајне средине који служи као помоћни део у изградњи модела. У дисертацији су приказани различити модели који се базирају на овој основној идеји случајне средине. Њихово увођење оправдано је применом над реалним подацима, а предност се огледа у флексибилности модела, која се постиже управо преко процеса случајне средине. Ови процеси су нестационарни, међутим, на одређеним деловима узорка могу се посматрати као стационарни. Ово својство доноси бројне олакшице у раду са њима. Дискутовано је о начину одређивања овог помоћног процеса у пракси. Најпре су дефинисани општи модели, а затим су детаљније разматрани одређени специјални случајеви ових модела. Непознати параметри су оцењивани на два начина, методом момената и методом условне максималне веродостојности. Конвергенција реализација ових оцена ка стварним вредностима потврђена је на симулираним подацима из ових модела.

Први и најједноставнији модел дискутован у овој дисертацији је првог реда и он је приказан у трећој глави. Он је такође први уведен модел овог типа у литератури. Као његово природно уопштење је дошао уопштени INAR модел вишег реда у случајној средини, који је тема четврте главе. Уопштење се огледа у томе да средина утиче на више параметара модела. Међутим, овај модел није јединствено одређен. Наиме, због одређених потешкоћа у одређивању расподеле иновационе случајне променљиве, предложена су

два типа овог модела који превазилазе тај проблем. Коначно, у петој глави је разматран дводимензионални модел који је заснован и на дводимензионалном процесу случајне средине, при чему је начин на који случајна средина утиче на модел у одређеном смислу “укрштен”, одакле потиче и назив.

Поред онога што је приказано овде, аутор је учествовао и у писању радова Laketa, Nastić (2018), као и Popović, Laketa, Nastić (2019) који у извесном смислу представљају проширење ове дисертације.

Даљи ток истраживања ове теме може ићи у више праваца. Један правац се тиче побољшања оцењивања процеса случајне средине, тј. утврђивања адекватног кластеровања података тако да примењени модел најбоље фитује одређене податке. Друго поље истраживања може се развијати задавањем неке друге маргиналне расподеле или расподеле бројачког низа укљученог у оператор истањивања. Такође, могуће је радити на побољшању постојећих INAR модела допуштањем случајности средине увођењем помоћног процеса. На крају, могуће је и сам процес случајне средине моделирати на неки други начин.

# Литература

- Al-Osh, M.A., Aly, E.E.A.A. (1992), First order autoregressive time series with negative binomial and geometric marginals, *Commun. Statist. Theory Meth.* **21**, 2483–2492.
- Al-Osh, M.A., Alzaid, A.A. (1987), First-order integer-valued autoregressive (IN-AR(1)) process, *J. Time Ser. Anal.* **8**, 261–275.
- Aly, E.A.A.A., Bouzar, N. (1994), Explicit stationary distributions for some Galton-Watson processes with immigration, *Commun. Statist. Stoch. Models* **10**, 499-517.
- Alzaid, A.A., Al-Osh, M.A. (1988), First-order integer-valued autoregressive (IN-AR(1)) process: distributional and regression properties, *Statist. Neerlandica* **42**, 53-61.
- Alzaid, A.A., Al-Osh, M.A. (1993), Some autoregressive moving average processes with generalized Poisson marginal distributions, *Ann. Inst. Statist. Math.* **45**, 223–232.
- Bakouch, H.S., Ristić, M.M. (2010), Zero truncated Poisson integer-valued AR(1) model, *Metrika* **72**, 265–280.
- Freeland, R.K., McCabe, B.P.M. (2004). Analysis of low count time series data by Poisson autoregression, *J. Time Ser. Anal.* **25**(5), 701–722.
- Jacobs, P.A., Lewis P.A.W. (1978a), Discrete time series generated by mixture I: correlational and runs properties, *J. R. Statist. Soc. B* **40**, 94–105.
- Jacobs, P.A., Lewis P.A.W. (1978b), Discrete time series generated by mixtures

II: asymptotic properties, *J. R. Statist. Soc. B* 40, 222–228.

