VILNIAUS UNIVERSITETAS

MATEMATIKOS IR INFORMATIKOS FAKULTETAS

Magistro baigiamasis darbas

Glodinimo metodų palyginimas

Comparison of graduation methods

Aistė Dargvilaitė

VILNIUS 2020

MATEMATIKOS IR INFORMATIKOS FAKULTETAS
MATEMATIKOS INSTITUTAS

Darbo vadovas lekt. dr. Aldona Skučaitė _______________
Darbo recenzentas ____

Darbas apgintas įrašoma data ________________________
Darbas įvertintas ______________________________________

Registravimo NR. ____________________________________
Įrašoma atidavimo į katedrą data _________________________

Turinys

Santrauka / Abstract

Glodinimo metodų palyginimas

Santrauka

Sudarant mirtingumo lenteles pirmuoju žingsniu gauti mirtingumo parametrų įverčiai vadinami grubiais, nes juos jungianti kreivė gali nebūtinai būti glodi, todėl gautus įverčius reikia suglodinti. Šiam tikslui naudojami parametriniai ir neparametriniai metodai. Šio darbo metu apžvelgsime keletą pavyzdžių iš kiekvienos grupės - Gompertzo ir Makehamo modelius (parametrinis glodinimas) bei Whittaker ir slenkamojo vidurkio metodus (neparametrinis glodinimas). Pateiksime skaitinius pavyzdžius, remiantis Lietuvos gyventojų duomenimis, bei palyginsime gautus rezultatus.

Raktiniai žodžiai : parametrinis ir neparametrinis glodinimas, Gompertzo modelis, Makehamo modelis, Whittaker metodas, slenkamojo vidurkio metodas, grubūs ir suglodinti įverčiai, išgyvenamumo modeliai.

Comparison of graduation methods

Abstract

When constructing mortality table, we get mortality estimates which we call crude because the curve connecting produced estimates may not be smooth. So crude rates should be graduated. We will use parametric and non-parametric approaches. We will provide examples from each group - Gompertz model, Makeham model (parametric graduation) and Whittaker method, moving-weight average method (non-parametric graduation). Numerical examples applied to Lithuanian mortality are given and we will compare got results.

Key words : parametric and non-parametric graduation, Gompertz model, Makeham model, Whittaker method, moving-weight average method, crude and graduated estimates, smooth, survival models.

1 Įvadas

Glodinimo problema iškyla, konstruojant mirtingumo lenteles, kai gauti mirtingumo parametrų įverčiai yra nepakankamai glodūs. Tyrėjai dirba su pradiniais stebėjimo duomenimis ir pirmu žingsniu gauna grubius įverčius. Jie atsiranda dėl mirčių skaičiaus, stebimų asmenų skaičiaus ir pan., tiek dėl atsitiktinių svyravimų, tiek ir dėl kitų priežasčių. Tokie duomenys yra suglodinami, kad būtų išvengta nepaaiškinamų pokyčių šuoliukų ar pan. Glodinant duomenis svarbu atsižvelgti į du veiksnius - kaip tiksliai glodinti duomenys atspindi realius stebėtus duomenis ir tai, kokio glodumo yra glodintus duomenis jungianti kreivė [4]. Kuo kreivė mažiau skiriasi nuo stebėtų duomenų, tuo ji mažiau glodi ir atvirkščiai.

Glodinimo procesą pirmą kartą 1918 metais aprašė amerikiečių matematikas Robertas Hendersonas. Mortonas D. Mileris 1946 metais taip pat pateikė savo idėjas apie glodinimą [4]. Šis procesas susideda bent jau iš dviejų žingsnių. Pirmas, apskaičiojami grūbūs tam tikro parametro duomenys. Antras, šie įverčiai glodinami.

Kompiuteriai suteikė galimybę apdoroti didelį duomenų kiekį ir atlikti sudėtingus skaičiavimus, o tai palengvina darbą, kai norima atlikti skaičiavimus su realiais duomenimis, naudojant glodinimo metodus.

