



**KAUNO TECHNOLOGIJOS UNIVERSITETAS  
MATEMATIKOS IR GAMTOS MOKSLŲ FAKULTETAS**

**Lina Kadikinaitė**

**PENSIJŲ FONDŲ VEIKLOS MARKOVO MODELIAI**

Baigiamasis magistro projektas

**Vadovas**  
prof. dr. Eimutis Valakevičius

Kaunas, 2016

**KAUNO TECHNOLOGIJOS UNIVERSITETAS**  
**MATEMATIKOS IR GAMTOS MOKSLŲ FAKULTETAS**

**PENSIJŲ FONDŲ VEIKLOS MARKOVO MODELIAI**  
Baigiamasis magistro projektas  
**Taikomoji matematika (621G10003)**

**Vadovas**  
prof. dr. Eimutis Valakevičius  
2016.05.27

**Recenzentas**  
doc. dr. Dalius Makackas  
2016.05.27

**Projektą atliko**  
Lina Kadikinaitė  
2016.05.27



**KAUNO TECHNOLOGIJOS UNIVERSITETAS  
MATEMATIKOS IR GAMTOS MOKSLŲ FAKULTETAS**

Lina Kadikinaityė  
Taikomoji matematika 621G10003

„Pensijų fondų veiklos Markovo modeliai“

**AKADEMINIO SAŽININGUMO DEKLARACIJA**

2016 m. gegužės mėn. 27 d.

Kaunas

Patvirtinu, kad mano, **Linos Kadikinaityės**, baigiamasis projektas tema „Pensijų fondų veiklos Markovo modeliai“ yra parašytas visiškai savarankiškai ir visi pateikti duomenys ar tyrimų rezultatai yra teisingi ir gauti sąžiningai. Šiame darbe nei viena dalis nėra plagijuota nuo jokių spausdintinių ar internetinių šaltinių, visos kitų šaltinių tiesioginės ir netiesioginės citatos nurodytos literatūros nuorodose. Įstatymų nenumatytų piniginių sumų už šį darbą niekam nesu mokėjęs.

Aš suprantu, kad išaiškėjus nesąžiningumo faktui, man bus taikomos nuobaudos, remiantis Kauno technologijos universitete galiojančia tvarka.

\_\_\_\_\_  
(vardą ir pavardę įrašyti ranka)

\_\_\_\_\_  
(parašas)

Kadikinaite, Lina. Pensijų fondų veiklos Markovo modeliai. Magistro baigiamasis projektas / vadovas prof. dr. Eimutis Valakevičius; Kauno technologijos universitetas, matematikos ir gamtos mokslų fakultetas.

Mokslo sritis: 000P Fiziniai mokslai.

Mokslo kryptis: 01P Matematika.

Reikšminiai žodžiai: Pensijų fondai, rizikos analizė, dalyvių migracija, tolydaus laiko Markovo grandinės, diskretaus laiko Markovo grandinės, homogeninis Markovo modelis, laike kintantis Markovo modelis, išplėstas laike kintantis Markovo modelis, įstrižainės korekcijos metodas, svertinės korekcijos modelis, EM algoritmas.

Kaunas, 2016. 59 p.

Lietuvos gyventojų demografinio senėjimo procesas kelia pavojų visiems pensijų fondams, todėl jų analizė, kaip niekada svarbi. Šio darbo tikslas – II pakopos pensijų fondų analizė. Projekto pradžioje pristatoma Lietuvoje veikiančių II pakopos pensijų fondų klasifikaciją ir atliekama dalyvių migracijos bei grynųjų aktyvų analizė paskutiniams 5 metams. Norint įvertinti II pakopos pensijų fondų rizikingumą siūloma analizuoti jų dalyvių migraciją naudojant Markovo grandines. Tyrimo metu pristatomi ir išbandomi šeši skirtingi modeliai šešioms skirtingoms grandinėms. Kiekvienam modeliui įvertinamos paklaidos, ir kiekvienai grandinei parenkamas geriausias modelis. Projekto išvadose pateikiama investavimo rekomendaciją įvertinus II pakopos pensijų fondų prognozes 2025 metams.

Kadikinaité, Lina. Pension funds' performance Markov models: Master's thesis in applied mathematics / supervisor prof. dr. Eimutis Valakevičius. The Faculty of Mathematics and Natural Sciences, Kaunas University of Technology.

Research area: 000P Physical sciences.

Research field: 01P Mathematics.

Key words: Pension funds, risk analysis, customers' migration, continuous time Markov chains, discrete time Markov chains, homogeneous Markov model, time-varying Markov model, extended time-varying Markov model, diagonal adjustment method, weighted adjustment method, EM algorithm.

Kaunas, 2016. 59 p.

Population ageing in Lithuania is causing serious problems for all pension funds making their analysis a foremost concern. The main objective of this thesis is the analysis of II pillar pension funds that are currently active in Lithuania. At the beginning of this thesis the reader is introduced to the classification of said pension funds as well as the analysis of the change in their net assets and number of participants during the last 5-year period. In order to estimate the risk associated with II pillar pension funds we introduce the notion of risk analysis based on participants' migrations. These migrations are introduced as six different Markov chains. Each chain is analyzed using three of six presented methods. The best method is chosen based on prediction errors. After analyzing all Markov chains the reader is presented with safest investment strategy for year 2025.

# Turinys

<b>IŽANGA</b> .....	<b>10</b>
<b>1. LITERATŪROS APŽVALGA</b> .....	<b>11</b>
<b>2. MEDŽIAGOS IR TYRIMŲ METODAI</b> .....	<b>14</b>
2.1. Diskretaus laiko Markovo grandinės .....	14
2.2. Tolydaus laiko Markovo grandinės.....	14
2.3. Markovo grandinių taikymai .....	16
2.3.1. Markovo grandinės modelis skirtas prekinio ženklo problemai spręsti .....	16
2.3.1.1. Homogeninis Markovo modelis.....	16
2.3.1.2. Laike kintantis Markovo modelis.....	16
2.3.1.3. Išplėstas laike kintantis Markovo modelis.....	17
2.3.2. Diskretaus laiko Markovo grandinės su diskrečių laiko momentų stebėjimais ...	18
2.3.3. Tolydaus laiko Markovo grandinės su diskrečių laiko momentų stebėjimais .....	18
2.3.3.1. Įstrižainės korekcijos ir svertinės korekcijos metodai.....	19
2.3.3.2. EM algoritmas .....	19
2.4. Stacionarus išsidėstymas .....	22
2.5. Paklaidų įvertinimas.....	22
2.5.1. Vidutinė absoliutinė procentinė paklaida.....	22
2.5.2. Vidutinė absoliutinė paklaida.....	23
<b>3. TYRIMŲ REZULTATAI IR JŲ APTARIMAS</b> .....	<b>24</b>
3.1. Duomenų paieška.....	24
3.2. Pensijų fondų pelningumo ir dalyvių migracijos analizė .....	25
3.3. Markovo grandinių identifikavimas.....	29
3.4. Markovo grandinių su išvestiniais duomenimis analizė .....	30
3.4.1. Pensijų fondus valdančių įmonių analizė.....	30
3.4.2. Konservatyvių pensijų fondų analizė.....	32
3.4.3. Mažos rizikos pensijų fondų analizė.....	34
3.4.4. Vidutinės rizikos pensijų fondų analizė.....	36
3.4.5. Didelės rizikos pensijų fondų analizė .....	38
3.4.6. Modelių trūkumai ir privalumai .....	40
3.5. Markovo grandinės, sudarytos atsižvelgiant į turimus duomenis, analizė .....	41
3.5.1. Įstrižainės korekcijos metodas .....	42
3.5.2. Svertinės korekcijos metodas .....	43
3.5.3. EM algoritmas .....	43
3.5.4. Modelių trūkumai ir privalumai .....	44
3.6. Programos vartotojo vadovas .....	45
<b>IŠVADOS</b> .....	<b>47</b>
<b>LITERATŪROS SĄRAŠAS</b> .....	<b>48</b>
<b>1 priedas. Funkcijos skirtos modelių realizacijai</b> .....	<b>49</b>

## Santrumpos ir žodynas

Darbe naudojamos santrumpos:

<b>Akronimas</b>	<b>Terminas</b>
<b>DLMG</b>	Diskretaus laiko Markovo grandinės
<b>MAE</b>	Vidutinė absoliutinė paklaida
<b>MAPE</b>	Vidutinė absoliutinė procentinė paklaida
<b>PF</b>	Pensijų fondas
<b>TLMG</b>	Tolydaus laiko Markovo grandinės

Naudojamos terminologijos vertimas į anglų kalbą:

<b>Lietuviškas terminas</b>	<b>Angliškas terminas</b>
<b>Diskretaus laiko Markovo grandinės</b>	<i>Discrete time Markov chains</i>
<b>Intensyvumų matrica</b>	<i>Intensity matrix / Transition rate matrix / Infinitesimal generator matrix</i>
<b>Istrižainės korekcijos metodas</b>	<i>Diagonal adjustment method</i>
<b>Nesvertinis apribotas mažiausių kvadratų metodas</b>	<i>Unweighted restricted ordinary least square technique</i>
<b>Perėjimo matrica</b>	<i>Transition matrix / probability matrix</i>
<b>Stacionarus išsidėstymas</b>	<i>Stationary distribution</i>
<b>Svertinės korekcijos metodas</b>	<i>Weighted adjustment method</i>
<b>Tolydaus laiko Markovo grandinės</b>	<i>Continuous time Markov chains</i>
<b>Vidutinė absoliutinė procentinė paklaida</b>	<i>Mean absolute percentage error</i>
<b>Vidutinė absoliutinė paklaida</b>	<i>Mean absolute error</i>
<b>Homogeninis Markovo modelis</b>	<i>Homogeneous Markov model</i>
<b>Laike kintantis Markovo modelis</b>	<i>Time-varying Markov model</i>
<b>Išplėstas laike kintantis Markovo modelis</b>	<i>Extended time-varying Markov model</i>
<b>Generatoriaus quasi - optimizacijos metodas</b>	<i>Quasi-optimization of the generator</i>

## Paveikslėlių sąrašas

1.1 pav. skirtingų pakopų pensijų fondų galimų pajamų palyginimas ir dabartinių pajamų palyginimas.....	11
2.1 pav. Markovo grandinė su trimis būsenomis.....	18
3.1 pav. Analizei naudota informacija apie pensijų fondų dalyvius.....	25
3.2 pav. Markovo grandinė atspindinti pensijų fondus valdančių įmonių dalyvių migraciją.....	31
3.3 pav. Markovo grandinė atspindinti konservatyvių pensijų fondų dalyvių migraciją.....	33
3.4 pav. Markovo grandinė atspindinti mažos rizikos pensijų fondų dalyvių migraciją.....	35
3.5 pav. Markovo grandinė atspindinti vidutinės rizikos pensijų fondų dalyvių migraciją.....	37
3.6 pav. Markovo grandinė atspindinti didelės rizikos pensijų fondų dalyvių migraciją.....	39
3.7 pav. Markovo grandinė atspindinti įmonės UAB „DNB investicijų valdymas“ pensijų fondų dalyvių migraciją.....	42



## Lentelių sąrašas

3.1 lentelė. Analizuojamų II pakopos pensijų fondų sąrašas .....	24
3.2 lentelė. Konservatyvių pensijų fondų grynujų aktyvų ir dalyvių pokyčiai .....	27
3.3 lentelė. Mažos rizikos pensijų fondų grynujų aktyvų ir dalyvių migracijos analizė.....	28
3.4 lentelė. Vidutinės rizikos pensijų fondų grynujų aktyvų ir dalyvių migracijos analizė.....	28
3.5 lentelė. Didelės rizikos pensijų fondų grynujų aktyvų ir dalyvių migracijos analizė .....	29
3.6 lentelė. Rinkos dalies pasiskirstymo tarp pensijų fondus valdančių įmonių 2015 metų istorinių duomenų ir prognozės palyginimas .....	30
3.7 lentelė. Rinkos dalies pasiskirstymo tarp pensijų fondus valdančių įmonių prognozių paklaidų įverčiai .....	31
3.8 lentelė. Rinkos dalies pasiskirstymas tarp pensijų fondus valdančių įmonių 2011 ir 2015 metais bei prognozė 2025 metams .....	32
3.9 lentelė. Rinkos dalies pasiskirstymo tarp konservatyvių pensijų fondų 2015 metų istorinių duomenų ir prognozės palyginimas .....	33
3.10 lentelė. Rinkos dalies pasiskirstymo tarp konservatyvių pensijų fondų prognozių paklaidų įverčiai.....	34
3.11 lentelė. Rinkos dalies pasiskirstymas tarp konservatyvių pensijų fondų 2011 ir 2015 metais bei prognozė 2025 metams .....	34
3.12 lentelė. Rinkos dalies pasiskirstymo tarp mažos rizikos pensijų fondų 2015 metų istorinių duomenų ir prognozės palyginimas .....	35
3.13 lentelė. Rinkos dalies pasiskirstymo tarp mažos rizikos pensijų fondų prognozių paklaidų įverčiai .....	35
3.14 lentelė. Rinkos dalies pasiskirstymas tarp mažos rizikos pensijų fondų 2011 ir 2015 metais bei prognozė 2025 metams .....	36
3.15 lentelė. Rinkos dalies pasiskirstymo tarp vidutinės rizikos pensijų fondų 2015 metų istorinių duomenų ir prognozės palyginimas .....	36
3.16 lentelė. Rinkos dalies pasiskirstymo tarp vidutinės rizikos pensijų fondų prognozių paklaidų įverčiai.....	37
3.17 lentelė. Rinkos dalies pasiskirstymas tarp vidutinės rizikos pensijų fondų 2011 ir 2015 metais bei prognozė 2025 metams .....	38
3.18 lentelė. Rinkos dalies pasiskirstymo tarp didelės rizikos pensijų fondų 2015 metų istorinių duomenų ir prognozės palyginimas .....	38
3.19 lentelė. Rinkos dalies pasiskirstymo tarp didelės rizikos pensijų fondų prognozių paklaidų įverčiai .....	39
3.20 lentelė. Rinkos dalies pasiskirstymas tarp didelės rizikos pensijų fondų 2011 ir 2015 metais ir prognozė 2025 metams .....	40
3.21 lentelė. Stacionarių tikimybių įverčiai naudojant įstrižainės korekcijos metodą .....	42
3.22 lentelė. Stacionarių tikimybių įverčiai naudojant svertinės korekcijos metodą.....	43
3.23 lentelė. Stacionarių tikimybių įverčiai naudojant EM algoritmą .....	44
3.24 lentelė. Modelių realizacijose naudojami parametrai .....	45

## IŽANGA

Vadovaujantis Lietuvos Statistikos departamento duomenimis per paskutinius 15 metų kiekvieną mėnesį Lietuvoje mirė daugiau žmonių, nei gimė naujagimių. Tai reiškia, kad artimiausiu metu į darbo rinką ateis mažiau darbuotojų, nei išeis į pensiją. Ši neigiama kaita kelia rimtą iššūkį visiems Lietuvoje veikiančioms pensijų fondams. Fondai, kurių augimas neigiamas gali atsidurti situacijoje, kai aktyvūs dalyviai taps nepajėgūs išlaikyti pasyvius. Dėl šios priežasties šiame darbe pristatysime pensijų fondų rizikos įvertinimą pagrįstą dalyvių migracija.

Šiame darbe pensijų fondų analizei siūloma naudoti Markovo grandinės. Jos leidžia modeliuoti dalyvių migraciją yra pakankamai universalios ir pasižymi stochastiškumu.

Darbo eigoje aptariama Lietuvos respublikoje egzistuojanti pensijų sistema ir išskiriami skirtingi pensijų fondų tipai. Pirmas skyrius skirtas literatūros apžvalgai. Antrame skyriuje pristatoma tolimesniuose skyriuose naudojama metodika. Trečiame skyriuje aprašoma tyrimo eiga, pritaikytų metodų rezultatai ir identifikuojamas geriausias pensijų fondas. Skyriaus pabaigoje pateikiamas sukurtų programų vartotojo vadovas.

Darbo tikslai ir uždaviniai:

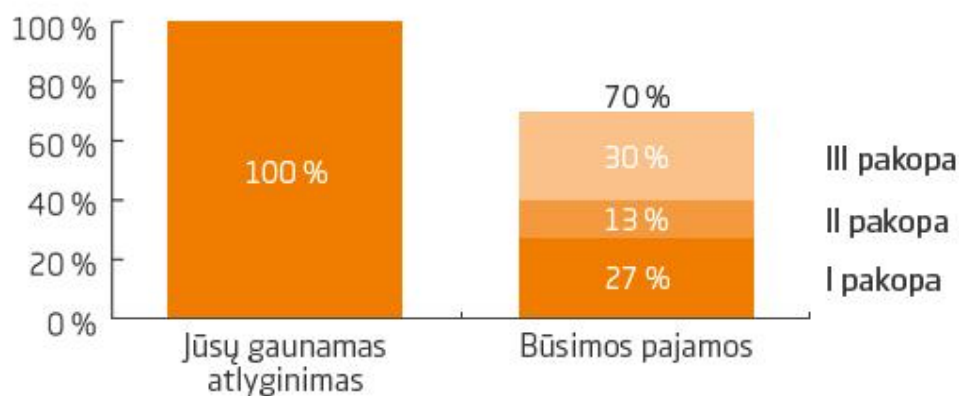
- Susipažinti su Lietuvos Respublikos pensijų sistema;
- Atlikti II pakopos pensijų fondų analizę;
- Peržvelgti galimas pensijų kaupimo strategijas;
- Išnagrinėti Markovo grandinėmis pagrįstą metodiką, kurią galima naudoti pensijų fondų analizėje;
- Sukurti programinę įrangą skirtą pensijų fondų analizei atlikti;
- Išnagrinėti modelių tikslumą;
- Pateikti prognozę 2025 metams;
- Rekomenduoti pensijų kaupimo strategiją.