Jacobs, P.A., Lewis P.A.W. (1978c), Discrete time series generated by mixtures III: autoregressive processes (DAR(p)), *Naval Postgraduate School Technical Report*, NPS55Lw 73061A.

Jowaheer, V., Khan, N.M., Sunecker, R.(2017), A BINAR(1) time series model with cross-correlated COM/Poisson innovations, *Commun. Stat. Theory Meth.*, 47:5, 1133-1154.

Khan, N.M., Sunecher, R., Jowaheer, V. (2016), Modelling a non-stationary BINAR(1) Poisson process, *Journal of Statistical Computation and Simulation* 86(15), 3106–3126.

Laketa, P.N. (2020), Crossed Bivariate Integer-Valued Autoregressive process based on bivariate Random Environment process, *Communications in Statistics-Theory and Methods*, прихваћен

Laketa, P.N., Nastić, A.S. (2018), Conditional least squares estimation of the parameters of higher order Random environment INAR models, *Facta Universitatis, Series: Math. Inf.*, 34(3) 525–535

Laketa, P.N., Nastić, A.S., Ristić, M.M. (2018), Generalized Random Environment INAR Models of Higher Order, *Mediterranean Journal of Mathematics*, 15:9

Latour, A. (1998), Existence and stochastic structure of a non-negative integer-valued autoregressive process, *J. Time Ser. Anal.* 19, 439–455.

MacDonald, I.L., Zucchini, W. (2009), *Hidden Markov Models for Time Series: An Introduction Using R*. CRC Press, Boca Raton.

MacQueen, J. B. (1967), Some Methods for Classification and Analysis of Multivariate Observations, *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, 1, Berkeley, CA: University of California Press, 281–297.

McKenzie, E. (1985), Some simple models for discrete variate time series, *Water*

- Resour. Bull. 21, 645–650.
- McKenzie, E. (1986), Autoregressive moving-average processes with negative binomial and geometric distributions, *Adv. Appl. Prob.* 18, 679–705.
- Möller, T.A., Weiß, C.H. (2015), Threshold models for integer-valued time series with infinite and finite range. In Steland et al. (eds.): *Stochastic Models, Statistics and Their Applications*, Springer Proceedings in Mathematics and Statistics 122, 327–334.
- Monteiro, M., Scotto, M.G., Pereira, I. (2012), Integer-valued self-exciting threshold autoregressive processes, *Commun. Stat. Theory Meth.* 41(15), 2717–2737.
- Nastić, A.S., Ristić, M.M. (2012), Some geometric mixed integer-valued autoregressive (INAR) models, *Statistics and Probability Letters* 82, 805–811.
- Nastić, A.S. Ristić, M.M., Bakouch, H.S. (2012), A combined geometric INAR( $p$ ) model based on negative binomial thinning, *Mathematical and Computer Modelling*, 55, 1665–1672.
- Nastić, A.S., Ristić, M.M., Bakouch, H.S. (2012), A combined geometric INAR( $p$ ) model based on negative binomial thinning, *Mathematical and Computer Modelling* 55, 1665–1672.
- Nastić, A.S., Laketa, P.N., Ristić, M.M. (2016), Random Environment Integer-Valued Autoregressive process, *J. Time Ser. Anal.* 37, 267–287.
- Nastić, A.S., Laketa, P.N., Ristić, M.M. (2017), Random environment INAR models of higher order, *RevStat:Statistical Journal*, 17, 35–65.
- Pedeli, X., Karlis, D. (2011), A bivariate INAR(1) process with application, *Stat. Model.* 11, 325–349.
- Popović P.M., Ristić M.M., Nastić A.S. (2016), A geometric bivariate time series with different marginal parameters, *Stat Papers* 57, 731–753.
- Popović P.M., Laketa, P.N., Nastić, A.S., (2019) Forecasting with two generalized