Grubūs mirtingumo duomenys gali būti glodinami, naudojant tiek parametrinius, tiek ir neparametrinius glodinimo metodus. Todėl šio darbo struktūra susideda iš trijų dalių. Pirmajame skyriuje trumpai aptarsime išgyvenamumo modelius. Antrame - nagrinėsime parametrinį glodinimą, pateiksime keletą konkrečių modelių, kurie bus pritaikyti Lietuvos gyventojų duomenims glodinti. Paskutiniame - apžvelgsime neparametrinį glodinimą, pateiksime keletą metodų ir praktinių skaičiavimų rezultatus. Ir galiausiai pateiksime viso darbo išvadas.

2 Išgyvenamumo teorijos pagrindai

Išgyvenamumo modeliu vadiname tam tikro atsitiktinio dydžio pasiskirstymo funkciją. Jei T - "mirties" momentas, tai išgyvenamumo funkcija žymima:

$$S(t) = P(T > t)$$

Čia $P(T > t)$ tikimybė, kad individas gyvas praėjus t laiko nuo stebėjimo pradžios. Momentu $t = 0$, tai išgyvenamumo funkcija lygi vienam, t.y. $S(0) = 1$. Taip pat, [10]

$$\lim_{t \to \infty} (S(t)) = 0.$$

Apžvelgsime keletą žymėjimų, kurie naudojami mirtingumo lentelėje [9]:

1. q_x - sąlyginė tikimybė, kad x-mečiui mirti per artimiausius metus su sąlyga, kad asmuo sulaukė amžiaus x.

2. d_x - mirčių skaičius amžiaus intervale [x, x+1].

3. E_x - ekspozicija, kuri gali būti matuojama stebimų asmenų skaičiumi arba stebimos kohortos bendra pragyventų metų trukme intervale [x,x+1].

4. m_x - centrinis mirtingumo dažnis, kuris apibrėžiamas kaip sąlyginio tankio μ_x svertinis vidurkis svoriais imant tikimybes išgyventi iki amžiaus y intervale [x,x+1]. Atsitiktinis dydis X žymi asmens amžių mirties dieną. Sąlyginis tankis dar vadinamas mirtingumo galia ir žymimas:

   $$\mu_x = \frac{f(x)}{S(x)}$$

   Čia $f(t)$ - atsitiktinio dydžio T tankis. Centrinis mirtingumo dažnis žymimas:

   $$m_x = \frac{\int_x^{x+1} S(y)\mu_y dy}{\int_x^{x+1} S(y) dy}.$$

Išgyvenamumo modeliai taikomi gyvybės draudime, pensijų sistemose bei kituose populiacijų tyrimuose [10].

3 Parametrinis glodinimas

Parametrinio glodinimo atveju, pasirenkama tam tikra matematinė formulė su vienu ar keliais parametrais, pavyzdžiui, Makehamo dėsnis, Gompertzo modelis, Heligmano-Pollardo dėsnis ar pan., ir tada pasirinkto modelio parametrai yra įvertinami iš turimų duomenų [8]. Glodinti mirtingumo duomenys turėtų parodyti laipsnišką gautų įverčių pasikeitimą be staigių pokyčių ar šuolių. Parametrinis glodinimas pritraukė aktuarų, statistikų, demografų dėmesį per praėjusį šimtmetį [1].

Glodinti duomenys išreiškiami per matematines funkcijas, priklausančias nuo amžiaus x. Pradinius įverčius žymėsime u_x, o glodintas reikšmes - simboliu v_x [4].

Toliau apžvelgsime keletą mirtingumo dėsnių, kurie plačiai naudojamos glodinimo procese. Tiems patiems grubiems įverčiams gali tikti keli mirtingumo dėsniai, tačiau reiktų pasirinkti labiausiai tinkantį modelį, kuris geriausiai suglodina grubius įverčius.