## 1. LITERATŪROS APŽVALGA

Šiame darbe analizuojami pensijų fondai (toliau – PF), tai investiciniai fondai, kurių tikslas – užtikrinti finansinį stabilumą pasiekus pensinį amžių ir išėjus į pensiją. Tokio tipo investiciniai fondai gali būti privatūs arba valstybiniai. Lietuvoje valstybinį pensijos fondą valdo „Sodra“ ir egzistuoja keletas privačių fondų. Pagal kaupimo specifiką jie skirstomi į pirmos, antros ir trečios pakopos pensijinius fondus:

- Pirmos pakopos PF, tai fondas, kuriame socialinės draudimo įmokos pervedamos tik „Sodrai“;
- Antros pakopos PF, tai fondas į kurį pervedama dalis socialinio draudimo įmokų, šiuo metu mokamų „Sodrai“. Šio fondo dalyviai ateityje gaus pensiją sudarytą iš dviejų dalių: socialinio draudimo ir iš savo sukauptų lėšų. Tokio tipo fondų pavyzdžiai – „DANSKE pensija 100“, „INVL MEZZO II 53+“ ir „SEB pensija 1“;
- Trečios pakopos pensijų fondai skiriasi nuo antros pakopos tuo, kad į juos pinigai pervedami savarankiškai pagal su valdymo įmone pasirašytą sutartį. Įmokos gali būti periodinės ir neperiodinės bei jas gali mokėti ne tik dalyvis, bet ir kiti asmenys, kaip pavyzdžiui, sutuoktiniai ar darbdaviai. Tokio tipo fondų pavyzdžiai – „Danske pensija plus“, „INVL III akcijų pensijų fondas“ ir „SEB pensija 1 plus“.

Daugiau informacijos apie skirtingus pensijų kaupimo metodus ir PF galima rasti šaltiniuose [1] ir [2]. Skirtingų pakopų pensijų fondų pajamų procentinis pasiskirstymas pateiktas žemiau esančiame paveikslėlyje.



1.1 pav. skirtingų pakopų pensijų fondų galimų pajamų palyginimas ir dabartinių pajamų palyginimas<sup>1</sup>

PF pasidaro naudingi, kai pensinį amžių pasiekęs žmogus nustoja dirbti. Jų tikslas – užtikrinti žmogaus finansinį saugumą jam nustojus gauti pastovias darbovietės išmokamas

<sup>1</sup> [https://www.swedbank.lt/images/Privatiems/taupymas\\_investavimas/pensijos\\_A3\\_grafikas\\_1.jpg](https://www.swedbank.lt/images/Privatiems/taupymas_investavimas/pensijos_A3_grafikas_1.jpg)

pajamas. Iš čia išplaukia, būtinybė įvertinti ir prižiūrėti pensijų fondus. Darykime prielaidą, kad vidutinis asmuo Lietuvoje pradeda dirbti būdamas 23 metų (įgijęs bakalauro laipsnį) ir į pensiją išeina pasiekęs pensinį amžių. Norint supaprastinti analizę, tarkime, kad tai yra 65 metai, vyrams ir moterims. Tokiu atveju žmogus 42 metus kaupia pinigus, kurie jam bus išmokėti pasiekus pensinį amžių. Taigi, žmogaus sugebėjimas apmokėti savo sąskaitas, nusipirkti maisto ar drabužių išėjus į pensiją tiesiogiai priklauso nuo fondų išmokų egzistavimo [3], [4], [5].

Šiame darbe pensijų fondų analizė atliekama naudojant Markovo grandines. Jos dažnai taikomos finansinių priemonių, eilių ir įvairių rinkų analizei dėl savo universalumo ir stochastiškumo [6]. Vienas iš Markovo grandinių taikymo pavyzdžių pateiktas šaltinyje [7]. Šio straipsnio autorius aptaria detalų rinkos dalies prognozavimo metodo realizavimo pavyzdį. Analizės metu naudojama homogeninė Markovo grandinė, kurios būsenos – pagrindinės konkuruojančių gamintojų kompiuterių plokštės. Kartu su prognoze kiekvienam laiko intervalui, taip pat pateikiamas stacionarus išsidėstymas atitinkantis gautus duomenis ir detales gautų reikšmių paaiškinimas.

Straipsniuose [8] ir [9] pateikti keturi Markovo grandinių modeliai, kurie tai pat yra skirti prognozuoti produktų užimama rinkos dalį ateityje. Straipsniuose autorius lygina ir analizuoja modelius, iš kurių vienas nėra paremtas Markovo grandinėmis. Kiekvieno modelio realizavimui pateiktas aiškus ir detalus pavyzdys, leidžiantis skaitytojui įsigilinti į modelius ir suprasti jų taikymo principą. Atlikus prognozę šešioms skirtingiems laiko intervalams, skaitytojui sudaroma galimybė pamatyti prognozuojamų dydžių kitimą laiko eilutėje. Straipsnio pabaigoje pateikiamos kiekvieno modelio savybės ir jų palyginimas. Autorius pabrėžia savo turimos grandinės analizės svarbą parenkant naudojamus modelius.

Kitas konkurencingas rinkos įvertinimo metodas pateiktas straipsnyje [10]. Straipsnio autoriaus pristatomas metodas, taip pat naudoja perėjimo tikimybių matricą ir jos dėka prognozuoja skalbimo miltelių rinkos pasiskirstymą įmonėms – „Ariel“, „Tix“, „Persil“, „Tomi“, „Omo“, „Biopon“. Naudojami 600 respondentų atsakymų duomenys. Autorius išskiria skirtingose perėjimo tikimybių matricos pozicijose esančių tikimybių interpretaciją, kuri atspindi rinkos padėtį. Išvadose pateikiama rinkos analizės nauda, reklamos efektingumo bei kokybės pokyčio interpretacija.

Toliau verta atkreipti dėmesį į G. Datong straipsnį [11]. Jame aptariama teisingos tikimybių matricos įvertinimo svarba. Norėdamas pasiekti geriausius rezultatus, autorius tikimybių perėjimo matricą įvertina naudodamas nesvertinį apribotą mažiausių kvadratų metodą. Pasinaudodamas šiuo metodu autorius apskaičiuoja tikimybių perėjimo matricą ir pateikia įžvalgas apie rinkos kitimą ateityje bei galimas rekomendacijas rinkos dalyviams.

Aukščiau išvardyti darbai pateikia rinkos įvertinimo modelius paremtus diskrečios Markovo grandinės perėjimo tikimybių matrica. Įvertinus analizuojamų dydžių ypatumus (dalyviai gali migruoti iš vienos būsenos į kitą bet kuriuo laiko momentu, tačiau informacija pateikiama fiksuotais laiko momentais) susidaro įspūdis, kad turimos tolydaus laiko Markovo grandinės. Norint jas teisingai įvertinti reikia pasitelkti metodus, kurie gali iš diskrečių duomenų įvertinti tolydžios Markovo grandinės parametrus.

Vienas iš tokių metodų detalai aprašytas straipsnyje [12]. Šio straipsnio autoriai pateikia du būdus skirtus įvertinti intensyvumų matricą: įstrižinės ir svertinės korekcijos metodus. Metodai logaritmuoja tikimybių perėjimo matricą ir atsižvelgdama į intensyvumų matricos savybes atlieka atitinkamas korekcijas. Autoriai palygina modelių pritaikymo efektyvumą su senesniais modeliais ir pateikia detalius jų siūlomų metodų įrodymus.

Abiejų metodų taikymas ganėtinai paprastas, kadangi jis nereikalauja sudėtingų algoritmų taikymo ir ilgo skaičiavimo. Straipsnyje [13] pateikiamas šių metodų tęsinys – generatoriaus quasi-optimizacijos metodas. Algoritmo tikslas – rasti mažiausią skirtumą tarp įvertinamos intensyvumų ir logaritmuotos perėjimo tikimybių matricų Darbe pateikiamas šio metodo skaičiavimo algoritmas pritaikytas greitam skaičiavimui.

Ketvirtas metodas, kuris sugeba įvertinti tolydžios Markovo grandinės intensyvumo matricą randamas autorių Bladt ir Sorensen straipsnyje [14] Autoriai šiai problemai pritaiko du metodus: EM ir Monte Carlo įvertinimo algoritmus. Straipsnyje pateikiamas modifikuotas algoritmas ir modifikavimo priežastys. Šio darbo tikslas – pateikti aiškų ir suprantama metodą intensyvumų matricai įvertinti.

## 2. MEDŽIAGOS IR TYRIMŲ METODAI

### 2.1. Diskretaus laiko Markovo grandinės

Šiame skyriuje apibrėšime Markovo grandines, kuriose procesas savo būseną keičia diskrečiais laiko momentais (toliau – DLMG). Šio tipo Markovo grandinės apibūdinamos naudojant du dydžius – laiko momentą ir žingsnio numerį.

- Tarkime, kad atsitiktinis procesas turi  $K$  būsenų. Visų galimų būsenų aibę žymėsime  $S = \{s_1, s_2, \dots, s_K\}$ .
- Proceso būseną laiko momentu  $t_n, n = 1, \dots, T$  yra žymima, kaip  $X(t_n)$ .

Proceso  $X = \{X(t_n) | n = 1, \dots, T\}$  judėjimas tarp būsenų, priklausančių aibei  $S$ , apibrėžiamas perėjimo tikimybių matrica  $P(t_n)$ .

$$P(t_n) = \begin{bmatrix} p_{11}(t_n) & p_{12}(t_n) & \dots & p_{1K}(t_n) \\ p_{21}(t_n) & p_{22}(t_n) & \dots & p_{2K}(t_n) \\ \vdots & \vdots & \ddots & \vdots \\ p_{K1}(t_n) & p_{K2}(t_n) & \dots & p_{KK}(t_n) \end{bmatrix} \quad (1)$$

Šioje matricoje  $p_{ij}(t_n) = P(X(t_n) = j | X(t_{n-1}) = i)$ . Toliau šią tikimybę vadinsime perėjimo iš  $i$ -osios būsenos į  $j$ -ąją būseną tikimybe. Aišku, kad  $\sum_j^K p_{ij}(t_n) = 1$ , kai sumuojama pagal visas galimas būsenas. Matricos  $P(t_n)$  pagalba galima rasti  $n$ -žingsninę perėjimo tikimybių matricą.

#### 2.1 Apibrėžimas

Perėjimo iš būsenos  $s_i$  į būseną  $s_j$  lygiai per  $m$  žingsnių tikimybė yra vadinama  $m$ -žingsnine perėjimo tikimybe ir žymima  $p_{ij}^{(m)}$ . Atitinkama matrica, sudaryta iš  $p_{ij}^{(m)}$  elementų yra vadinama  $m$ -žingsnine perėjimo tikimybių matrica.

#### 2.1 Teorema

Jei  $P = P(t_n)$  yra vienažingsnė Markovo grandinės (kurios perėjimo tikimybės nepriklauso nuo žingsnio numerio, o priklauso tik nuo to, iš kokios į kokią būseną vyksta perėjimas) perėjimo tikimybių matrica, tai  $P^m$  yra  $m$ -žingsninė perėjimo tikimybių matrica [6].

### 2.2. Tolydaus laiko Markovo grandinės

Tolydaus laiko Markovo grandinės (toliau – TLMG) yra DLMG tęsinys, jų skirtumas – laiko momentas. Diskrečiu atveju procesas savo būsenas keičia diskrečiu, o tolydžiu atveju – atsitiktiniu laiko momentu.

Kaip ir DLMG atveju, TLMG yra apibūdinamas stochastiniu procesu  $X = \{X(t)|0 \leq t\}$ , kuris visiems  $t \geq 0, s \geq 0$  ir  $i, j \in S$  tenkina lygtį (2).

$$P(X(s+t) = j|X(s) = i, \{X(u): 0 \leq u < s\}) = P(X(s+t) = j|X(s) = i) \quad (2)$$

## 2.2 Apibrėžimas

TLMG procesas vadinamas homogeniniu, jeigu tenkina lygtį,

$$P(X(s+t) = j|X(s) = i) = P(X(t) = j|X(0) = i).$$

TLMG apibūdinimui yra naudojama:

- Proceso būsenų aibė -  $S = \{s_1, s_2, \dots, s_K\}$ ;

$$\mathbf{Q} = \begin{bmatrix} -q_1 & q_{12} & \dots & q_{1K} \\ q_{21} & -q_{22} & \dots & q_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ q_{K1} & q_{K2} & \dots & -q_{KK} \end{bmatrix} \quad (3)$$

- Intensyvumų matrica  $\mathbf{Q}$ .

Intensyvumų matrica  $\mathbf{Q}$  tenkina šias savybes:

1.  $0 \leq -q_{ii} < \infty$ ;
2.  $0 \leq q_{ij}$ : kai  $i \neq j$ ;
3.  $\sum_j q_{ij} = 0$ : visiems  $i$ .

Perėjimo tikimybių matricą galime apibrėžti, kaip

$$\mathbf{P}(t, t + \Delta t) = \mathbf{I} + \Delta t \mathbf{Q} + o(\Delta t); \quad (4)$$

čia  $\mathbf{I}$  yra vienetinė matrica.  $M$ -tojo žingsnio perėjimo tikimybių matrica apskaičiuojama panašiai [6]. Tarkime, kad  $s = t + m\Delta t$ , tuomet

$$\begin{aligned} \mathbf{P}(t, s) &\approx (\mathbf{I} + \Delta t \mathbf{Q})^m \\ &= \left( \mathbf{I} + \frac{s-t}{m} \mathbf{Q} \right)^m \end{aligned} \quad (5)$$

ir kai  $m \rightarrow \infty$

$$\mathbf{P}(t, s) = \exp((s-t)\mathbf{Q}). \quad (6)$$

## 2.3. Markovo grandinių taikymai

### 2.3.1. Markovo grandinės modelis skirtas prekinio ženklo problemai spręsti

Šio modelio tikslas – įvertinti klientų lojalumą ir prognozuoti klientų elgesį. Analizės metu įvertinami istoriniai duomenys, o gauti rezultatai atspindi modelio ateities prognozę. Sekančiame skyriuje aptarsime tris modelio įvertinimo algoritmus:

- Homogeninis Markovo modelis;
- Laike kintantis Markovo modelis;
- Išplėstas laike kintantis Markovo modelis.

#### 2.3.1.1. Homogeninis Markovo modelis

Šiame modelyje klientai tarp būsenų juda fiksuotais laiko momentais. Tai pat jie gali pasilikti esamoje būsenoje arba pereiti į kitą būseną (pradėti naudotis konkurento paslaugomis). Modelio įvertinimui naudojami du dydžiai – perėjimo tikimybių matrica  $P(t_n)$ , kuri pateikta lygtyje (1) ir rinkos dalies vektorius  $Y_t$ , kuris apibrėžiamas, kaip

$$Y_t = [y_t(s_1), y_t(s_2), \dots, y_t(s_{K-1}), y_t(s_K)]. \quad (7)$$

Vektoriaus  $Y_t$  elementai  $y_t(s_1)$  įvertinami apskaičiuojant tiekėjo  $s_i$  rinkos dalį laiko momentu  $t$ . Kadangi įvertinama rinkos dalis, tai  $\sum_{i=1}^K y_t(s_i) = 1$ . Norint rasti kito etapo rinkos dalį,  $Y_{t+1}$ , reikia sudauginti  $Y_t$  ir  $P$ , kur  $P$  yra šio etapo perėjimų tikimybių matrica. Ji turi būti pastovi skirtingoms  $t$  reikšmėms [8], [9].

Norint įvertinti  $t + 1$  etapo rinkos dalies vektorius naudojant esamo etapo,  $t = 1$ , duomenis naudojamos žemiau pateiktos išraiškos,

$$\begin{aligned} Y_2 &= Y_1 P, \\ Y_3 &= Y_2 P = (Y_1 P) P = Y_1 P^2, \\ Y_{t+1} &= Y_1 P^t. \end{aligned} \quad (8)$$

#### 2.3.1.2. Laike kintantis Markovo modelis

Kai  $P$  matrica kinta laike, tai homogeninis modelis nebetinka ir tada naudojamas laike kintantis Markovo modelis. Šis modelis naudoja modifikuotą (8) formulę, kur



$$\begin{aligned}
Y_2 &= Y_1 P_1, \\
Y_3 &= Y_2 P_2 = Y_1 P_1 P_2, \\
Y_{t+1} &= Y_1 \prod_{n=1}^t P_n.
\end{aligned} \tag{9}$$

Formulėje (9) norint įvertinti  $Y_{t+1}$  kiekvienam laiko periodui  $t$  reikia apskaičiuoti  $P_n$ . Tai galima atlikti naudojant žemiau pateiktas išraiškas [8], [9].

$$\begin{aligned}
P_{t+1} &= P_t + \Delta P; \\
\Delta P(i, j) &= \delta p_{ij}; \\
\Delta p_{ij} &= \frac{p_{ij}(t)}{\sum_{i=1}^K p_{ij}(t)} - \frac{p_{ij}(t-1)}{\sum_{i=1}^K p_{ij}(t)}.
\end{aligned} \tag{10}$$

Sujungus išraiškas (9) ir (10), gauname galutinę modelio išraišką,

$$Y_{t+1} = Y_1 \prod_{n=1}^t P_n = Y_1 \prod_{n=1}^t (P_1 + n\Delta P). \tag{11}$$

### 2.3.1.3. Išplėstas laike kintantis Markovo modelis

Abu anksčiau minėti modeliai skirti įvertinti tiekėjo  $i$  rinkos dalį laiko momentu  $t + 1$ . Šio modelio privalumas – išsamesnė prognozė. Įvertinama ne tik rinkos dalis, bet ir tikslus klientų priklausančių tiekėjui  $i$  kiekis laiko momentu  $t + 1$ .

Šiam modeliui pirmiausiai reikia įvertinti rinkos augimo koeficientą  $G$ ,

$$G = [g_1, g_2, \dots, g_K]. \tag{12}$$

Sekantis modelio žingsnis – populiacijos vektoriaus  $Q_1$  įvertinimas. Kito etapo populiacijos vektorius įvertinamas naudojant žemiau pateiktas formules,

$$\begin{aligned}
Q_2 &= G \odot Q_1 P_1 = [g_1 q_1(s_1), g_2 q_1(s_2), \dots, g_K q_1(s_K)] P_1, \\
Q_3 &= G \odot Q_2 P_2 = G \odot G \odot Q_1 P_1 P_2 \\
&= [g_1^2 q_1(s_1), g_2^2 q_1(s_2), \dots, g_K^2 q_1(s_K)] P_1 P_2, \\
Q_{t+1} &= G^{\odot t} \odot Q_1 \prod_{n=1}^t P_n = [g_1^t q_1(s_1), g_2^t q_1(s_2), \dots, g_K^t q_1(s_K)] P_1 P_2 \dots P_t, \\
Q_{t+1} &= G^{\odot t} \odot Q_1 \prod_{n=1}^t (P_1 + n\Delta P),
\end{aligned} \tag{13}$$

kur  $\odot$  žymi matricų daugybą, kur matricos  $C = A \odot B$   $c_{ij}$  elementas randamas sudauginus atitinkamus matricų  $A$  ir  $B$  elementus, t.y.  $c_{ij} = a_{ij} b_{ij}$  [8], [9].