- integer-valued autoregressive processes of order one in the mutual random environment, *SORT*, 43:2, 355-384.
- Ristić, M.M., Bakouch, H.S., Nastić, A.S. (2009), A new geometric first-order integer-valued autoregressive (NGINAR(1)) process, *J. Stat. Plan. Inference* 139, 2218–2226.
- Ristić, M.M., Nastić, A.S., Jayakumar, K., Bakouch, H.S. (2012), A bivariate INAR(1) time series model with geometric marginals. *Appl Math Lett* 25, 481-485.
- Tang, M., Wang, Y. (2014), Asymptotic Behaviour of Random Coefficient INAR Model under Random Environment defined by Difference Equation, *Advances in Difference Equations*, 2014(1), 1–9.
- Weiß, C.H. (2008), The combined INAR(p) models for time series of counts, *Statistics and Probability Letters* 72, 1817–1822.
- Weiß, C.H. (2014), A Poisson INAR(1) Model with Serially Dependent Innovations, *Metrika*, 78(7), 829-851.
- Zheng, H., Basawa, I.V., Datta, S. (2006), Inference for  $p$ th-order random coefficient integer-valued autoregressive processes, *J. Time Ser. Anal.* 27, 411–440.
- Zheng, H., Basawa, I.V., Datta, S. (2007), First-order random coefficient integer-valued autoregressive processes, *J. Stat. Plann. Inf.* 137, 212–229.
- Zhu, R., Joe, H. (2006), Modelling count data time series with Markov processes based on binomial thinning, *J. Time Ser. Anal.* 27, 727–738.
- Zhu, R., Joe, H. (2010), Negative binomial time series models based on expectation thinning operators, *J. Stat. Plann. Inf.* 140, 1874–1888.



# Биографија

Петра Лакета је рођена 22.10.1991. године у Нишу. Завршила је основну школу “Учитељ Таса” у Нишу као Вуковац и ученик генерације. Од 2006. до 2010. похађала је и завршила као носилац Вукове дипломе Специјализовано одељење за ученике са посебним способностима за физику и природне науке у гимназији “Светозар Марковић”.

Током основне и средње школе постигла је бројне успехе на државним и савезним такмичењима из математике и физике, од којих је најзначајнији освајање бронзане медаље на Српској математичкој олимпијади.

Основне студије Математике уписала је 2010. године на Природно-математичком факултету у Нишу и исте завршила 2013. године са просечном оценом 10,00. Исте године је уписала мастер студије Примењене математике, модул Математика у физици, на истом факултету и ове студије завршила 2015. просечном оценом 10,00. Докторске студије математике уписала је 2015. године на истом факултету и досадашњи просек оцена је 10,00. Поред тога је 2010. године уписала основне студије Физике на Природно-математичком факултету у Нишу и исте завршила 2017. године са просечном оценом 10,00.

Од 2015. године је почела са радом на Природно-математичком факултету у Нишу, у звању Сарадник у настави, да би од 2016. године прешла у звање Асистента, и још увек је у истом звању. Такође, од 2016. године ради у гимназији “Светозар Марковић” са децом у Специјализованом одељењу за ученике са посебним способностима за физику и природне науке, држећи им наставу из предмета Математика.

У наставку се налази списак радова и конференција.

Радови:

1. Petra N. Laketa, Aleksandar S. Nastić, Miroslav M. Ristić, Generalized random environment INAR models of higher order, *Mediterranean Journal of Mathematics*, 2018, 15:1, art.num. 9. (M21)
2. Aleksandar S. Nastić, Petra N. Laketa, Miroslav M. Ristić, Random Environment Integer-Valued Autoregressive process, *Journal of Time Series Analysis*, 2016, 37:2, 267-287. (M22)
3. Anwar Ja'afar Mohamad Jawad, Marko D. Petković, Petra N. Laketa, Anjan Biswas Dynamics of the shallow water waves with Boussinesq equation, *Scientia Iranica (Transaction B)*, 2013, 20:1, 179-184. (M22)
4. Predrag M. Popović, Petra N. Laketa, Aleksandar S. Nastić, Forecasting with two generalized integer-valued autoregressive processes of order one in the mutual random environment, *SORT*, 2019, 43:2, 355-384. (M22)
5. Petra N. Laketa, Crossed Bivariate Integer-Valued Autoregressive process based on bivariate Random Environment process, *Communications in Statistics-Theory and Methods*, 2020. прихваћен (M23)
6. Aleksandar S. Nastić, Petra N. Laketa, Miroslav M. Ristić, Random environment INAR models of higher order, *RevStat:Statistical Journal*, 2019, 17:1, 35-65.(M23)
7. Petra N. Laketa, Aleksandar S. Nastić, Conditional least squares estimation of the parameters of higher order Random environment INAR models, *Facta Universitatis, Series: Mathematics and Informatics*, 2019, 34:3 525-535. (M51)