3.1 Gompertzo modelis

Gompertzas 1825 metais [4] užrašė modelį tokia formule

$$\mu_x = ac^x \tag{1}$$

Čia $a > 0$, $c > 1$. Šis modelis buvo pavadintas jo vardu. Tada Gompertzo modelio (1) logaritminė transformacija būtų:

$$\log \mu_x = \log a + x \log c$$

Gompertzo mirtingumo galia μ_x gali būti išreiškiama [6] taip:

$$\mu_x = ae^{bx} \tag{2}$$

Čia a žymi mirtingumo parametrą, kuris nepriklauso nuo amžiaus x, $a > 0$, o b - mirtingumo parametras priklausantis nuo amžiaus x, $b > 0$. $x = 0$ nurodo pradinį amžių. Gompertzo modelis buvo plačiai taikomas daugumoje šalių per paskutinius 170 metų.

Gompertzo modelis plačiai naudojamas ne tik sudarant mirtingumo lenteles, o ir biologijoje - augalų, gyvūnų populiacijoms augimui stebėti [10]. Gompertzo modelį pradėjo taikyti draudimo kompanijos, sudarant mirtingumo lenteles.

3.2 Makehamo modelis

Makehamo modelis yra Gompertzo modelio modifikacija. Šis modelis buvo pasiūlytas 1860 metais ir išreiškiamas tokia formule [4]:

$$\mu_x = a + be^{cx} \tag{3}$$

Čia $a > 0$ skirtas modeliuoti mirtingumą jaunesniame amžiuje, pavyzdžiui, dėl smurtinių mirčių. O be^{cx} - senėjmo komponentė, kai $b, c > 0$. Taip pat Makehamo modelį galima užrašyti taip [10]:

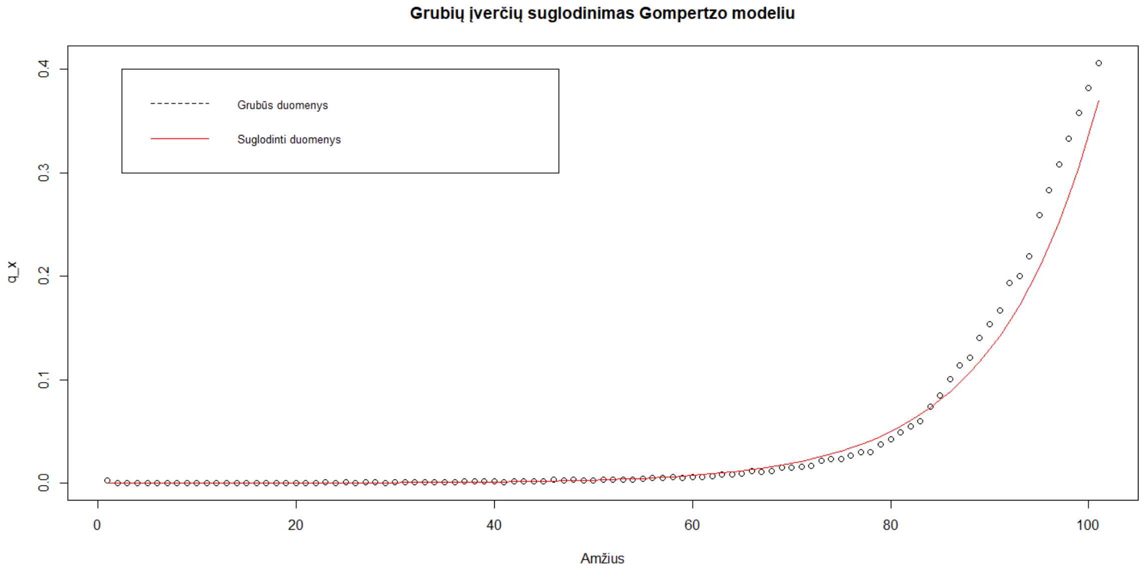
$$\mu_x = a + bd^x$$

Čia $a, b > 0$, $d > 1$ ir $d = e^c$.