### 2.3.2. Diskretaus laiko Markovo grandinės su diskrečių laiko momentų stebėjimais

Tai lengviausias skaičiavimo būdas. Pirmas sprendimo žingsnis – tikėtinumo funkcijos sudarymas. Jos išraiška pateikta žemiau.

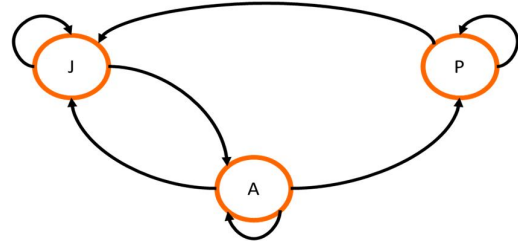
$$L(\mathbf{P}) = \prod_{i=1}^K \prod_{j=1}^K p_{ij}^{N_{ij}(m)}, \quad (14)$$

kur  $N_{ij}(m)$  – perėjimų iš būsenos  $i$  į būseną  $j$  skaičius. Iš čia gaunama perėjimo tikimybių įvertinio išraiška [6]:

$$\widehat{p}_{ij} = \frac{N_{ij}(m)}{\sum_{j=1}^K N_{ij}(m)}. \quad (15)$$

#### Pavyzdys

Turime 2.1. pav. pavaizduotą Markovo grandinę su trimis būsenomis, taigi  $S = \{s_1 = J, s_2 = A, s_3 = P\}$  ir  $J \rightarrow J$  perėjo 10 dalyvių ir  $N_{11} = 10$ ,  $J \rightarrow A = 5 = N_{12}$ ,  $A \rightarrow A = 20 = N_{22}$ ,  $A \rightarrow J = 15 = N_{21}$ ,  $A \rightarrow P = 1 = N_{23}$ ,  $P \rightarrow P = 2 = N_{33}$ ,  $P \rightarrow J = 20 = N_{31}$ .



2.1 pav. Markovo grandinė su trimis būsenomis

Turint šiuos duomenis galime įvertinti perėjimo tikimybes,

$$\begin{aligned} \widehat{p}_{11} &= \frac{N_{11}}{\sum_{j=1}^K N_{1j}} = \frac{10}{10 + 5} = \frac{2}{3} \approx 0.67, \widehat{p}_{12} \approx 0.33, \widehat{p}_{13} = 0, \\ \widehat{p}_{21} &\approx 0.42, \widehat{p}_{22} \approx 0.55, \widehat{p}_{23} \approx 0.03, \\ \widehat{p}_{31} &\approx 0.91, \widehat{p}_{32} = 0, \widehat{p}_{33} \approx 0.09, \end{aligned} \quad (16)$$

ir gaunama perėjimo tikimybių matrica,

$$\mathbf{P} = \begin{bmatrix} 0.67 & 0.33 & 0 \\ 0.42 & 0.55 & 0.03 \\ 0.91 & 0 & 0.09 \end{bmatrix}. \quad (17)$$

### 2.3.3. Tolydaus laiko Markovo grandinės su diskrečių laiko momentų stebėjimais

Šiame skyriuje aptarsime tris metodus skirtus įvertinti intensyvumų matricą, kai dirbama su stebėjimais užfiksuotais diskrečiais laiko momentais.

### 2.3.3.1. Įstrižainės korekcijos ir svertinės korekcijos metodai

Pirmas šio metodo žingsnis - perėjimo matricos logaritmo radimas. Šiame darbe bus naudojamas [12] pateiktas metodas, kuriame

$$\log(\widehat{\mathbf{P}}) = (\widehat{\mathbf{P}} - \mathbf{I}) - \frac{(\widehat{\mathbf{P}} - \mathbf{I})^2}{2} + \frac{(\widehat{\mathbf{P}} - \mathbf{I})^3}{3} - \frac{(\widehat{\mathbf{P}} - \mathbf{I})^4}{4} + \dots \quad (18)$$

ir

$$\widehat{\mathbf{Q}} = \frac{1}{s-t} \log(\log(\widehat{\mathbf{P}}(t,s))). \quad (19)$$

Sudarius  $\widehat{\mathbf{Q}}$  gauname, kad matrica netenkina anksčiau apibrėžtų jos savybių. Šiai problemai išspręsti taikomas įstrižainės korekcijos ir svertinės korekcijos metodai.

Įstrižainės korekcijos metodo algoritmas,

1.  $\hat{q}_{ij} = \begin{cases} 0, & \text{jeigu } i \neq j \text{ ir } q_{ij} < 0 \\ q_{ij}, & \text{kitais atvejais} \end{cases}$ , čia  $q_{ij}$  – matricos  $\widehat{\mathbf{Q}}$  elementai, o  $\hat{q}_{ij}$  – matricos  $\widehat{\mathbf{Q}}_{IK}$  elementai;
2.  $\hat{q}_{ii}^{IK} = -\sum_{j=1, i \neq j}^K \hat{q}_{ij}$ , visiems  $i = 1, 2, \dots, K$ .

Svertinės korekcijos metodo algoritmas:

1.  $\hat{q}_{ij} = \begin{cases} 0, & \text{jeigu } i \neq j \text{ ir } q_{ij} < 0 \\ q_{ij}, & \text{kitais atvejais} \end{cases}$ , čia  $q_{ij}$  – matricos  $\widehat{\mathbf{Q}}$  elementai, o  $\hat{q}_{ij}$  – matricos  $\widehat{\mathbf{Q}}_{SK}$  elementai;
2.  $\hat{q}_{ii}^{SK} = \hat{q}_{ij} - |\hat{q}_{ij}| \left( \frac{\sum_{j=1}^K \hat{q}_{ij}}{\sum_{j=1}^K |\hat{q}_{ij}|} \right)$  visiems  $i = 1, 2, \dots, K$ . [12]

### 2.3.3.2. EM algoritmas

Naudojant šį algoritmą yra siekiama rasti didžiausią tikėtimumo funkcijos reikšmę. Algoritmas sprendžiamas inicijuojant du žingsnius:

- Žingsnis E (tikėtimumo žingsnis). Žingsnio tikslas – apskaičiuoti įvertinį trūkstamiems dydžiams;
- Žingsnis M (didžiausios vertės paieškos žingsnis). Naudojant E žingsnio rezultatus įvertinami parametrai, kurie toliau bus naudojami žingsnyje E.

Algoritmą galima pradėti tiek nuo E, tiek nuo M žingsnio priklausomai, kokius pradinis duomenis turime. Yra įrodyta, kad jeigu apskaičiuotume tikėtimumo reikšmę su kiekviename

algoritmo cikle gautu parametru, tai šios tikėtinumo reikšmės sudarytų nemažėjančią seką. Taigi, EM algoritmas maksimizuoja tikėtinumo funkciją tik lokaliai.

Kadangi algoritmas konverguoja į lokalių maksimumą svarbu ir pradinės matricos  $\mathbf{Q}$  parinkimas. Tam atlikti naudojami keli inicializavimo būdai: atliekamas  $k$ -vidurkių klasterizavimas ir gautas sprendinys naudojamas kaip pradinis EM algoritmo parametras; arba pradinis parametras parenkamas atsitiktinai, ir algoritmas atliekamas iš kelių pradinių taškų. Galutinis sprendinys pasirenkamas pagal didžiausią tikėtinumo funkcijos reikšmę.

EM algoritmo taikymas:

Tikėtinumo funkcijos logaritmas pateiktas žemiau esančioje lygtyje,

$$E[\log L(\mathbf{Q}) | x^{steb}] = \sum_{i=1}^K \sum_{j \neq i} \log(q_{ij}) E[N_{ij}(T) | x^{steb}] - \sum_{i=1}^K \sum_{j \neq i} q_{ij} E[R_i(T) | x^{steb}] \quad (20)$$

M žingsnio įverčiai gaunami naudojantis formulę,

$$\hat{q}_{ij} = \frac{E[N_{ij}(T) | x^{steb}]}{E[R_j(T) | x^{steb}]} \quad (21)$$

Šios sąlyginės tikimybės gali būti išreikštos, kaip

$$\begin{aligned} E[N_{ij}(T) | x^{steb}] &= \sum_h^N E[N_{ij}^h(T) | x^h] \text{ ir } E[R_i(T) | x^{steb}] \\ &= \sum_h^N E[R_i^h(T) | x^h]. \end{aligned} \quad (22)$$

Lygtyje (22)  $E[N_{ij}^h(T) | x^h]$  ir  $E[R_i^h(T) | x^h]$  yra sąlyginis perėjimų iš būsenos  $i$  į būseną  $j$  vidurkis ir sąlyginis laiko, praleisto būsenoje  $i$  vidurkis, kai turime diskretaus laiko stebėjimus  $x^h$ . Iš to seka, kad sunkiausia algoritmo dalis – šių dydžių įvertinimas E žingsnyje. Pažymėkime  $x^h(t_{n+1})$  ir  $x^h(t_n)$  stebimo proceso  $h$  būseną laiko momentais  $t_{n+1}$  ir  $t_n$ , tuomet pagal Markovo savybę gauname kad,

$$\begin{aligned} E[R_i^h(T) | x^h] &= \sum_{n=1}^{T-1} E[R_i^h(\Delta t) | X^h(t_{n+1}) = x^h(t_{n+1}), X^h(t_n) \\ &= x^h(t_n)] \end{aligned} \quad (23)$$

ir

$$\begin{aligned} E[N_{ij}^h(T)|x^h] &= \sum_{n=1}^{T-1} E[N_{ij}^h(\Delta t)|X^h(t_{n+1}) = x^h(t_{n+1}), X^h(t_n) \\ &= x^h(t_n)]. \end{aligned} \quad (24)$$

Tuomet  $E[N_{ij}^h(T)|x^h]$  ir  $E[R_i^h(T)|x^h]$  skaičiavimas tampa sąlyginio vidurkio skaičiavimu laiko periodui tarp stebėjimų momentų  $t_{n+1}$  ir  $t_n$ . Jeigu laikysime, kad  $e_i$  yra  $K$  ilgio vektorius, kurio visi elementai yra lygūs nuliui, o elementas  $i$  yra lygus vienam, tai išraišką (23) galime užrašyti taip:

$$\begin{aligned} &\sum_{n=1}^{T-1} E[R_i^h(\Delta t)|X^h(t_{n+1}) = x^h(t_{n+1}), X^h(t_n) = x^h(t_n)] \\ &= \sum_{n=1}^{T-1} \frac{1}{D} e_{x^h(t_n)}^T \left( \int_{t_n}^{t_{n+1}} \exp((s \right. \\ &\quad \left. - t_n)\mathbf{Q}(e_i e_i^T) \exp((t_{n+1} - s)\mathbf{Q}) ds) e_{x^h(t_{n+1})} \right). \end{aligned} \quad (25)$$

Tada (25) išraiška tampa tokia:

$$\begin{aligned} &\sum_{n=1}^{T-1} E[N_{ij}^h(\Delta t)|X^h(t_{n+1}) = x^h(t_{n+1}), X^h(t_n) = x^h(t_n)] \\ &= \sum_{n=1}^{T-1} \frac{1}{D} e_{x^h(t_n)}^T (q_{ij} \int_{t_n}^{t_{n+1}} \exp((s \\ &\quad - t_n)\mathbf{Q}(e_i e_j^T) \exp((t_{n+1} - s)\mathbf{Q}) ds) e_{x^h(t_{n+1})} \right). \end{aligned} \quad (26)$$

(25) ir (26) lygtyse  $D$  yra lygus

$$D = e_{x^h(t_n)}^T \exp((t_n - t_{n+1})\mathbf{Q}) e_{x^h(t_{n+1})}. \quad (27)$$

EM algoritmo taikymo žingsniai:

1. Pasirinktu metodu nustatoma pradinė  $\mathbf{Q}_0$  matrica ir ji prilyginama matricai  $\mathbf{Q}_k$ ;
2. Visiems stebėjimams apskaičiuojamos (25) ir (26) reikšmės;

3. Įvertinami  $\widehat{q}_{ij} = \frac{E[N_{ij}(T)|x^{stab}]}{E[R_j(T)|x^{stab}]}$  visoms  $i$  ir  $j$  reikšmėms. Gauta matrica  $\widehat{Q}$  prilyginama matricai  $Q_k$  ir grįžtama prie 2 žingsnio;
4. Procesas kartojamas, kol pasiekiami stabilūs rezultatai. [14]

## 2.4. Stacionarus išsidėstymas

Tarkime, kad nagrinėjamam procesui egzistuoja ribinės tikimybės

$$p_i = \lim_{t \rightarrow \infty} p_i(t). \quad (28)$$

Tokiu atveju bėgant laikui sistemoje nusistovi stacionarus režimas, kurio metu ji pereina iš vienos būsenos į kitą, bet būsenų tikimybės jau nesikeičia.

DTMG stacionarios tikimybės  $\pi_i$  randamos išsprendus lygtį

$$\pi = P\pi. \quad (29)$$

TLMG stacionarios tikimybės  $\pi_i$  randamos išsprendus lygtį

$$\pi Q = 0. \quad (30)$$

Abiem atvejais  $\pi_i > 0$  ir  $\sum_{i=1}^K \pi_i = 1$ . [15]

## 2.5. Paklaidų įvertinimas

Modelio paklaidų įvertinimas reikalingas norint nustatyti metodo tinkamumą. Šiame darbe paklaidų įvertinimui bus naudojami du paklaidų įverčiai: vidutinė absoliutinė procentinė paklaida (toliau – MAPE) ir vidutinė absoliutinė paklaida (toliau – MAE).

### 2.5.1. Vidutinė absoliutinė procentinė paklaida

Paklaida apskaičiuojama pagal formulę,

$$MAPE = \frac{1}{n} \sum_{t=0}^{n-1} \frac{|T - T'|}{T} * 100; \quad (31)$$

kur  $T$  – tikra reikšmė,  $T'$  – prognozės rezultatai.

Paklaidos reikšmių interpretacija:

- Iki 10 % didelis tikslumas;
- 10 – 20 % geras tikslumas;
- 20 – 50 % pakankamas tikslumas;

- 50 % + blogas tikslumas.

### 2.5.2. Vidutinė absoliutinė paklaida

Paklaida apskaičiuojama naudojantis žemiau pateikta formule,

$$MAE = \frac{1}{n} \sum_{t=0}^{n-1} |T' - T|. \quad (32)$$

### 3. TYRIMŲ REZULTATAI IR JŲ APTARIMAS

Tyrimo tikslas – rasti rizikingiausių ir mažiausiai rizikingą PF. Tyrimo metu buvo analizuojamos penkios įmonės, valdančios 2 pakopos PF. Įmonių ir PF sąrašas pateiktas 3.1 lentelėje

#### 3.1 lentelė. Analizuojamų II pakopos pensijų fondų sąrašas

Pensijų fondą valdanti įmonė	II pakopos pensijų fondas
UAGDPB „AVIVA Lietuva“	Aviva Europensija Aviva Europensija ekstra Aviva Europensija plus
UAB „Danske Capital investicijų valdymas“	Danske Pensija 100 Danske Pensija 50 Konservatyvaus valdymo Danske pensija
UAB „DNB investicijų valdymas“	DNB Pensija 1 DNB Pensija 2 DNB Pensija 3
UAB „INVL Asset Management“	INVL STABILO II 58+ INVL EXTREMO II 16+ INVL MEDIO II 47+
UAB „SEB investicijų valdymas“	SEB Pensija 1 SEB Pensija 2 SEB Pensija 3
UAB „Swedbank investicijų valdymas“	Swedbank Pensija 1 Swedbank Pensija 2 Swedbank Pensija 3 Swedbank Pensija 4 Swedbank Pensija 5

#### 3.1. Duomenų paieška

Pirmas tyrimo etapas – duomenų paieška. Siekiant gauti duomenis reikalingus analizei buvo susisiepta su visomis 3.1 lentelėje išvardytomis įmonėmis, Lietuvos Respublikos centriniu banku, Lietuvos Respublikos statistikos departamentu ir Sodra. Duomenis tyrimui pateikė Lietuvos Respublikos centrinis bankas, Lietuvos Respublikos statistikos departamentas ir UAB "DNB investicijų valdymas".

Toliau buvo apdorojami turimi duomenys. Lietuvos Respublikos centrinis bankas ir UAB "DNB investicijų valdymas" pateikė duomenis, esančius 3.1 pav. Lentelėje pateikti duomenys atspindi PF dalyvių migraciją. Didžioji dalis analizuojamų PF visuomenei šiuos duomenis pateikia kiekvienų kalendorinių metų veiklos ataskaitoje, tačiau minėti šaltiniai suteikė galimybę



analizuoti mažesnio laikotarpio duomenis ir tokiu būdu atkurti trūkstamus duomenis. Taip pat buvo naudojami duomenys apie grynujų aktyvų pokytį, infliacijos lygį Lietuvoje ir papildomi

<b>Dalyvių skaičius ataskaitinio laikotarpio pradžioje</b>		
<b>Dalyvių skaičius ataskaitinio laikotarpio pabaigoje</b>	dalyviai, pasinaudoję teise nukelti pensijų išmokos mokėjimą	
	dalyviai, gaunantys periodines pensijų išmokas	
	kiti dalyviai	
	Iš viso	
<b>Dalyvių skaičiaus pokytis</b>		
<b>Bendras prisijungusių dalyvių skaičius</b>		
<b>Dalyviai, pensijų kaupimo sutartis sudarę pirmą kartą</b>		
<b>Iš kitų pensijų fondų atėję dalyviai</b>	iš tos pačios pensijų kaupimo bendrovės	
	iš kitos pensijų kaupimo bendrovės	
<b>Bendras pasitraukusių dalyvių skaičius</b>		
<b>Iš kitus pensijų fondus išėjusių dalyvių skaičius</b>	valdomus tos pačios pensijų kaupimo bendrovės	
	valdomus kitos pensijų kaupimo bendrovės	
<b>Baigusią dalyvavimą dalyvių skaičius</b>	išstojusių (pildoma tik papildomo savanoriško pensijų kaupimo pensijų fondo atveju)	
	sulaukusių nustatyto pensijos amžiaus ir pasinaudojusių teise į pensijų išmoką (vienkartinę išmoką, anuitetą ir jų derinį)	
	dalyvių, pasinaudojusių Pensijų kaupimo įstatymo suteikta galimybe vienašališkai nutraukti pirmą kartą sudarytą sutartį	
	mirusių dalyvių	

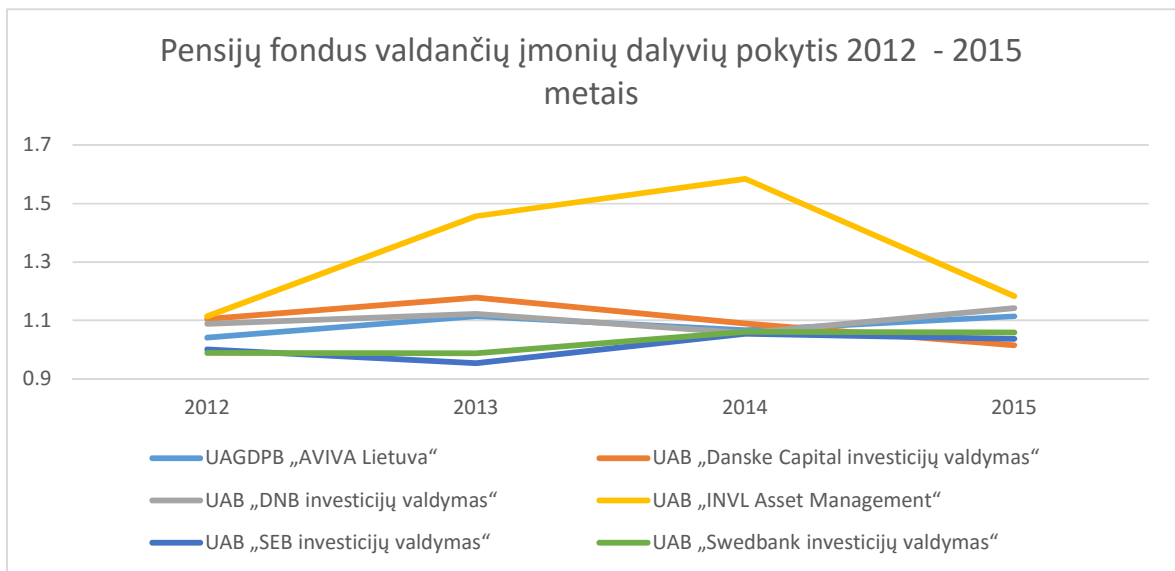
3.1 pav. Analizei naudota informacija apie pensijų fondų dalyvius

Lietuvos Respublikos centrinio banko teikiami išvestiniai duomenys apie PF dalyvių judėjimą.