Конференције:

1. Aleksandar S. Nastić, Petra N. Laketa, Miroslav M. Ristić, Random Environment Integer-Valued Autoregressive process, *Applied International Conference*, Ohrid, August 21-24, 2016. (M34)
2. Aleksandar S. Nastić, Petra N. Laketa, Conditional least squares estimation of the parameters of Random environment INAR models of higher order, *The 14th Serbian Mathematical Congress*, 2018. (M34)

3. Petra N. Laketa, Predrag M. Popović, Aleksandar S. Nastić, Bivariate Integer-Valued Autoregressive Process in Random Environment, the 9th International Congress on Industrial and Applied Mathematics, Universitat de Valencia, July 15th-19th, 2019. (M34)
4. Petra N. Laketa, On random environment integer-valued autoregressive models - a survey, the 21th European Young Statisticians Meeting, Belgrade, July 29th - August 2nd, 2019. (M33)

## ИЗЈАВА О АУТОРСТВУ

Изјављујем да је докторска дисертација, под насловом

### **Ненегативни целобројни ауторегресивни процеси у случајној средини генерисани геометријским бројачким низовима**

која је одбрањена на Природно-математичком факултету Универзитета у Нишу:

- резултат сопственог истраживачког рада;
- да ову дисертацију, ни у целини, нити у деловима, нисам пријављивао/ла на другим факултетима, нити универзитетима;
- да нисам повредио/ла ауторска права, нити злоупотребио/ла интелектуалну својину других лица.

Дозвољавам да се објаве моји лични подаци, који су у вези са ауторством и добијањем академског звања доктора наука, као што су име и презиме, година и место рођења и датум одбране рада, и то у каталогу Библиотеке, Дигиталном репозиторијуму Универзитета у Нишу, као и у публикацијама Универзитета у Нишу.

У Нишу, 13.03.2020.

Потпис аутора дисертације:



---

Др Петра Н. Лакета

**ИЗЈАВА О ИСТОВЕТНОСТИ ШТАМПАНОГ И ЕЛЕКТРОНСКОГ ОБЛИКА  
ДОКТОРСКЕ ДИСЕРТАЦИЈЕ**

Наслов дисертације:

**Ненегативни целобројни ауторегресивни процеси у случајној средини генерисани  
геометријским бројачким низовима**

Изјављујем да је електронски облик моје докторске дисертације, коју сам предао/ла за уношење у **Дигитални репозиторијум Универзитета у Нишу**, истоветан штампаном облику.

У Нишу, 13.03.2020.

Потпис аутора дисертације:



---

Др Петра Н. Лакета

## ИЗЈАВА О КОРИШЋЕЊУ

Овлашћујем Универзитетску библиотеку „Никола Тесла“ да у Дигитални репозиторијум Универзитета у Нишу унесе моју докторску дисертацију, под насловом:

### **Ненегативни целобројни ауторегресивни процеси у случајној средини генерисани геометријским бројачким низовима**

Дисертацију са свим прилозима предао/ла сам у електронском облику, погодном за трајно архивирање.

Моју докторску дисертацију, унету у Дигитални репозиторијум Универзитета у Нишу, могу користити сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons), за коју сам се одлучио/ла.

1. Ауторство (CC BY)

2. Ауторство – некомерцијално (CC BY-NC)

**3. Ауторство – некомерцијално – без прераде (CC BY-NC-ND)**

4. Ауторство – некомерцијално – делити под истим условима (CC BY-NC-SA)

5. Ауторство – без прераде (CC BY-ND)

6. Ауторство – делити под истим условима (CC BY-SA)

У Нишу, 13.03.2020.

Потпис аутора дисертације:



Др Петра Н. Лакета