3.3 Skaičiavimų rezultatai

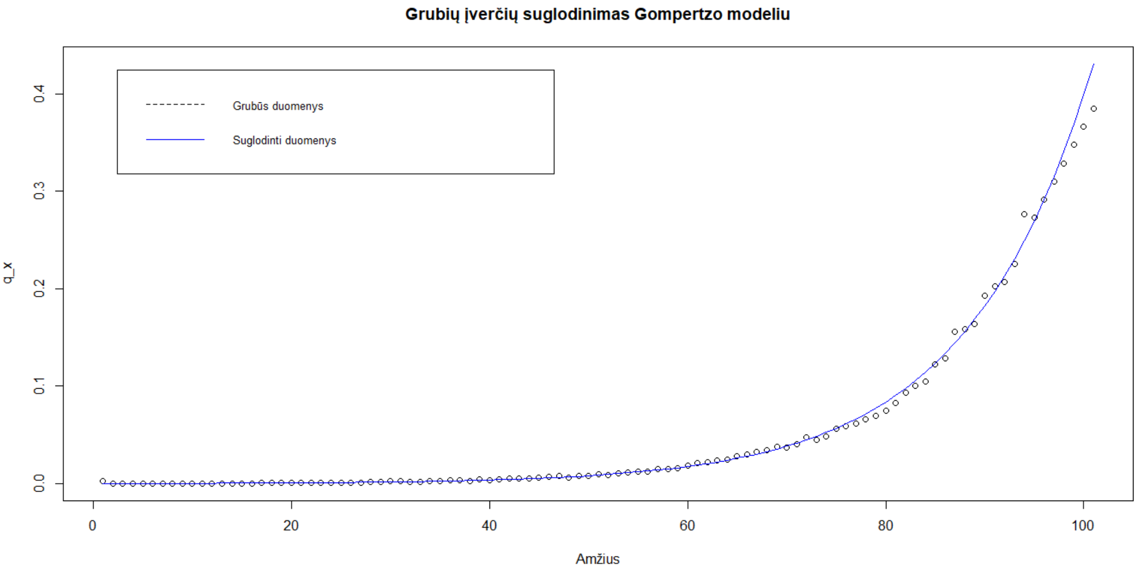
Gompertzo ir Makehamo modelius panaudosime duomenų apie 2017 metų Lietuvos gyventojų mirtingumą glodinimui [7]. Nagrinėjamas amžiaus intervalas nuo 0 iki 100 metų, tiek moterų, tiek vyrų duomenys bus nagrinėjami atskirai. Skaičiavimai bus atliekami naudojant paketą R.

Gompertzo modeliui naudosime (2) formulę. Tiek moterų, tiek vyrų atveju bus glodinami mirtingumo tikimybių įverčiai. Grafike (1 pav.) pavaizduoti moterų mirtingumo grubūs ir suglodinti įverčiai. Šiuo atveju Gompertzo modelio parametrai yra $a = 0.000024$, o $b = 0.0955$. Moterų atveju gavosi, kad suglodintos reikšmės vyresniuose amžiuose yra žemiau grubių reikšmių, galime manyti, kad šis modelis netinka moterų mirtingumo duomenims glodinti.