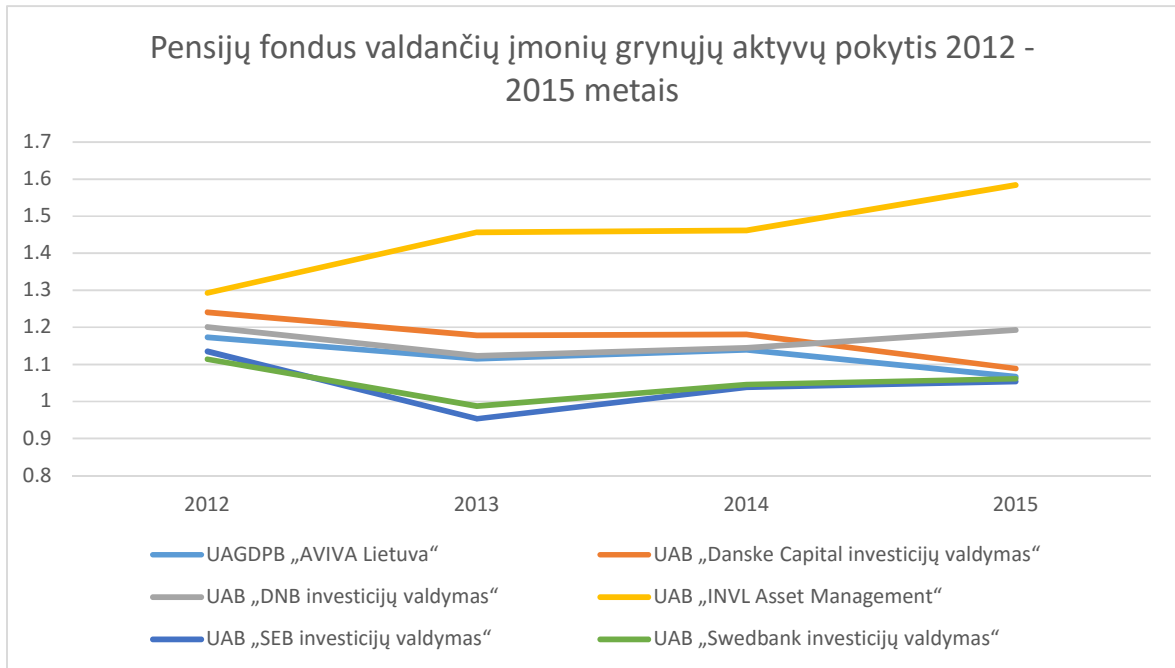
### 3.2. Pensijų fondų pelningumo ir dalyvių migracijos analizė

Šiame tyrimo etape analizuojami PF ir juos valdančių įmonių dalyvių grynujų aktyvų pokyčiai 2011 – 2015 metais. Grynieji aktyvai, tai PF turto vertės ir ilgalaikių bei trumpalaikių finansinių įsipareigojimų, susijusių su visomis valdymo ir administravimo išlaidomis (išskyrus įsipareigojimus PF dalyviams) skirtumas. 2015 metais vidutinė grynujų aktyvų dalis vienam PF dalyviui buvo 1 771 Eur., o šiuo metu vidutinė senatvės pensija Lietuvoje yra 255,15 Eur. Tai reiškia, kad vienam PF dalyviui priklausanti grynujų aktyvų dalis yra mažesnė nei septynios vidutinės senatvės pensijos išmokos.

Žemiau pateiktuose grafikuose matome PF valdančių įmonių ir PF dalyvių, bei grynujų aktyvų pokytį lyginant su praeitu laikotarpiu.



3.1 grafikas. Pensijų fondus valdančių įmonių dalyvių pokytis 2012 – 2015 metais.



3.2 grafikas. Pensijų fondus valdančių įmonių grynujų aktyvų vertės pokytis 2012 – 2015 metais.

Iš šių grafikų matome, kad UAB „INVL Asset Management“ įmonės aktyvai augo lėčiau nei dalyvių skaičius. Tai reiškia, kad 2014 metais grynujų aktyvų kiekis vienam dalyviui sumažėjo. Tačiau 2015 metais šio fondo grynujų aktyvų augimas buvo didžiausias. Remiantis istoriniais duomenimis žinome, kad 2014 metais grynujų aktyvų kiekis vienas asmeniui buvo 17 % mažesnis nei 2013 metais. Įmonės, kurių grynujų aktyvų pokytis ir grynujų aktyvų pokytis vienam asmeniui analizuojamu laikotarpiu nesumažėjo – UAB „Danske Capital investicijų

valdymas“, UAB „Aviva Lietuva“ ir UAB „DNB investicijų valdymas“. O UAB „Aviva Lietuva“ patyrė stabiliausią grynujų aktyvų ir dalyvių augimą.

Toliau apskaičiuojama analizuojamų fondų statistiką. Tolimesniuose tyrimo etapuose atsiradus poreikiui ji bus naudojama, kaip papildomas kriterijus, bei atskaitinis taškas rezultatų interpretacijai.

Analizuoti PF vadovavosi šiomis investavimo strategijomis:

- Konservatyvaus investavimo strategija – šios investavimo strategijos PF savo turtą investuoja tik į Lietuvos Respublikos centrinio banko ar kitų šalių centrinių bankų išleistus vertybinius popierius [16]. Žemiau esančioje lentelėje pateiktas konservatyvių PF sąrašas ir duomenys apie dalyvių ir grynujų aktyvų pokytį analizuojamu laikotarpiu;

**3.2 lentelė.** Konservatyvių pensijų fondų grynujų aktyvų ir dalyvių pokyčiai

<b>Pensijų fondas</b>	<b>Veiklos pradžia</b>	<b>Grynujų aktyvų vertės pokytis</b>	<b>Dalyvių pokytis</b>	<b>Grynujų aktyvų vertės pokytis vienam dalyviui (įvertinus infliaciją) ↓</b>
<b>Konservatyvaus valdymo Danske pensija</b>	2006 m. balandžio 6 d.	+ 90 % (+ 78 %)	+ 16 %	+ 54 %
<b>Aviva Europensija</b>	2004 m. birželio 15 d	+ 58 % (+ 48 %)	+ 1 %	+ 46 %
<b>DNB Pensija 1</b>	2003 m. rugsėjo 11 d	+ 224 % (+ 204 %)	+ 129 %	+ 33 %
<b>Swedbank Pensija 1</b>	2003 m. rugpjūčio 28 d	+ 4 % (- 2 %)	- 23 %	+ 27 %
<b>INVL STABILO II 58+</b>	2003 m. rugpjūčio 28 d	+ 1911 % (+ 1784 %)	+ 2551 %	- 29 %

- Mažos akcijų dalies investavimo strategija (iki 30 proc. lėšų investuojama į akcijas) – šio tipo PF yra laikomi mažai rizikingais, kadangi juose investicijos į akcijas siekia iki 20-30 proc. viso jų turto ir labiausiai tinka investuotojams, kurie nori didesnio pajamingumo, tačiau nenori prisiimti papildomos rizikos [16]. Duomenys apie šio tipo fondus pateikti 3.3 lentelėje.

### 3.3 lentelė. Mažos rizikos pensijų fondų grynujų aktyvų ir dalyvių migracijos analizė

Pensijų fondas	Veiklos pradžia	Grynujų aktyvų vertės pokytis	Dalyvių pokytis	Grynujų aktyvų vertės pokytis vienam dalyviui (įvertinus infliaciją) ↓
Swedbank Pensija 2	2003 m. rugpjūčio 28 d.	+ 39 % (+ 30 %)	- 15 %	+ 54 %
Aviva Europensija plus	2004 m. birželio 15 d.	+ 111 % (+ 97 %)	+ 33 %	+ 49 %
DNB Pensija 2	2004 m. birželio 15 d.	+ 107 % (+ 94 %)	+ 45 %	+ 33 %

- Vidutinės akcijų dalies investavimo strategija (į akcijas investuojama 30 – 70 proc. viso turto) – šios strategijos besilaikantys PF yra laikomi vidutinės rizikos PF, jie labiausiai tinka vidutinio amžiaus klientams [16]. Žemiau esančioje lentelėje pateikti duomenys apie šio tipo fondų grynujų aktyvų ir dalyvių pokytį 2011 – 2015 metais;

### 3.4 lentelė. Vidutinės rizikos pensijų fondų grynujų aktyvų ir dalyvių migracijos analizė

Pensijų fondas	Veiklos pradžia	Grynujų aktyvų vertės pokytis	Dalyvių pokytis	Grynujų aktyvų vertės pokytis vienam dalyviui (įvertinus infliaciją) ↓
Swedbank Pensija 4	2005 m. lapkričio 25 d.	+ 138 % (+ 123 %)	+ 15 %	+ 94 %
Aviva Europensija ekstra	2006 m. vasario 8 d.	+ 176 % (+ 158 %)	+ 43 %	+ 80 %
Swedbank Pensija 3	2003 m. rugpjūčio 28 d.	+ 69 % (+ 59 %)	- 3 %	+ 63 %
Danske pensija 50	2004 m. birželio 15 d.	+ 102 % (+ 89 %)	+ 18 %	+ 61 %
SEB Pensija 2	2003 m. rugpjūčio 28 d.	+ 62 % (+ 51 %)	+ 2%	+ 48 %
DNB Pensija 3	2003 m. rugsėjo 11 d.	+ 204 % (+ 185 %)	+ 108 %	+ 37 %
INVL MEDIO II 47+	2007 m. rugsėjo 20 d.	+ 446 % (+ 411 %)	+ 344 %	+ 15 %

- Akcijų investavimo strategija (visas turtas investuojamas į akcijas) – tai rizikingiausia investavimo strategija, kurioje investicijos į akcijas siekia nuo 70 iki 100 proc., tokio tipo PF renkasi investuotojai, kurie nori didžiausio pajamingumo ir

yra mažai jautrūs rizikai. Šie PF fondai dažniausiai rekomenduojami į darbo rinką įžengusiems asmenims [16]. Detalesnė informacija apie šio tipo PF pateikta žemiau esančioje lentelėje.

**3.5 lentelė.** Didelės rizikos pensijų fondų grynujų aktyvų ir dalyvių migracijos analizė

Pensijų fondas	Veiklos pradžia	Grynujų aktyvų vertės pokytis	Dalyvių pokytis	Grynujų aktyvų vertės pokytis vienam dalyviui (įvertinus infliaciją) ↓
<b>INVL EXTREMO II 16+</b>	2007 m. rugsėjo 24 d.	+ 341 % (+313 %)	+ 131 %	+ 79 %
<b>SEB Pensija 3</b>	2006 m. kovo 27 d.	+ 126 % (+ 112 %)	+ 30 %	+ 63 %
<b>Danske pensija 100</b>	2004 m. birželio 15 d.	+ 141 % (+ 126 %)	+ 55 %	+ 46 %
<b>Swedbank Pensija 5</b>	2011 m. balandžio 21 d.	+ 1312 % (+ 1223 %)	+ 3155 %	- 59 %

Kiekvienam asmeniui rekomenduojama keisti PF pereinant iš labiau rizikingo į mažiau rizikingą besikeičiant jo poreikiams ir aplinkybėms. Tai reikštų, kad kiekvienas asmuo turėtų pradėti akcijų tipo PF ir iš jo pereiti į vidutinės akcijų dalies PF, o išeiti į pensiją būnantis mažos akcijų dalies PF.

Rekomenduojamas PF keitimo amžius:

- Iki 40 metų – didelės rizikos PF (akcijos sudaro 70 – 100 proc. portfelio);
- Nuo 36 iki 60 metų – vidutinės rizikos PF (akcijos sudaro 30 – 70 proc. portfelio);
- Nuo 46 iki 60 metų – mažos rizikos PF (akcijos sudaro iki 30 proc. portfelio);
- Nuo 60 metų – konservatyvus PF [17].

### 3.3. Markovo grandinių identifikavimas

Šio etapo metu buvo nustatomos galimos proceso būsenos ir sudaromos Markovo grandinės. Sudarant pirmuosius modelius – įtraukiami tik tie dalyviai, kurie lieka sistemoje, tai yra, ignoruojami PF klientai, kurie dėl įvairių priežasčių pasitraukė iš PF. Pirmuoju atveju dalyvis migruoja tarp PF valdančių įmonių, o antruoju – tarp skirtingų fondų, priklausančių tai pačiai rizikos grupei. Šios analizės tikslas – nustatyti optimaliausią PF keitimo seką.

Iš pateiktų istorinių duomenų, galima įvertinti tikimybes, jog dalyvis nekeis PF, taigi, likusios perėjimo tikimybės parenkamos atsižvelgiant į turimus duomenis. Įėjimo ar išėjimo iš fondo tikimybės apskaičiuojamos, naudojant žinomas išėjimo į kitus fondus reikšmes ir jas išskirstant analizuojamiems fondams, atsižvelgiant į  $(t - 1)$  laikotarpio augimo tendencijas.

Trečias analizuojamas modelis - UAB "DNB investicijų valdymas" įmonės valdomų PF dalyvių judėjimas. Šio modelio būsenos nustatytos atsižvelgiant į turimus duomenis. Verta pastebėti, jog šį modelį galima pritaikyti ir visai pensijų sistemai apibrėžiant papildomas būsenas.

### 3.4. Markovo grandinių su išvestiniais duomenimis analizė

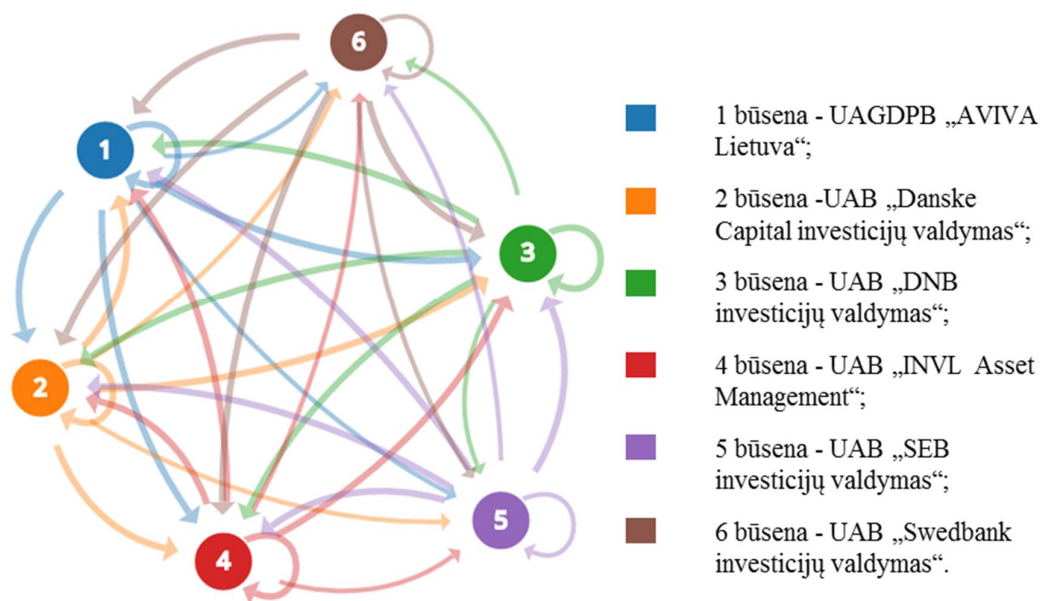
Šiame poskyryje taikomi 2.3.1. poskyryje pateikti metodai ir analizuojami gauti rezultatai.

#### 3.4.1. Pensijų fondus valdančių įmonių analizė

Pirmas analizuojamas modelis pavaizduotas 3.2 pav. Jis skirtas prognozuoti klientų kiekį PF valdančiose įmonėse. Prognozei naudojami 2011 m. ir 2012 m. duomenys. Gauti rezultatai bus lyginami su 2015 m. duomenimis. Atliktos prognozės rezultatai pateikti 3.6 lentelėje.