1 pav.: Mirtingumo tikimybių (moterys) pradiniai ir suglodinti įverčiai

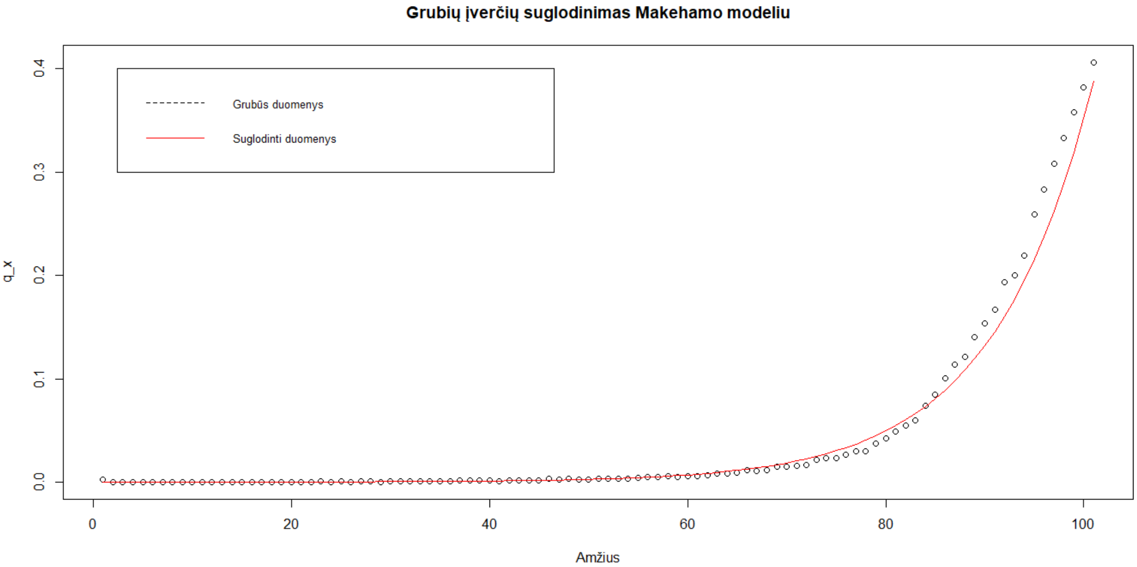
Toliau kitame grafike (2 pav.) pavaizduoti vyrų pradiniai įverčiai bei suglodinti duomenys.



2 pav.: Mirtingumo tikimybių (vyrai) pradiniai ir suglodinti įverčiai

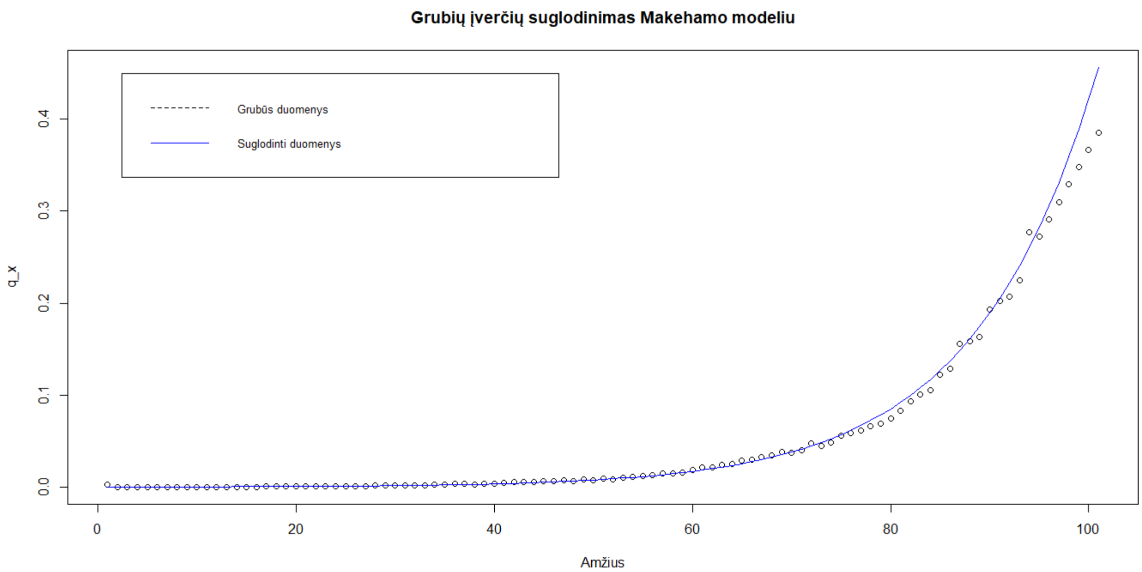
Vyrų atveju, parametrai - $a = 0.000163$, o $b = 0.77999$. Šie du grafikai, kur naudojamas Gompertzo modelis, nėra panašūs, nors naudojama ta pati struktūra. Bet yra žinoma, kad vyrai gyvena trumpiau ir mirčių skaičius yra didesnis.

Grafike (3 pav.) matome moterų duomenis, kur pirminiai įverčiai yra suglodinami Makehamo modeliu (3). Šiam būdui naudojami tie patys pradiniai įverčiai, t.y. mirtingumo tikimybės, kaip ir Gompertzo modeliui. Makehamo modelyje naudojami tokie parametrai: $a = 0.000019$, $b = 0.0981$, o $c = 0.000115$. Šiame grafike taip pat suglodintos reikšmės žemiau už pradinius įverčius.