**3.6 lentelė.** Rinkos dalies pasiskirstymo tarp pensijų fondus valdančių įmonių 2015 metų istorinių duomenų ir prognozės palyginimas

<b>Pensijų fondą valdanti įmonė</b>	<b>UAGDPB „AVIVA Lietuva“</b>	<b>UAB „Danske Capital investicijų valdymas“</b>	<b>UAB „DNB investicijų valdymas“</b>	<b>UAB „INVL Asset Management“</b>	<b>UAB „SEB investicijų valdymas“</b>	<b>UAB „Swedbank investicijų valdymas“</b>
<b>Metodas</b>						
<b>2015 m. duomenys</b>	17%	2%	11%	9%	22%	39%
<b>Homogeninis Markovo modelis</b>	18%	3%	9%	5%	25%	40%
<b>Laike kintantis Markovo modelis</b>	17%	3%	9%	6%	25%	40%
<b>Išplėstas laike kintantis Markovo modelis</b>	18%	2%	10%	5%	25%	40%



3.2 pav. Markovo grandinė atspindinti pensijų fondus valdančių įmonių dalyvių migraciją<sup>2</sup>

Aukščiau esančioje lentelėje pateikiami prognozės rezultatai, kurie parodo kokia rinkos dalis ateityje priklausys PF valdančiai įmonei. Skirtumas tarp prognozių atliktų skirtingais metodais tai pačiai įmonei – keli procentai. Tai parodo, kad naudojami modeliai yra tinkami, tačiau norint parinkti geriausią vertinimo būdą reikia įvertinti paklaidų įverčius, jie pateikti 3.7 lentelėje.

**3.7 lentelė.** Rinkos dalies pasiskirstymo tarp pensijų fondus valdančių įmonių prognozių paklaidų įverčiai

Metodas	Paklaidų įvertis	MAPE (vidutinė absoliutinė procentinė paklaida)	MAE (vidutinė absoliutinė paklaida)
Homogeninis Markovo modelis		26,6 %	1,8
Laike kintantis Markovo modelis		25,8 %	2
Išplėstas laike kintantis Markovo modelis		17 %	1,6

Įvertinus modelių paklaidas geriausias metodas – išplėstas laike kintantis Markovo modelis. Modelio privalumas – statistiškai reikšmingai mažesnė MAPE paklaida. Lyginant su kitais modeliais ji pusantro karto mažesnė, o MAE paklaida visiems modeliams panaši. Tačiau šios paklaidos skirtos įvertinti rinkos dalies prognozei, o išplėstas laike kintantis Markovo modelis prognozuoja ir kiekvienos PF valdančios įmonės dalyvių skaičių. Pasak modelio atliktos prognozės 2015 m. PF turės 1 055 582 klientus, kadangi mes šiuos skaičius žinome galime įvertinti paklaidą. Remiantis istoriniais duomenimis 2015m. PF veikloje dalyvavo 1 176 471

<sup>2</sup> Vaizdas sugeneruotas naudojant vizualizacijos įrankį „Markov Chains“. Įrankį galima rasti internetiniu adresu – <http://setosa.io/markov>

asmenys, tai reiškia, kad modelio paklaida – 120 889 klientai, t.y. 10%. Dėl šios priežasties išplėstas laike kintantis Markovo modelis yra atmetamas ir iš likusių dviejų išrenkamas modelis, kurio paklaidos yra mažiausios. Abiejų likusių modelių MAE paklaidos vienodos, tačiau laike kintančio Markovo modelio MAPE paklaida vienu punktu mažesnė. Taigi, galime teigti, kad šis metodas yra tinkamiausias PF valdančių įmonių analizei, kai turimas 3.2 pav. pavaizduotas modelis.

Parinkus tinkamiausią modelį atliekama prognozė 2025 m. Šios prognozės rezultatai pateikti 3.8 lentelėje.

**3.8 lentelė.** Rinkos dalies pasiskirstymas tarp pensijų fondus valdančių įmonių 2011 ir 2015 metais bei prognozė 2025 metams

<b>Pensijų fondą valdanti įmonė</b> <b>Laikotarpis</b>	<b>UAGDPB „AVIVA Lietuva“</b>	<b>UAB „Danske Capital investicijų valdymas“</b>	<b>UAB „DNB investicijų valdymas“</b>	<b>UAB „INVL Asset Management“</b>	<b>UAB „SEB investicijų valdymas“</b>	<b>UAB „Swedbank investicijų valdymas“</b>
<b>2011 m.</b>	16% (162 924)	2% (14 940)	8% (73 981)	4% (35 015)	27% (257 613)	44% (431 076)
<b>2015 m.</b>	17% (211 894)	2% (21 505)	11% (134 974)	9% (106 164)	22% (267 772)	39% (471 529)
<b>2025 m.</b>	24%	2%	17%	15%	17%	24%

Vadovaujantis šios prognozės rezultatais, pensijos kaupimas rekomenduojamas UAGDPB "AVIVA Lietuva" įmonei priklausančiuose PF, nes jų rinkos dalies prieaugis bus didžiausias (7 punktai). Kitos įmonės, kurių fondų rinkos dalis padidės yra: UAB " DNB investicijų valdymas " ir UAB " INVL Asset Management ", abiem įmonėms priklausanti rinkos dalis padidės po 6 punktus.

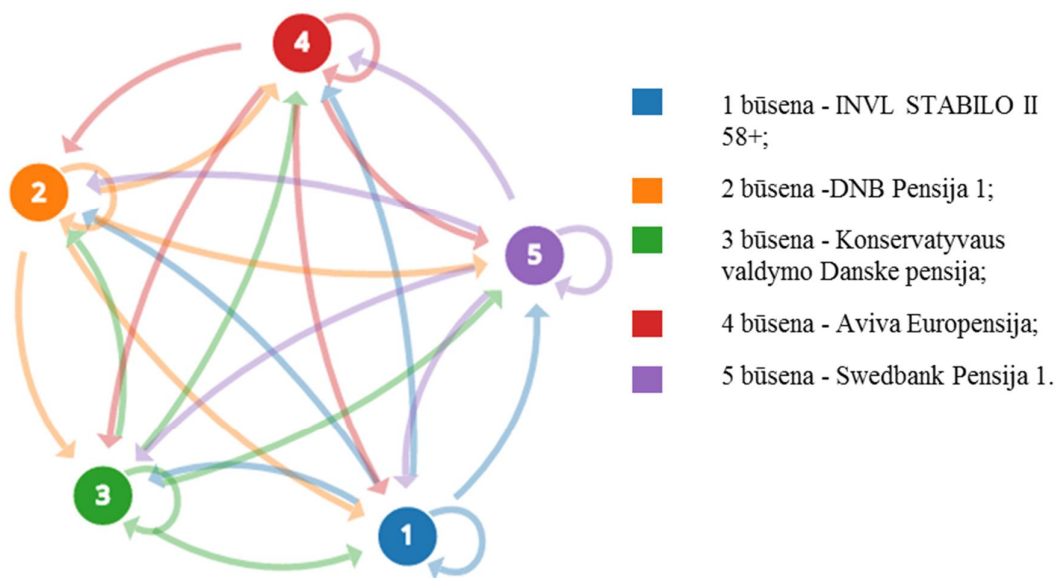
### **3.4.2. Konservatyvių pensijų fondų analizė**

Toliau analizuojamos Markovo grandinės pavaizduotos 3.3 – 3.6 paveikslėliuose. Modeliai skirti prognozuoti klientų kiekį PF. Kaip ir anksčiau prognozei naudojami 2011 m. ir 2012 m. duomenys, gauti rezultatai bus lyginami su 2015 m. duomenimis. Pirmiausiai nagrinėjama Markovo grandinė, kuri atspindi dalyvių judėjimą tarp PF, kurie vadovaujasi konservatyvia investavimo strategija.



**3.9 lentelė.** Rinkos dalies pasiskirstymo tarp konservatyvių pensijų fondų 2015 metų istorinių duomenų ir prognozės palyginimas

Pensijų fondas \ Metodas	2015 m. duomenys	Homogeninis Markovo modelis	Laike kintantis Markovo modelis	Išplėstas laike kintantis Markovo modelis
INVL STABILO II 58+	8%	5%	4%	3%
DNB Pensija 1	21%	12%	12%	15%
Konservatyvaus valdymo Danske pensija	1%	2%	3%	4%
Aviva Europensija	35%	39%	39%	46%
Swedbank Pensija 1	33%	42%	42%	32%



3.3 pav. Markovo grandinė atspindinti konservatyvių pensijų fondų dalyvių migraciją

Aukščiau esančioje lentelėje pateikti 3.3 pav. pavaizduotos Markovo grandinės įvertinimo rezultatai. Visų trijų metodų rezultatai panašūs. Didžiausia modelių įvertinių paklaidą – „DNB Pensija 1“ PF rinkos dalies įvertinime. Mažiausia paklaidą įvertinant šią reikšmę – 7 punktai. Mažiausia paklaidą – „Konservatyvaus valdymo Danske pensija“ PF rinkos dalies įvertinime, nes šio PF didžiausia paklaidą yra 3%.

Norint parinkti geriausią metodą įvertinamos MAPE ir MAE paklaidos. Įvertinimo rezultatai pateikti 3.10 lentelėje.

**3.10 lentelė.** Rinkos dalies pasiskirstymo tarp konservatyvių pensijų fondų prognozių paklaidų įverčiai

Metodas	Paklaidų įvertis	MAPE (vidutinė absoliutinė procentinė paklaida)	MAE (vidutinė absoliutinė paklaida)
Homogeninis Markovo modelis		36,3 %	5,3
Laike kintantis Markovo modelis		48,6 %	5,6
Išplėstas laike kintantis Markovo modelis		69,5 %	5,4

Iš šios lentelės duomenų galime teigti, kad šiuo atveju tiksliausias yra homogeninis Markovo modelis. Parinkus tiksliausią metodą, buvo atlikta prognozė 2025 m. Jos rezultatai pateikti 3.11 lentelėje.

**3.11 lentelė.** Rinkos dalies pasiskirstymas tarp konservatyvių pensijų fondų 2011 ir 2015 metais bei prognozė 2025 metams

Pensijų fondas	INVL STABILO II 58+	DNB Pensija 1	Konservatyvaus valdymo Danske pensija	Aviva Europensija	Swedbank Pensija 1
2011 m.	0,3 % (218)	10,5 % (6 535)	1 (644)	39,2 % (24 427)	49 % (30 472)
2015 m.	8 % (5 780)	21 % (14 935)	1 % (744)	36 % (24 749)	34 % (23 382)
2025 m.	7 %	22 %	2 %	35 %	34 %

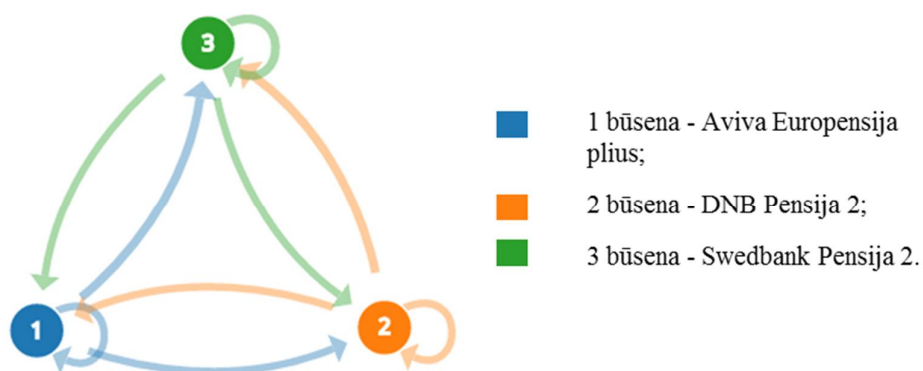
Pasak atliktos prognozės 2025 metais rinkos situacija beveik nepasikeis. Didžiausias užfiksuotas pokytis – rinkos dalis padidės / sumažės vienu punktu. Tai reiškia, kad visi fondai ganėtinai stabilūs. Norint išvengti rizikos, jog fonde neliks klientų, renkamės tarp „Aviva Europensija“ ir „Swedbank Pensija 1“ fondų, tačiau jų rinkos dalies skirtumas tik vienas punktas, todėl renkantis tarp šių fondų neužtenka dalyvių migracijos analizės. Norint parinkti geriausią fondą reikia atsižvelgti ir į grynųjų aktyvų pokytį pateiktą 3.2 lentelėje. Vadovaujantis šios lentelės duomenimis „Aviva Europensija“ PF grynųjų aktyvų augimas vienam PF dalyviui (atsižvelgus į infliaciją) yra 19 punktų didesnis už „Swedbank Pensija 1“ grynųjų aktyvų augimą. Todėl jis rekomenduojamas kaip mažiausiai rizikingas PF.

### 3.4.3. Mažos rizikos pensijų fondų analizė

Kitame žingsnyje tokia pati analizė taikoma Markovo grandinei, kuri skirta mažos akcijų dalies PF dalyvių analizei. Lentelėje 3.12 pateikiami 2015 m. rinkos dalies įvertinimo rezultatai.

**3.12 lentelė.** Rinkos dalies pasiskirstymo tarp mažos rizikos pensijų fondų 2015 metų istorinių duomenų ir prognozės palyginimas

Metodas Pensijų fondas	2015 m. duomenys	Homogeninis Markovo modelis	Laike kintantis Markovo modelis	Išplėstas laike kintantis Markovo modelis
Aviva Europensija plus	46,1 %	38,7 %	38,4 %	41,3 %
DNB Pensija 2	16 %	18 %	18 %	19,3 %
Swedbank Pensija 2	37,9 %	43,3 %	43,6 %	39,4 %



3.4 pav. Markovo grandinė atspindinti mažos rizikos pensijų fondų dalyvių migraciją

Homogeninis ir laike kintantis Markovo modeliai gauna panašius rinkos dalies įverčius. Jų įverčiai skiriasi 0-0,3 punktais, o tuo tarpu išplėsto laike kintančio Markovo modelio įverčiai nuo ankščiau minėtų skiriasi 1,3-4,2 punktais. Norint parinkti tinkamiausią variantą, įvertinamos MAPE ir MAE paklaidos, naudojant 2015 m. duomenis. Paklaidų įverčiai pateikti 3.13 lentelėje.

**3.13 lentelė.** Rinkos dalies pasiskirstymo tarp mažos rizikos pensijų fondų prognozių paklaidų įverčiai

Metodas	Paklaidų įvertis	MAPE (vidutinė absoliutinė procentinė paklaida)	MAE (vidutinė absoliutinė paklaida)
Homogeninis Markovo modelis		12,9 %	4,7
Laike kintantis Markovo modelis		13,2 %	4,8
Išplėstas laike kintantis Markovo modelis		10,1 %	2,9

Iš šioje lentelėje pateiktų duomenų matome, kad išplėsto laike kintančio Markovo modelio MAPE paklaida mažiausia. Pagal prognozę 2015 metais mažos rizikos PF dalyvaus 292

433 asmenys, o remiantis istoriniais duomenimis veikloje dalyvavo 288 209 asmenys, tai reiškia, kad modelio paklaida – 4 224 klientai, t.y. 1,5%. Taigi šis modelis yra tinkamiausias norint atlikti mažos rizikos PF prognozę 2025 metams. Modelio prognozė pateikta 3.14 lentelėje

**3.14 lentelė.** Rinkos dalies pasiskirstymas tarp mažos rizikos pensijų fondų 2011 ir 2015 metais bei prognozė 2025 metams

<b>Pensijų fondas</b> <b>Laikotarpis</b>	<b>Aviva Europensija plus</b>	<b>DNB Pensija 2</b>	<b>Swedbank Pensija 2</b>
<b>2011 m.</b>	38,4 % (100 251)	12,2 % (31 836)	49,4 % (128 952)
<b>2015 m.</b>	46,1 % ( 132 862)	16 % (46 219)	37,9 % (109 128)
<b>2025 m.</b>	45 %	16,7 %	38,3 %

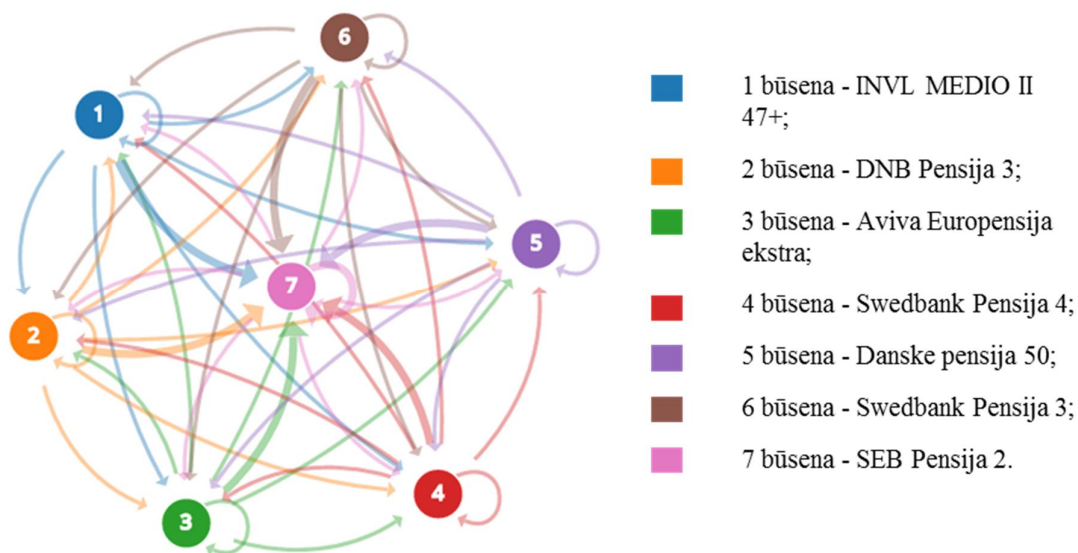
Modelis prognozuoja, jog 2025 metais PF rinkos dalies tarp mažos rizikos PF pasiskirstymas bus panašus į 2015 metų rinkos pasiskirstymą. PF „Aviva Europensija plus“ rinkos dalis sumažės 1,1 punktu, o PF „Swedbank Pensija 2“ rinkos dalis padidės 0,4 punkto, tačiau dominuojantis PF nepasikeis. Taigi rekomenduojama investuoti į PF „Aviva Europensija plus“.

#### **3.4.4. Vidutinės rizikos pensijų fondų analizė**

Toliau analizuojama Markovo grandinė, kuri skirta modeliuoti PF dalyvių judėjimą tarp vidutinės rizikos PF. Atliktos prognozės 2015 metams rezultatai pateikti 3.15 lentelėje.