3 pav.: Mirtingumo tikimybių (moterys) pradiniai ir suglodinti įverčiai

Toliau grafike (4 pav.) pavaizduoti vyrų duomenys. Vyrų atveju, $a = 0.00014$, $b = 0.08009$, o $c = 0.000067$.



4 pav.: Mirtingumo tikimybių (vyrai) pradiniai ir suglodinti įverčiai

Abiejų modelių atveju, jaunesniuose amžiuose mirtingumas didėja palaispniui ir suglodintos reikšmės, lyginant su grubiais įverčiais, yra gana panašios. Moterų suglodinti duomenys labiau skiriasi nuo pradinių nei vyrų atveju. Makehamo modelio atveju, moterų suglodinti įverčiai labiau panašesni į pradinius nei Gopertzo modelyje. Vyrų mirtingumo suglodinimas yra panašus abiejuose grafikuose.

Žemiau pateikta (1 lentelė, 2 lentelė) asmens likusio gyvenimo trukmė (atskirai moterims ir vyrams). Pagal Gompertzo ir Makehamo modelius moters likusi gyvenimo trukmė lygi 18 metų, o pagal grubius įverčius beveik 19 metų. Vyrų atveju, naudojant pradiniams įverčiams Gompertzo modelį likusi gyvenimo trukmė yra panašesnė į neglodintų duomenų.

	Neglodinti duomenys	Gompertzo modelis	Makehamo modelis
e_{65}	18.8077	18.1305	18.1622

1 lentelė: 65 metų moters likusio gyvenimo trukmė grubių įverčių atveju ir taikant modelius.

	Neglodinti duomenys	Gompertzo modelis	Makehamo modelis
e_{65}	13.8748	13.7215	13.6686

2 lentelė: 65 metų vyro likusio gyvenimo trukmė grubių įverčių atveju ir taikant modelius.

4 Neparametrinis glodinimas

Šioje dalyje kalbėsime apie neparametrinį glodinimą, kuriam atlikti nereikia prielaidos apie mirtingumo dėsnio matematinę formulę. Neparametriniams priskiriami tokie metodai kaip slenkamojo vidurkio (angl. weighted moving average), kernelio, Whittaker modelis bei kiti metodai, kurie paremti tam tikromis splaino funkcijomis (angl. spline functions) [8].

4.1 Whittaker metodas

E. T. Whittaker 1923 metais aprašė glodinimo metodą, kuris buvo pavadintas jo vardu. Šis metodas siekia balanso tarp gautos funkcijos glodumo ir atitikimo realiems duomenims.

Whittaker būdas yra paremtas tam tikros funkcijos minimizavimu. Pagrindinė šio metodo formulė ([4], [5]):

$$M = F + hS = \sum_{x=1}^{n} w_x(v_x - u_x)^2 + h\sum_{x=1}^{n-z}(\Delta^z v_x)^2 \tag{4}$$

Čia v_x - suglodinti duomenys, o u_x - grubūs įverčiai, kai $x = 1, 2, ..., n$. Keli pastebėjimai (4) formulėje:

1. Amžių žymėsime x, laikysime, kad $x = 1, 2, ..., n$.

2. Parametras z žymi polinomo laipsnį, kuris parodo duomenų glotnumą. Paprastai z imamas 2, 3 arba 4.

3. Svoriai $w_1, w_2, ..., w_n$ yra teigiami skaičiai ir lygūs 1.

4. Parametras h ($h > 0$) turi įtakos funkcijos M minimizavimui. Jei $h = 0$, tai $hS = 0$, bet kokiai S reikšmei. Kai M yra minimizuojame taške $F = 0$, t.y. $v_x = u_x$ visiems x, tada graduoti duomenys sutampa su grubiais įverčiais. Kuo aukštesnės h reikšmės, tuo labiau suglodinami duomenys.