**3.15 lentelė.** Rinkos dalies pasiskirstymo tarp vidutinės rizikos pensijų fondų 2015 metų istorinių duomenų ir prognozės palyginimas

<b>Pensijų fondas</b> <b>Metodas</b>	<b>2015 m. duomenys</b>	<b>Homogeninis Markovo modelis</b>	<b>Laike kintantis Markovo modelis</b>	<b>Išplėstas laike kintantis Markovo modelis</b>
<b>INVL MEDIO II 47+</b>	5,9 %	3,8 %	3,8 %	3 %
<b>DNB Pensija 3</b>	11,8 %	8,2 %	8,2 %	8,5 %
<b>Aviva Europensija ekstra</b>	8,7 %	8,4 %	8,3 %	7,9 %
<b>Swedbank Pensija 4</b>	17,5 %	17,6 %	17,7 %	17,5 %
<b>Danske pensija 50</b>	0,7 %	2,6 %	2,6 %	2,1 %
<b>Swedbank Pensija 3</b>	26,9 %	29,1 %	29,2 %	29 %
<b>SEB Pensija 2</b>	28,5 %	30,3 %	30,2 %	32 %



3.5 pav. Markovo grandinė atspindinti vidutinės rizikos pensijų fondų dalyvių migraciją

Iš lentelėje pateiktų duomenų matome, kad visi trys metodai pateikia panašias prognozes, kai naudojami 2011 ir 2012 metų duomenys, prognozuojant 2015 metų rinkos pasiskirstymą. Tačiau modelių ir realių duomenų įverčių skirtumas atskiriems PF svyruoja tarp 0,1 iki 4,9 punktų. Modelių paklaidos pateiktos 3.16 lentelėje.

**3.16 lentelė.** Rinkos dalies pasiskirstymo tarp vidutinės rizikos pensijų fondų prognozių paklaidų įverčiai

Metodas	Paklaidų įvertis	MAPE (vidutinė absoliutinė procentinė paklaida)	MAE (vidutinė absoliutinė paklaida)
Homogeninis Markovo modelis		36,2 %	1,9
Laike kintantis Markovo modelis		36,3 %	1,9
Išplėstas laike kintantis Markovo modelis		32,6 %	2,2

Iš šios lentelės duomenų matome, kad prognozei atlikti tinkamiausias yra išplėstasis laike kintantis Markovo modelis. Jo MAPE paklaida rinkos proporcijai – 32,6 %, tačiau klientų kiekio 2015 m. paklaida – 12,2%. Todėl 2025 m. prognozei atlikti bus naudojamas homogeninis Markovo modelis. Verta pastebėti, kad visų trijų modelių MAPE ir MAE paklaidos panašios, jų skirtumas svyruoja tarp 0 ir 0,3 punkto. 3.17 lentelėje pateikiama homogeninio Markovo modelio prognozė 2025 metams.

**3.17 lentelė.** Rinkos dalies pasiskirstymas tarp vidutinės rizikos pensijų fondų 2011 ir 2015 metais bei prognozė 2025 metams

<b>Pensijų fondas</b> <b>Laikotarpis</b>	<b>INVL MEDIO II 47+</b>	<b>DNB Pensija 3</b>	<b>Aviva Europensija ekstra</b>	<b>Swedbank Pensija 4</b>	<b>Danske pensija 50</b>	<b>Swedbank Pensija 3</b>	<b>SEB Pensija 2</b>
<b>2011 m.</b>	1,58 % (8 412)	6,7 % (35 610)	7,2 % (38 246)	18 % (95 806)	0,7 % (3 774)	32,7 % (173 984)	33,1 % (175 791)
<b>2015 m.</b>	5,9 % (37 356)	11,8 % (74 042)	8,7 % (54 883)	17,5 % (109 903)	0,7 % (4 449)	26,9 % (169 469)	28,5 % (179 739)
<b>2025 m.</b>	16 %	17 %	11,2 %	12,3 %	8,7 %	17,3 %	17,5 %

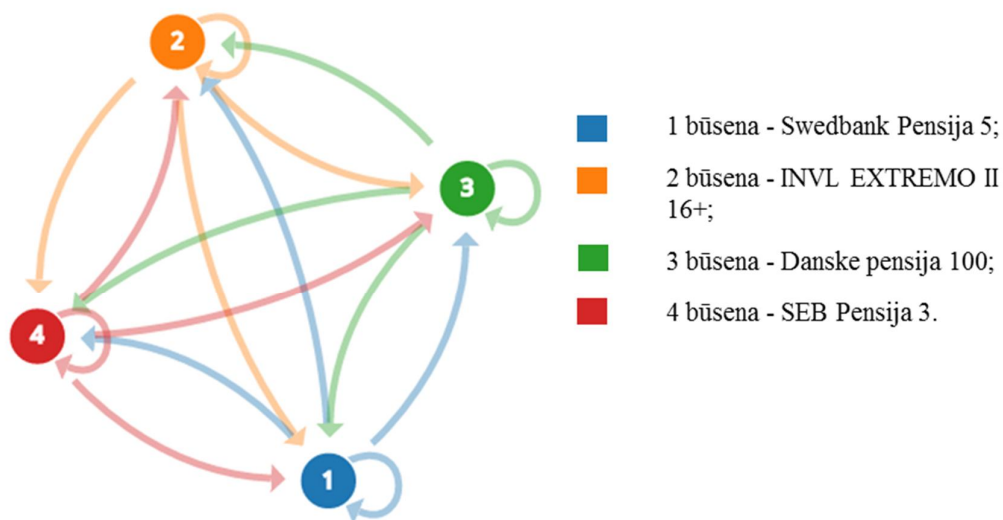
Pasak šios prognozės per ateinančius 10 metų vidutinės rizikos PF rinkos pasiskirstymas smarkiai pasikeis. Didžiausią rinkos dalį užimančiu fondu išliks “SEB Pensija 2”, tačiau jis praras 11 proc. Didžiausias augimas prognozuojamas PF “INVL MEDIO II 47+”. Jo rinkos dalis padidės 10,1 punktu ir beveik pasivys “SEB Pensija 2”. Dėl šios priežasties rekomenduojama investuoti į PF “INVL MEDIO II 47+”.

### 3.4.5. Didelės rizikos pensijų fondų analizė

Paskutinė Markovo grandinė, kuri analizuojama naudojant homogeninį, laike kintantį ir išplėstą laike kintantį Markovo modelius, modeliuoja didelės rizikos PF dalyvių judėjimą tarp PF. Šios grandinės 2015 metų prognozė, kai naudojami 2011 ir 2012 metų duomenys, pateikta 3.18 lentelėje.

**3.18 lentelė.** Rinkos dalies pasiskirstymo tarp didelės rizikos pensijų fondų 2015 metų istorinių duomenų ir prognozės palyginimas

<b>Pensijų fondas</b> <b>Metodas</b>	<b>2015 m. duomenys</b>	<b>Homogeninis Markovo modelis</b>	<b>Laike kintantis Markovo modelis</b>	<b>Išplėstas laike kintantis Markovo modelis</b>
<b>Swedbank Pensija 5</b>	31,6 %	11,1 %	10 %	33,4 %
<b>INVL EXTREMO II 16+</b>	31,7 %	31,8 %	32,2 %	26,9 %
<b>Danske pensija 100</b>	8,5 %	13 %	13,1 %	12,4 %
<b>SEB Pensija 3</b>	28,2 %	44,1 %	44,7 %	27,6 %



3.6 pav. Markovo grandinė atspindinti didelės rizikos pensijų fondų dalyvių migraciją

Iš lentelėje pateiktų duomenų matome, kad analizuojamų modelių paklaidos didelės. Didžiausias nuokrypis nuo realių duomenų – 21,6 punktai, o mažiausias – 0,1. Norint palyginti modelių tikslumą apskaičiuojamos MAPE ir MAE paklaidos. Jos pateiktos žemiau esančioje lentelėje.

**3.19 lentelė.** Rinkos dalies pasiskirstymo tarp didelės rizikos pensijų fondų prognozių paklaidų įverčiai

Metodas	Paklaidų įvertis	MAPE (vidutinė absoliutinė procentinė paklaida)	MAE (vidutinė absoliutinė paklaida)
Homogeninis Markovo modelis		42,2 %	10,2
Laike kintantis Markovo modelis		44,9 %	10,8
Išplėstas laike kintantis Markovo modelis		15 %	2,6

Įvertinus paklaidas, geriausias modelis – išplėstas laike kintantis Markovo modelis, jo MAPE įvertis - 15 %. Vadovaujantis prognozės rezultatais 2015 m. didžiausios rizikos PF dalyvaus 191 943 asmenys, o remiantis istoriniais duomenimis veikloje dalyvavo 168 494 asmenys, tai reiškia, kad modelio paklaida – 23 449 klientai, t.y. 12,2%. Dėl šios priežastis šis modelis netinkamas ir renkamas modelis su antra mažiausia MAPE paklaida. 2025 m. prognozei bus naudojamas homogeninis Markovo modelis. Prognozės rezultatai pateikti 3.20 lentelėje.

**3.20 lentelė.** Rinkos dalies pasiskirstymas tarp didelės rizikos pensijų fondų 2011 ir 2015 metais ir prognozė 2025 metams

<b>Pensijų Fondas</b> <b>Laikotarpis</b>	<b>Swedbank Pensija 5</b>	<b>INVL EXTREMO II 16+</b>	<b>Danske pensija 100</b>	<b>SEB Pensija 3</b>
<b>2011 m.</b>	2,3 % (1 862)	32,8 % (26 385)	13,1 % (10 522)	51,8 % (41 710)
<b>2015 m.</b>	31,6 % (60 616)	31,7 % (60 939)	8,5 % (16 309)	28,2 % (54 079)
<b>2025 m.</b>	30,2 %	31,5 %	9 %	29 %

Prognozuojamas 2025 m. rinkos pasiskirstymas panašus į 2015 m. To priežastis – stabilūs PF. Ateinančius dešimt metų didžiausią rinkos dalį užims „INVL EXTREMO II 16+“ PF. Verta pastebėti, kad 2011 – 2015 m. šio fondo grynujų aktyvų vertės pokytis vienam dalyviui (įvertinus infliaciją) buvo didžiausias tarp didžiausios rizikos PF. Dėl šių priežasčių norint investuoti į didelės rizikos fondus rekomenduojama pasirinkti būtent šį PF.

#### **3.4.6. Modelių trūkumai ir privalumai**

Naudotų modelių trūkumai:

- Homogeninis ir laike kintantis Markovo modeliai suteikia mažai informacijos apie prognozuojamus PF;
- Išplėstas laike kintantis Markovo modelis įvertina dalyvių kiekį PF prognozuojamu laikotarpiu, tačiau visais atvejais paklaida buvo apie 10%;
- Visų trijų modelių MAPE paklaidos didelės tam tikroms grandinėms, tačiau reikia atsižvelgti į tai, kad dalis naudojamų duomenų nėra realūs, o išskaičiuoti iš išvestinių.

Naudotų modelių pliusai:

- Modeliai pateikia informaciją apie rinkos pasiskirstymą ateityje;
- Galima labai lengvai pritaikyti didesnėms sistemoms, kaip pavyzdžiui visai pensijų sistemai;
- Net ir su netiksliais duomenimis modeliai pateikia rezultatus, kurie atspindi rinkos tendencijas.



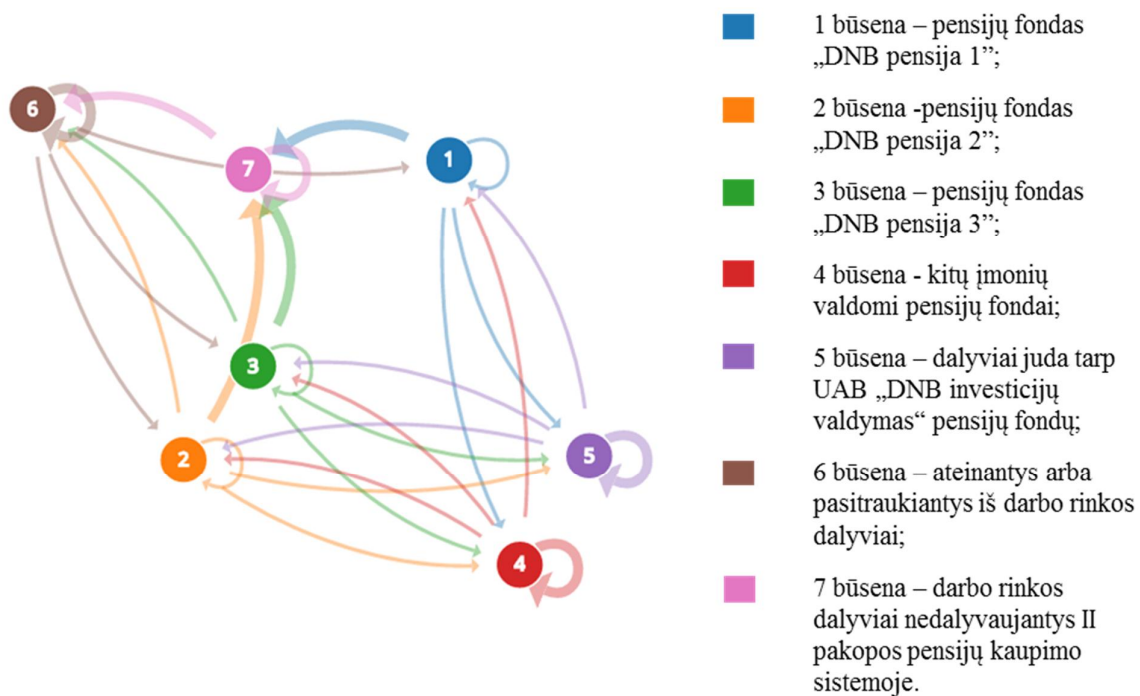
### 3.5. Markovo grandinės, sudarytos atsižvelgiant į turimus duomenis, analizė

Analizės metu buvo taikomi 2.3.3. poskyryje pateikti metodai ir analizuojami gauti rezultatai. Analizės metu naudojamas modelis pateiktas 3.7 pav. Šį modelį sudaro 7 skirtingos būsenos. Žemiau pateiktas išsamus būsenų aprašymas. Punkto numeris atitinka būsenos numerį Markovo grandinėje.

- 1) „DNB pensija 1” PF;
- 2) „DNB pensija 2” PF;
- 3) „DNB pensija 3” PF;
- 4) Būsena atspindinti kitų įmonių PF. Šios būsenos įtraukimo tikslas – turimi duomenys. Norint tyrime naudoti tik turimus duomenis ir gauti kuo realistiškesnius rezultatus ši būsena skirta atspindėti informacijai apie dalyvių srautą, kai nežinoma iš kokio/ į kokį PF perėjo dalyviai;
- 5) Būsena atspindinti dalyvių judėjimą tame pačiame fonde. Ši būsena įtraukta į modelį dėl tų pačių priežasčių, kaip ir būsena Nr. 4.;
- 6) Į šią būseną pereina iš PF pasitraukę dalyviai. Iš šios būsenos išeinantys asmenys simbolizuoja naujai į darbo rinką ateinančius asmenis. Tai daroma norint modelyje atspindėti mažėjančio gimstamumo problemą;
- 7) Į paskutinę būseną patenka asmenys pasinaudoję galimybe nutraukti II pakopos pensijos kaupimo sutartį.

Kadangi dalyviai iš PF gali išeiti ar į juos ateiti jiems tinkamu momentu, o ne iš anksto numatytais laiko momentais, šiai grandinei analizuoti labiausiai tinka TLMG, tačiau turimi duomenys yra diskretūs. Todėl tolimesniuose modelio etapuose bus taikomi metodai, kurie gali iš diskrečių duomenų įvertinti intensyvumą matricą  $Q$ .

Analizės metu naudojama UAB "DNB investicijų valdymas" įmonės suteikta informacija apie jų valdomų PF dalyvių migraciją.



3.7 pav. Markovo grandinė atspindinti įmonės UAB „DNB investicijų valdymas“ pensijų fondų dalyvių migraciją

### 3.5.1. Įstrižinės korekcijos metodas

Pirmiausiai, įvertinama  $Q$  matrica naudojant 2015 m. II pusmečio duomenis ir įstrižinės korekcijos metodą. Detalus metodo aprašymas pateiktas skyriuje 2.3.3.1. Tarkime, kad vienas periodas – šeši mėnesiai. Žemiau pateikta įvertinta  $Q$  matrica.

$$Q = \begin{pmatrix} -0,0038 & 0 & 0 & 0,00114 & 0,00007 & 0 & 0,00260 \\ 0 & -0,00092 & 0 & 0,00053 & 0,00007 & 0,00002 & 0,00031 \\ 0 & 0 & -0,00168 & 0,00142 & 0,00007 & 0,00001 & 0,00018 \\ 0,00001 & 0,00011 & 0,00011 & -0,00023 & 0 & 0 & 0 \\ 0,44183 & 0,20067 & 0,20067 & 0 & -0,84317 & 0 & 0 \\ 0,00005 & 0,00077 & 0,00196 & 0 & 0 & -0,00278 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0,40124 & -0,40124 \end{pmatrix}$$

Įvertinus intensyvumų matricą, galime įvertinti stacionarias šios Markovo grandinės perėjimo tikimybes, kai  $t \rightarrow \infty$ . Tikimybių įverčiai pateikti žemiau esančioje lentelėje.

3.21 lentelė. Stacionarių tikimybių įverčiai naudojant įstrižinės korekcijos metodą

Būseną, $i$	1	2	3	4	5	6	7
Tikimybė, $\pi_i$	0,00478	0,11452	0,07921	0,77786	0,00002	0,02346	0,00015

Iš aukščiau pateiktų rezultatų matome, kad ateityje įmonei UAB "DNB investicijų valdymas" priklausys 19,85 % rinkos. Šį rezultatą galime palyginti su 3.8 lentėje pateiktais prognozės įverčiais 2025 m., naudojant laike kintančio Markovo modelio prognozę. Šių dviejų prognozių skirtumas – 19,85 % - 17 % = 2,85 %.

### 3.5.2. Svertinės korekcijos metodas

Šiame tyrimo etape intensyvumų matrica įvertinama naudojant svertinės korekcijos metodą, detalus modelio aprašymas pateiktas skyriuje 2.3.3.1. Gauta  $Q$  matrica pateikta žemiau.

$$Q = \begin{pmatrix} -0,00381 & 0 & 0 & 0,00114 & 0,00007 & 0 & 0,00261 \\ 0 & -0,00092 & 0 & 0,00053 & 0,00007 & 0,00002 & 0,00031 \\ 0 & 0 & -0,00168 & 0,00142 & 0,00007 & 0,00001 & 0,00018 \\ 0,00001 & 0,00011 & 0,00011 & -0,00023 & 0 & 0 & 0 \\ 0,83476 & 0,37913 & 0,37913 & 0 & -1,59302 & 0 & 0 \\ 0,00005 & 0,00077 & 0,00196 & 0 & 0 & -0,00278 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0,50484 & -0,50484 \end{pmatrix}$$

Įvertinus intensyvumų matricą, galime įvertinti stacionarias šios Markovo grandinės perėjimo tikimybes, kai  $t \rightarrow \infty$ . Tikimybių įverčiai pateikti 3.22 lentelėje.

**3.22 lentelė.** Stacionarių tikimybių įverčiai naudojant svertinės korekcijos metodą

Būsena, $i$	1	2	3	4	5	6	7
Tikimybė, $\pi_i$	0,00478	0,11452	0,07921	0,77786	0,00002	0,02346	0,00015

Analizės rezultatai identiški įstrižainės korekcijos metodo analizės rezultatams.

### 3.5.3. EM algoritmas

Paskutiniame tyrimo etape intensyvumų matrica įvertinama naudojant EM algoritmą, detalus modelio aprašymas pateiktas skyriuje 2.3.3.2. Stabilūs rezultatai pasiekti algoritmą pakartojus 28 kartus. Skaičiavimai užtruko 18h. Gauta  $Q$  matrica pateikta žemiau.

$$Q = \begin{pmatrix} -0,00379 & 0 & 0 & 0,00114 & 0,00011 & 0 & 0,00260 \\ 0 & -0,00092 & 0 & 0,00053 & 0,0001 & 0,0002 & 0,00031 \\ 0 & 0 & -0,00168 & 0,00142 & 0,00011 & 0,00001 & 0,00018 \\ 0,00001 & 0,00011 & 0,00011 & -0,00023 & 0 & 0 & 0 \\ 0,44016 & 0,20067 & 0,20067 & 0 & -1,33253 & 0 & 0 \\ 0,00005 & 0,00077 & 0,00196 & 0 & 0 & -0,00278 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0,40126 & -0,40124 \end{pmatrix}$$

### 3.23 lentelė. Stacionarių tikimybių įverčiai naudojant EM algoritmą

Būsena, $i$	1	2	3	4	5	6	7
Tikimybė, $\pi_i$	0,0048	0,115	0,07954	0,78111	0,00002	0,02355	0,00015

Šio modelio prognozė panaši į matytas anksčiau. Tačiau verta pastebėti, kad dėl duomenų stygiaus dalis šio modelio privalumų nebuvo išnaudoti. Norint gauti geresnius rezultatus reikėtų naudoti kelių etapų vieno dalyvio duomenis.

Iš 3.23 lentėje pateiktų rezultatų matome, kad ateityje įmonei UAB "DNB investicijų valdymas" priklausys 19,93 % rinkos. Jeigu prisiminsime 3.8 lentelėje pateiktus duomenis galime teigti, kad PF valdanti bendrovė UAB "DNB investicijų valdymas" per ateinančius 10 metų augs ir todėl ją galima laikyti stabilia.

### 3.5.4. Modelių trūkumai ir privalumai

Naudotų modelių trūkumai:

- Įstrižinės ir svertinės korekcijos metodų tikslumas mažesnis už EM algoritmo [18];
- EM algoritmo skaičiavimai užtrunka žymiai ilgiau nei visų kitų šiame darbe analizuotų metodų.

Naudotų modelių plusai:

- Modeliai įvertina intensyvumų matrica, kuri turi daug panaudojimo būdų;
- Galima gauti rezultatus naudojant tik viešai prieinamus duomenis;
- EM algoritmas gali atsižvelgti į dalyvių migraciją praeityje, tačiau šiuos duomenis gauti sunku.

### 3.6. Programos vartotojo vadovas

Tyrimo metu buvo naudota statistinės ir grafinės analizės programa R. Visi skaičiavimai atlikti su 3.2.4. programos versija. Šiame skyriuje pateiktas atskirų funkcijų aprašymas. Modelių įkėlimui į aplinką sukurtas R paketas „PensijuFonduModeliai\_0.1.0.tar“. Žemiau esančioje lentelėje pateiktas modeliuose naudojamų parametrų paaiškinimas.

**3.24 lentelė.** Modelių realizacijose naudojami parametrai

Parametras	Parametro interpretacija
<b>M1 (=M2,M3)</b>	Matrica, kurioje elementas $\{i, j\}$ simbolizuoja dalyvių perėjusių iš būsenos $i$ į būseną $j$ skaičių
<b>T</b>	Prognozuojamų periodų skaičius. Pvz. jeigu naudojami 2011 m. duomenys ir analizė atliekama 2015 m., o vieno periodo trukmė metai, tai parametras $t = 4$
<b>Q</b>	Intensyvumų matrica. Jos detalus aprašymas pateiktas skyriuje 2.2.
<b>perejimu_matrica</b>	Matrica, aprašanti kiekvieno dalyvio judėjimą, t.y. pirmas dalyvis laiko momentu $t$ buvo būsenoje $i$ ir laiko momentu $t + 1$ perėjo į būseną $j$
<b>dalyviai1</b>	Fondo dalyvių skaičius laiko momentu $t - 1$
<b>dalyviai2</b>	Fondo dalyvių skaičius laiko momentu $t$
<b>isejo</b>	Iš fondo veiklos pasitraukusių dalyvių skaičius laiko momentu $t$

#### 1. Homogeninis Markovo modelis

Pilna šio kodo realizacija pateikta pirmame priede. Skaičiavimui naudojami dviejų etapų duomenys. Jie priskiriami matricoms  $M1$  ir  $M2$ . Parametras  $T$  keičiamas pagal prognozuojamų periodų kiekį, kur matrica  $M1$  sudaroma iš pradinio periodo duomenų. Norint įvertinti paklaidas, modeliui priskiriama matrica  $M3$ , kuri sudaroma iš istorinių prognozuojamo periodo duomenų. Matematinis modelio aprašymas pateiktas skyriuje 2.3.1.1. Naudojamų paklaidų aprašymas pateiktas skyriuje 2.5.

Analizei naudojamas R paketas „expm“. Paketas skirtas atlikti matricos kėlimo laipsniu operaciją.

#### 2. Laike kintantis Markovo modelis

Kodo realizacija pateikta pirmame priede. Skaičiavimuose naudojami homogeniniam Markovo modelyje aprašyti parametrai. Matematinis modelio aprašymas pateiktas skyriuje 2.3.1.2. Naudojamų paklaidų aprašymas pateiktas skyriuje 2.5.

#### 3. Išplėstas laike kintantis Markovo modelis

Kodo realizacija pateikta pirmame priede. Analizei naudojamos aukščiau pateiktos matricos  $M1$ ,  $M2$  ir  $M3$ . Matematinis modelio aprašymas pateiktas skyriuje 2.3.1.3. Naudojamų paklaidų aprašymas pateiktas skyriuje 2.5.

#### 4. Įstrižainės korekcijos metodas

Modelio tikslas – intensyvumų matricos  $Q$  įvertinimas iš diskrečių istorinių duomenų. Kodo realizacija pateikta pirmame priede. Matricai įvertinti naudojama matrica  $M1$ . Matematinis modelio aprašymas pateiktas skyriuje 2.3.3.1.

#### 5. Svertinės korekcijos metodas

Kodo realizacija pateikta pirmame priede. Modelis naudoja tuos pačius duomenis, kaip ir įstrižainės korekcijos metodas. Matematinis modelio aprašymas pateiktas skyriuje 2.3.3.1.

#### 6. EM algoritmas

Kodo realizacija pateikta prieduose. Algoritmas, naudodamas duomenis, pateiktus matricoje perejimų lentelė ir pradinį  $Q$  matricos reikšmių spėjimą (siūloma naudoti įstrižainės korekcijos ar svertinės korekcijos metodų  $Q$  matricos įverčius), įvertina matricą  $Q$ . Esant galimybei rekomenduojama naudoti daugiau negu vieno periodo perėjimo matricą. Matematinis modelio aprašymas pateiktas skyriuje 2.3.3.2.

Analizei naudojamas R paketas „doParallel“. Paketas skirtas kodo optimizacijai.

#### 7. Stacionarių tikimybių įvertinimas

Kodo realizacija pateikta prieduose. Stacionariosios  $Q$  matricos tikimybės įvertinamos, imituojant vieno dalyvio judėjimą tarp būsenų. Tikimybių matematinis aprašymas pateiktas skyriuje 2.4.

#### 8. Trūkstančių duomenų atkūrimas

Kodo realizacija pateikta prieduose. Naudodamas vektorius *dalyviai1*, *dalyviai2* ir *isejo*. Algoritmas apskaičiuoja fondų dalyvių santykį laiko momentais  $t$  ir  $(t - 1)$  ir pagal tai išskirsto bendrą pasitraukusių dalyvių kiekį atskiriems fondams.

## IŠVADOS

Darbo eigoje buvo pristatyti Lietuvos Respublikoje egzistuojantys II pakopos pensijų fondai ir atlikta jų grynujų aktyvų bei dalyvių pokyčių analizė. Pateiktos skirtingos pensijų fondų investavimo strategijos ir išskirtos rekomenduotinos amžiaus grupės.

Sukurtas ir 1 priede pateiktas programos realizacijos kodas. Kiekvienam metodui realizuota atskira funkcija ir visos funkcijos pateiktos R pakete. Programos sukurtos naudojant statistinės analizės programą R ir jos paketus „doParallel“ bei „expm“.

Analizės metu naudojami šeši skirtingi modeliai. Pirmi trys modeliai pritaikyti grandinėms su išskaičiuotais duomenimis. Kiekvienai grandinei parinktas geriausias metodas atsižvelgiant į modelio paklaidų įverčius, bei identifikuoti naudotų modelių trūkumai ir privalumai. Analizės rezultatai pateikti, kaip investavimo rekomendacijos:

- Rekomenduojama PF valdanti įmonė - UAGDPB "AVIVA Lietuva";
- Rekomenduojamas konservatyvus PF - „Aviva Europensija“;
- Rekomenduojamas mažos rizikos (akcijos sudaro iki 30 proc. portfelio investicijų) PF – „Aviva Europensija plus“;
- Rekomenduojamas vidutinės rizikos (akcijos sudaro 30-70 proc. portfelio investicijų) PF – “INVL MEDIO II 47+”;
- Rekomenduojamas didžiausios rizikos (akcijos sudaro 70-100 proc. portfelio investicijų) PF - „INVL EXTREMO II 16+“.

Likę trys analizės metodai pritaikyti įmonei UAB "DNB investicijų valdymas". Grandinės įvertinime naudojami tik realūs duomenys. Tyrimo rezultatai ir įžvalgos:

- Analizės metu įvertintas įmonės UAB "DNB investicijų valdymas" stabilumas. Pasak visų trijų naudotų prognozės modelių ši įmonė augs ir ateityje jai priklausys didesnė rinkos dalis, todėl ši pensijų fondus valdanti įmonė yra laikoma stabilia dalyvių augimo atžvilgiu;
- Prognozės tendencijos atitinka anksčiau naudotų metodų rezultatus.

Gauti rezultatai atitiko grynujų aktyvų ir dalyvių pokyčio analizės rezultatus. Identifikuoti pensijų fondai priklausantys įmonėms, kurių grynujų aktyvų pokytis buvo didžiausias arba stabiliausias.

## LITERATŪROS SĄRAŠAS

1. Lietuvos Respublikos Centrinis Bankas. *Pensijų Fondai*. URL:[https://www.lb.lt/pensiju\\_fondai](https://www.lb.lt/pensiju_fondai).
2. Pinigų Bitė. URL:<http://www.pinigubite.lt/planuojant-pensija/antros-pakopos-pensiju-fondai>.
3. BERTSCHI, L., SVEN E., ANDREAS, R. Dynamic Asset Liability Management: A Profit Testing Model for Swiss Pension Funds. 2015
4. MACCIONI, A.F. A Stochastic Model for the Analysis of Demographic Risk in Pay-as-You-Go Pension Funds. *Mathematical Methods in Economics and Finance*, 2011, vol. 3, pp 41-60.
5. VISCO, I. Retirement Saving and the Payout Phase. *OECD Journal: Financial Market Trends*, 2009, vol. 1. pp. 143-162.
6. PRANEVIČIUS, H. ir VALAKEVIČIUS, E. *Markovo Procesų Teorijos Taikymas Sistemoms Modeliuoti: Paskaitų Konspektas*. Kaunas: Technologija, 1991.
7. BUCEA-MANEA-ȚONIȘ, R. Computer Assisted Si Computer Assisted Simulation of Market Shares for Competing Products using Markov Chains. INT. Symposium of Electrical Engineering and Energy Converters - ELS 2003
8. CHAN, K.C. Markov and Non-Markov Models for Brand Switching and Market Share Forecasting. Sydney, Australia: 2014 07 1-4, 2014 ISBN 978-0-9803267-5-8.
9. CHAN, K.C. Market Share Modelling and Forecasting using Markov Chains and Alternative Models. *International Journal of Innovative Computing, Information and Control*, 2015, vol. 11, no. 4.
10. GÉZÁNÉ, H. Prediction of Customers' Loyalty and Market Share by using Markov Chain Model. *Professional Papers*, 2007, vol. 9. pp. 91.
11. DATONG, G. A Markov Chain Model Analysis of GSM Network Service Providers Marketing Mix. *International Journal of Engineering and Sciences*, 2011, vol. 11, no. 4. pp. 49-57.
12. ISRAEL, R.B., ROSENTHAL, J.S. and WEI, J.Z. Finding Generators for Markov Chains Via Empirical Transition Matrices, with Applications to Credit Ratings. *Mathematical Finance*, 2001, vol. 11, no. 2. pp. 245-265.
13. KREININ, A. and SIDELNIKOVA, M. Regularization Algorithms for Transition Matrices. *Algo Research Quarterly*, 2001, vol. 4, no. 1/2. pp. 23-40.
14. BLADT, M. and SØRENSEN, M. Statistical Inference for Discretely Observed Markov Jump Processes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2005, vol. 67, no. 3. pp. 395-410.
15. ALBERT, A. Estimating the Infinitesimal Generator of a Continuous Time, Finite State Markov Process. *The Annals of Mathematical Statistics*, 1962. pp. 727-753.
16. VAIČIULIS, M. *Pensijų Fondai*, 2008.
17. AB "Swedbank". *Kokį Pensijų Fondą Pasirinkti?* URL:[https://www.swedbank.lt/lt/pages/privatiems/swedbank\\_pensiju\\_fondai](https://www.swedbank.lt/lt/pages/privatiems/swedbank_pensiju_fondai).
18. INAMURA, Y. *Estimating Continuous Time Transition Matrices from Discretely Observed Data*, 2006.



## 1 priedas. Funkcijos skirtos modelių realizacijai

### 1. Homogeninis Markovo modelis, „Homogeninis.R“

```
homogeninis<-function(M1, M2, T, M3=NULL, tikri=NULL)
{
#Naudojama paketu ikelimas i R aplinka

if(!require("expm")){install.packages("expm")}
library(expm)

# #Reikalingu duomenų ikelimas
# M1 #Dalyvių migracija laiko momentu T=-1, pvz 2011 m.
# M2 #Dalyvių migracija laiko momentu T=0, pvz 2012 m.
# T=9 #Prognoze atliekama periodui T+1
#
# M3 #Dalyvių migracija laiko momentu T+1, pvz. 2015m.
# tikri #Imonių / pensijų fondų rinkos dalis laiko momentu 2015m.

#####Modelio realizacija #####
if (is.null(M3)) {remove(M3)}
if (is.null(tikri)) {remove(tikri)}

#Nereikalingu duomenų pasalinimas iš programos aplinkos
rm(list=setdiff(ls(), c("M1", "M2", "M3","tikri", "T")))

#Prognozuojamo periodo rinkos dalis (ivertinama jeigu pateikta matrica M3)
if (exists("M3"))
{
realus<-M3
mshare=matrix(0, nrow=1,ncol=ncol(realus))
for (i in 1:nrow(realus))
{
mshare[1,i]=sum(realus[,i])
}
tikri=round(mshare/sum(mshare),4)
}

#T=0 periodo rinkos dalies ivertinimas
mshare2=matrix(0, nrow=1,ncol=ncol(M2))
for (i in 1:nrow(M2))
{
mshare2[1,i]=sum(M2[,i])
}

#Periodo T=-1 tikimybių pereinimo matricos ivertinimas
P1<-M1
for (i in 1:nrow(M1))
{
P1[i,]=M1[i,]/sum(M1[i,])
}
P1[is.na(P1)]<-0

#Periodo T=0 tikimybių pereinimo matricos ivertinimas
P2<-M2
for (i in 1:nrow(M2))
{
P2[i,]=M2[i,]/sum(M2[i,])
}
P2[is.na(P2)]<-0

#Pereinimo tikimybių matricų vidurkinimas
P=round((P1+P2)/2,5)

#Prognoze
prognoze=round((mshare2/sum(mshare2))%*%(P%^(T-1)),4)

#Jeigu pateikta matrica M3 arba vektorius tikri apskaičiuojamos modelio
paklaidos
```

```

if (exists("tikri"))
{
#MAPE (1/n)*sum((A-F)/A)
MAPE=sum(abs((tikri-prognoze)/tikri))/length(tikri)

#MAE (1/n)*sum(|F-A|)
MAE=sum(abs(prognoze-tikri))/length(tikri)
}

rm(list=setdiff(ls(), c("M1", "M2", "prognoze", "realus", "tikri", "MAPE",
"MAE")))

####Prognozes rezultatu spausdinimas####
if (exists("tikri"))
{
print("prognoze")
print(round(prognoze*100,2))
print("MAPE")
print(MAPE)
print("MAE")
print(MAE)

rezultatai <- list("prognoze" = round(prognoze*100,2), "MAPE" = MAPE,
"MAE"=MAE)
return(rezultatai)

} else {
print("prognoze")
print(round(prognoze*100,2))

rezultatai <- list("prognoze" = round(prognoze*100,2),"MAPE" = "truksta
duomenu", "MAE"="truksta duomenu")
return(rezultatai)
}
}

```

## 2. Laike kintantis Markovo modelis, „Laike\_kintantis.R“

```

laike_kintantis<-function(M1, M2, T, M3=NULL, tikri=NULL)
{
#Reikalingu duomenu ikelimas
# M1 #Dalyviu migracija laiko momentu T=-1, pvz 2011 m.
# M2 #Dalyviu migracija laiko momentu T=0, pvz 2012 m.
# T=10 #Prognoze atliekama periodui T+1
#
# M3 #Dalyviu migracija laiko momentu T+1, pvz. 2015m.
# tikri #Imoni / pensiju fondu rinkos dalis laiko momentu 2015m.

#####Modelio realizacija #####
if (is.null(M3)) {remove(M3)}
if (is.null(tikri)) {remove(tikri)}
#Nereikalingu duomenu pasalinimas is programos aplinkos
rm(list=setdiff(ls(), c("M1", "M2", "M3", "tikri", "T")))

#Prognozuojamo periodo rinkos dalis (ivertinama jeigu pateikta matrica M3)
if (exists("M3"))
{
realus<-M3
mshare=matrix(0, nrow=1,ncol=ncol(realus))
for (i in 1:nrow(realus))
{
mshare[1,i]=sum(realus[,i])
}
tikri=round(mshare/sum(mshare),4)
}
}

```

```

#T=0 periodo rinkos dalies ivertinimas
mshare2=matrix(0, nrow=1,ncol=ncol(M2))
for (i in 1:nrow(M2))
{
  mshare2[1,i]=sum(M2[,i])
}

#Periodo T=-1 tikimybiu perejimo matricos ivertinimas
P1<-M1
for (i in 1:nrow(M1))
{
  P1[i,]=M1[i,]/sum(M1[i,])
}
P1[is.na(P1)]<-0

#Periodo T=0 tikimybiu perejimo matricos ivertinimas
P2<-M2
for (i in 1:nrow(M2))
{
  P2[i,]=M2[i,]/sum(M2[i,])
}
P2[is.na(P2)]<-0

#Tikimybiu pokycio ivertinimas
P_delta=P2-P1

#Prognoze
a=diag(nrow(M1))
for (i in 1:(T-1))
{
  P_t=P1+(i-1)*P_delta
  for (x in 1:nrow(P_t))
  {
    for (j in 1:ncol(P_t))
    {
      if(P_t[x,j]<0)
      {
        P_t[x,j]=0
      }
    }
    P_t[x,]=P_t[x,]/sum(P_t[x,])
  }
  a=a%%P_t
}
prognoze=round((mshare2/sum(mshare2))%%a,4)

#Jeigu pateikta matrica M3 arba vektorius tikri apskaičiuojamos modelio
paklaidos
if (exists("tikri"))
{
  #MAPE (1/n)*sum((A-F)/A)
  MAPE=sum(abs((tikri-prognoze)/tikri))/length(tikri)

  #MAE (1/n)*sum(|F-A|)
  MAE=sum(abs(prognoze-tikri))/length(tikri)
}

rm(list=setdiff(ls(), c("M1", "M2", "prognoze", "realus", "tikri", "MAPE",
"MAE")))

####Duomenu spausdinimas####
if (exists("tikri"))
{
  print("prognoze")
  print(round(prognoze*100,2))
  print("MAPE")
  print(MAPE)
  print("MAE")
  print(MAE)
  rezultatai <- list("prognoze" = round(prognoze*100,2), "MAPE" = MAPE,
"MAE"=MAE)
}

```

```

return(rezultatai)

} else {
  print("prognoze")
  print(round(prognoze*100,2))

  rezultatai <- list("prognoze" = round(prognoze*100,2),"MAPE" = "truksta
duomenu", "MAE"="truksta duomenu")
  return(rezultatai)
}
}

```

### 3. Išplėstas laike kintantis Markovo modelis, „Išplestas\_laike\_kintantis.R“

```

isplestas_laike_kintantis<-function(M1, M2, T, M3=NULL, tikri=NULL)
{
# #Reikalingu duomenu ikelimas
# M1 #Dalyviu migracija laiko momentu T=-1, pvz 2011 m.
# M2 #Dalyviu migracija laiko momentu T=0, pvz 2012 m.
# T=10 #Prognozuojamu periodu kiekis
#
# M3 #Dalyviu migracija laiko momentu T+1, pvz. 2015m.a
# tikri #Imonių / pensiju fondu rinkos dalis laiko momentu 2015m.

#####Modelio realizacija #####
#Nereikalingu duomenu pasalinimas is programos aplinkos
if (is.null(M3)) {remove(M3)}
if (is.null(tikri)) {remove(tikri)}

rm(list=setdiff(ls(), c("M1", "M2", "M3", "T", "tikri")))

#Prognozuojamo periodo rinkos dalis (ivertinama jeigu pateikta matrica M3)
if (exists("M3"))
{
realus<-M3

mshare=matrix(0, nrow=1,ncol=ncol(realus))
for (i in 1:nrow(realus))
{
mshare[1,i]=sum(realus[,i])
}
tikri=round(mshare/sum(mshare),4)
}

#T=0 periodo rinkos dalies ivertinimas
mshare2=matrix(0, nrow=1,ncol=ncol(M2))
for (i in 1:nrow(M2))
{
mshare2[1,i]=sum(M2[,i])
}

#Periodo T=-1 tikimybiu perejimo matricos ivertinimas
P1<-M1
for (i in 1:nrow(M1))
{
P1[i,]=M1[i,]/sum(M1[i,])
}
P1[is.na(P1)]<-0

#Periodo T=0 tikimybiu perejimo matricos ivertinimas
P2<-M2
for (i in 1:nrow(M2))
{
P2[i,]=M2[i,]/sum(M2[i,])
}
P2[is.na(P2)]<-0

```

```

#Tikimybiu pokycio ivertinimas
P_delta=P2-P1

#Rinkos augimo ivertinimas
G=0
for (i in 1:nrow(M1))
{
G[i]=sum(M2[i,])/sum(M1[,i])
}

#Rinkos dalies
Q1=0
for (i in 1:nrow(M1))
{
Q1[i]=sum(M2[,i])
}

#Prognoze
Q_t=G^(T-1)*Q1
for (i in 1:(T-1))
{
P_t=P1+(i-1)*P_delta
for (x in 1:nrow(P_t))
{
for (j in 1:ncol(P_t))
{
if(P_t[x,j]<0)
{
P_t[x,j]=0
}
}
P_t[x,]=P_t[x,]/sum(P_t[x,])
}
Q_t=Q_t%%P_t
}
prognoze=round(Q_t/sum(Q_t),4)

#Jeigu pateikta matrica M3 arba vektorius tikri apskaičiuojamos modelio
paklaidos
if (exists("tikri"))
{
#MAPE (1/n)*sum((A-F)/A)
MAPE=sum(abs((tikri-prognoze)/tikri))/length(tikri)

#MAE (1/n)*sum(|F-A|)
MAE=sum(abs(prognoze-tikri))/length(tikri)
}

rm(list=setdiff(ls(), c("M1", "M2", "prognoze", "realus", "tikri", "MAPE",
"MAE","Q_t")))

####Duomenu spausdinimas####
if (exists("tikri"))
{
print("prognoze")
print(prognoze)
print ("Dalyviu pasiskirstymas")
print(Q_t)
print("Dalyviu kiekis")
print(sum(Q_t))
print("MAPE")
print(MAPE)
print("MAE")
print(MAE)

rezultatai <- list("prognoze" = prognoze, "Dalyviu
pasiskirstymas"=Q_t,"Dalyviu kiekis"= sum(Q_t), "MAPE" = MAPE, "MAE"=MAE)
return(rezultatai)

} else {
print("prognoze")

```

```

print(prognose)

rezultatai <- list("prognoze" = prognoze, "Dalyviu
pasiskirstymas"=Q_t,"Dalyviu kiekis"= sum(Q_t), "MAPE" = "truksta
duomenu", "MAE"="truksta duomenu")
return(rezultatai)
}
}

```

#### 4. Įstrižinės korekcijos metodas, „Istrizaines\_korekcija.R“

```

Istrizaines_korekcija<-function(M1)
{
#Reikalingu duomenu ikelimas
# M1 #Periodo T=0 dalyviu migracijos duomenys

#####Modelio realizacija #####
#Perejimo tikimybiu matricos ivertinimas
P<-M1
for (i in 1:nrow(M1))
{
P[i,]=M1[i,]/sum(M1[i,])
}
P[is.na(P)]<-0

#log(P) ivertinimas
I<-diag(nrow(M1))
P_log=0
for (i in 1:10^6)
{
a=((P-I)^(i))/i
P_log=P_log+a*(-1)^(i+1)
}

#Neigiamu elementu pasalinimas
Q_IK<-P_log
for (i in 1:nrow(P_log))
{
for (j in 1:ncol(P_log))
{
if((i!=j)&&(P_log[i,j]<0))
{
Q_IK[i,j]=0
}
}
}
}
m=P_log-diag(P_log)*diag(nrow(P_log))

#Istrizaines elementu perskaiciavimas
for (i in 1:nrow(m))
{
sum=0
suma=sum(m[i,])
Q_IK[i,i]=-suma
}

rm(list=setdiff(ls(), c("Q_IK","Q_WA", "Q_EM", "M1", "P")))

####Duomenu spausdinimas####
print("Q matrica")
print(round(Q_IK,6))

rezultatai <- list("prognoze" = round(Q_IK,6))
return(rezultatai)
}

```

## 5. Svertinės korekcijos metodas, „Svertine\_korekcija.R”

```
Svertine_korekcija<-function(M1)
{
#Reikalingu duomenų įkėlimas
# M1 #Periodo T=0 dalyvių migracijos duomenys

#####Modelio realizacija #####
#Perejimo tikimybių matricos įvertinimas
P<-M1
for (i in 1:nrow(M1))
{
  P[i,]=M1[i,]/sum(M1[i,])
}
P[is.na(P)]<-0

#log(P) įvertinimas
I<-diag(nrow(M1))
P_log=0
for (i in 1:10^6)
{
  a=((P-I)^(i))/i
  P_log=P_log+a*(-1)^(i+1)
}

#Neigiamų elementų pašalinimas
Q_SK<-P_log
for (i in 1:nrow(P_log))
{
  for (j in 1:ncol(P_log))
  {
    if((i!=j)&&(P_log[i,j]<0))
    {
      Q_SK[i,j]=0
    }
  }
}

#Svertinę elementų korekciją
Q_SK<-P_log
for (i in 1:nrow(P_log))
{
  for (j in 1:ncol(P_log))
  {
    Q_SK[i,j]=P_log[i,j]-
abs(P_log[i,j])*sum(P_log[i,])/sum(abs(P_log[i,]))
  }
}

rm(list=setdiff(ls(), c("Q_DA","Q_SK", "Q_EM", "M1", "P", "tikri")))

####Duomenų spausdinimas####
print("Q matrica")
print(round(Q_SK,6))
rezultatai <- list("prognoze" = round(Q_SK,6))
return(rezultatai)
}
```

## 6. EM algoritmas, „EM\_algoritmas.R“

```
EM_algoritmas<-function(Q, M1, kartojuimai)
{
#Naudojamu paketu ikelimas i R aplinka
#Pradinis paketo irasymas
if(!require("doParallel")){install.packages("doParallel")}
library(doParallel)

#####Reikalingu duomenu ikelimas#####
#Pradinis Q matricos spejimas, naudojamas vienas is zemiau pateiktu metodu
#Q<-Q_SK
#Q<-Q_IK
# Q<-matrix(c(...), nrow=nrow(M1), ncol=ncol(M1), byrow=TRUE)

#Periodo T=0 dalyviu migracijos duomenys
# M1

#Nurodoma kiek kartu reikia kartoti EM algoritma
# kartojuimai=52

#####Modelio realizacija #####
#Procesoriaus branduoliu kiekis paraleliniam skaiciavimui,
#jeigu pocesorius turi 4 branduolius skaiciavimui siuloma paskirti 3
c1 <- makeCluster(2)
registerDoParallel(c1)

#Perejimu matricos generavimas
if (!exists("perejimu_lentele"))
{
#perejimu duomenys lenteleje
perejimu_lentele<-matrix(nrow=sum(M1), ncol=2)
colnames(perejimu_lentele) <- c("t_n", "t_n+1")

s=0
foreach (x=1:nrow(M1)) %do%
{
for (y in 1:ncol(M1))
{
if (M1[x,y]>0)
{
for (i in 1:M1[x,y])
{
perejimu_lentele[i+s,1]=x
perejimu_lentele[i+s,2]=y
}
s=s+i
}
}
}
remove(i,s,x,y)
}

####EM algoritmas
bandymas=c(rep(0, kartojuimai))

foreach (z=1:kartojuimai) %do%
{
#####N11###
E_N_ij=matrix(0,nrow=nrow(M1),ncol=ncol(M1))

m=0
m=foreach(i=1:nrow(M1))%do%
{
m=foreach(j=1:nrow(M1))%do%
{
for (h in 1:nrow(perejimu_lentele))
{
```



```

if (h>1)
{
  if((perejimu_lentele[h-
1,1]==perejimu_lentele[h,1])&&(perejimu_lentele[h-
1,2]==perejimu_lentele[h,2]))
  {
    E_N_ij[i,j]=E_N_ij[i,j]+a*(M1[perejimu_lentele[h-
1,1],perejimu_lentele[h-1,2]]-1)
    h=h+M1[perejimu_lentele[h-1,1],perejimu_lentele[h-1,2]]-2
  }
}
else
{
  e_tn<-rep(0,nrow(M1)); e_tn[perejimu_lentele[h,1]]=1
  e_tn1<-rep(0,nrow(M1)); e_tn1[perejimu_lentele[h,2]]=1
  e_i<-rep(0,nrow(M1)); e_i[i]=1
  e_j<-rep(0,nrow(M1)); e_j[j]=1
  #analitine integracija
  tn1=2
  tn=1
  res=0
  f<-function(s){exp((s-tn)*Q)**e_i**(t(e_j))**exp((tn1-s)*Q)}
  for (ii in seq(1,2,1/10000)[-10001])
    res=res+f(ii)
  res=res/10000

  #D reiksme
  D=t(e_tn)**exp(Q)**(e_tn1)

  #laiko momento suma
  aa=t(e_tn)**res**e_tn1
  a=(1/D)*aa[1,1]*(Q[i,j])
  E_N_ij[i,j]=E_N_ij[i,j]+a
}
}
}

# Rii ###
R_N_j=c(rep(0, nrow(M1)))

m=foreach (j=1:nrow(M1))%do%
{
  for (h in 1:nrow(perejimu_lentele))
  {
    e_tn<-rep(0,nrow(M1)); e_tn[perejimu_lentele[h,1]]=1
    e_tn1<-rep(0,nrow(M1)); e_tn1[perejimu_lentele[h,2]]=1
    e_i<-rep(0,nrow(M1)); e_i[j]=1

    if (h>1)
    {
      if((perejimu_lentele[h-
1,1]==perejimu_lentele[h,1])&&(perejimu_lentele[h-
1,2]==perejimu_lentele[h,2]))
      {
        R_N_j[j]=R_N_j[j]+b*(M1[perejimu_lentele[h-
1,1],perejimu_lentele[h-1,2]]-1)
        h=h+M1[perejimu_lentele[h-1,1],perejimu_lentele[h-1,2]]-2
      }
    }
  }
}
else
{
  #analitine integracija
  tn1=2
  tn=1
  res=0
  f<-function(s){exp((s-tn)*Q)**e_i**(t(e_j))**exp((tn1-s)*Q)}
  for (ii in seq(1,2,1/10000)[-10001])

```

```

    for (ii in seq(1,2,1/10000)[-10001])
      res=res+f(ii)
    res=res/10000

    #D reiksme
    D=t(e_tn)%*%exp(Q)%*%(e_tn1)

    #laiko momento suma
    bb=t(e_tn)%*%res%*%e_tn1
    b=(1/D)*bb[1,1]
    R_N_j[j]=R_N_j[j]+b
  }
}

Q_nauja<-matrix(0,nrow=nrow(M1), ncol=ncol(M1))
for (i in 1:nrow(M1))
{
  Q_nauja[i,]<-E_N_ij[i,]/R_N_j[i]
}
Q=Q_nauja
bandymas[z]=sum(Q)
}

Q_EM<-Q
rm(list=setdiff(ls(), c("Q_DA","Q_WA", "Q_EM", "M1", "bandymas")))

#####Rezultatu spausdinimas#####
print("Q matrica")
print(round(Q_EM,6))

rezultatai <- list("prognoze" = round(Q_EM,6))
return(rezultatai)
}

```

## 7. Stacionarių tikimybių įvertinimas, „Stacionarus\_issidestymas.R“

```

Stacionarus_Issidestymas<-function(Q)
{
  #Naudojamu paketu ikelimas i R aplinka
  if(!require("expm")){install.packages("expm")}
  library(expm)

  # #Analizuojama intensyvumu matrica
  # Q<-...

  #####Stacionarus issidestymo ivertinimas####
  a=Q*1/6+diag(nrow(Q))
  b=c(1, rep(0,nrow(Q)-1))%*%a

  for (i in 1:10^7)
  {
    b=b%*%a
  }
  round(b,5)

  #####Rezultatai####
  print("stacionari")
  print(round(b,5))
  print("DNB dalis")
  print(sum(b[1:3]))

  rezultatai <- list("prognoze" = round(b,5),"DNB dalis" = sum(b[1:3]))
  return(rezultatai)
}

```

## 8. Trūkstatų duomenų atkūrimas, „Duomenų\_atkurimas.R“

```
Duomenų_atkurimas<-function(dalyviai1, dalyviai2, isejo)
{
  #####Atkuriamos grandines duomenys#####
  # #Laiko periodo T=-1 dalyviu skaicius
  # dalyviai1=c( 1, 21027, 8639, 39163)
  # #Laiko periodo T=0 dalyviu skaicius
  # dalyviai2=c(1862, 26385, 10522, 41710)
  # #Laiko period T=0 pasitraukusiu dalyviu skaicius
  # isejo=c(18, 278, 224, 1081)

  #####Duomenų atkurimas#####
  matrica=diag(dalyviai2)

  #Augimo koeficianto radimas
  delta=dalyviai2/dalyviai1

  #Isejusių dalyvių išskirstymas fondams (imonems)
  for (j in 1:nrow(matrica))
  {
    for (i in 1:ncol(matrica))
    {
      if (i!=j)
      {
        matrica[j,i]=isejo[j]*delta[i]/(sum(delta)-delta[j])
      }
    }
  }
  matrica<-round(matrica,0)
  matrica[is.na(matrica)]<-0
  M1<-matrica

  #####Rezultatu spausdinimas#####
  cat("\014")
  print("Atkurti duomenys")
  print(M1)
  return(M1)
}
```

## 9. Sukurto R paketo aprašymas

```
Package: PensijuFonduModeliai
Type: Package
Title: Baigiamojo projekto dalis
Version: 0.1.0
Author: Lina Kadikinaite
Description: Pensiju fondu analize naudojant Markovo grandines
LazyData: TRUE
```