4.2 Slenkamojo vidurkio metodas

Slenkamojo vidurkio (ang. moving-weighted average) metodas priklauso neparametriniam glodinimui ir jis yra vienas iš anksčiausiai pradėtų naudoti. Šis metodas pakankamai lengvas, nes galima išsiversti be kompiuterinių skaičiavimų [4].

Prieš nagrinėjant slenkamojo vidurkio metodą, reiktų aptarti įverčio paklaidas (angl. error of estimation) arba kitaip liekanas (angl. residuals). Iš mūsų turimos lentelės apie žymėjimus (žr. A.1 priede), turime:

$$u_x = t_x + e_x$$

Čia u_x - pradiniai įverčiai, t_x - "tikroji" reikšmė, e_x - paklaidos. Jei laikysime, jog t_x mūsų nagrinėjami duomenys u_x, tai jie nusakomi tokia slenkamojo vidurkio formule:

$$v_x = \frac{1}{2n+1}\sum_{r=-n}^{n} u_{x+r} \tag{5}$$

Čia

1. n neneigiamas sveikas skaičius.

2. v_x - empirinis $2n+1$ dydžių u_x vidurkis, kitaip suglodintos vertinamo parametro reikšmės.

3. u_{x+r} - grubios reikšmės.

Pasinaudoję (5) lygybę ([3], [4]), turime

$$v_x = \frac{1}{2n+1}\sum_{r=-n}^{n}(t_{x+r}+e_{x+r}) = \frac{1}{2n+1}\sum_{r=-n}^{n}t_{x+r} + \frac{1}{2n+1}\sum_{r=-n}^{n}e_{x+r} \tag{6}$$

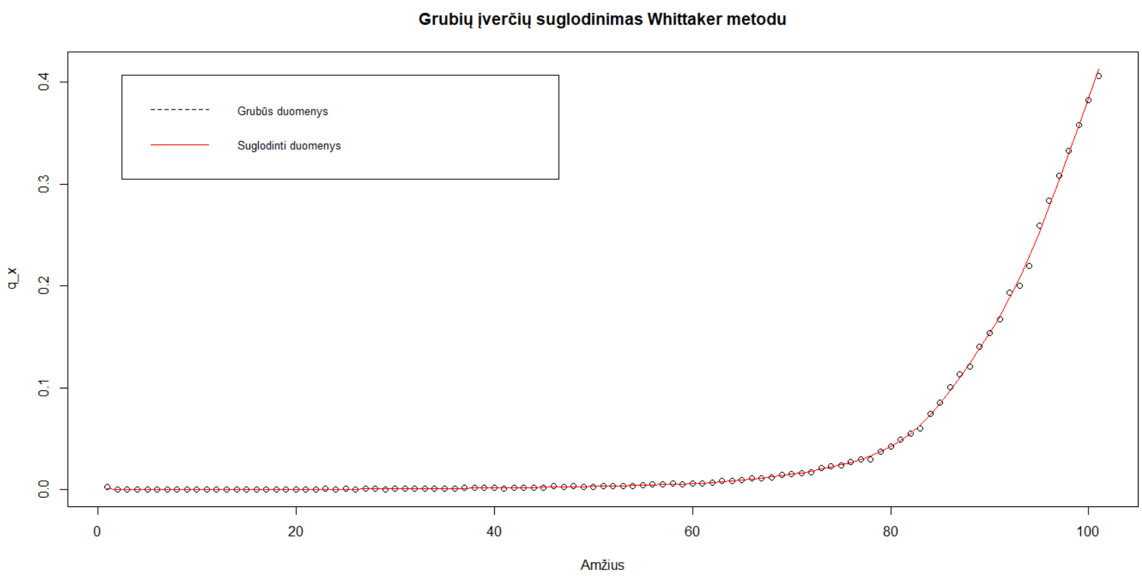
Lygybėje (6) liekanos paklaida

$$e'_x = \frac{1}{2n+1}\sum_{r=-n}^{n}e_{x+r}$$

kuri yra gali būti mažesnė arba didesnė nei e_x, o e'_x žymime liekanos paklaidą.

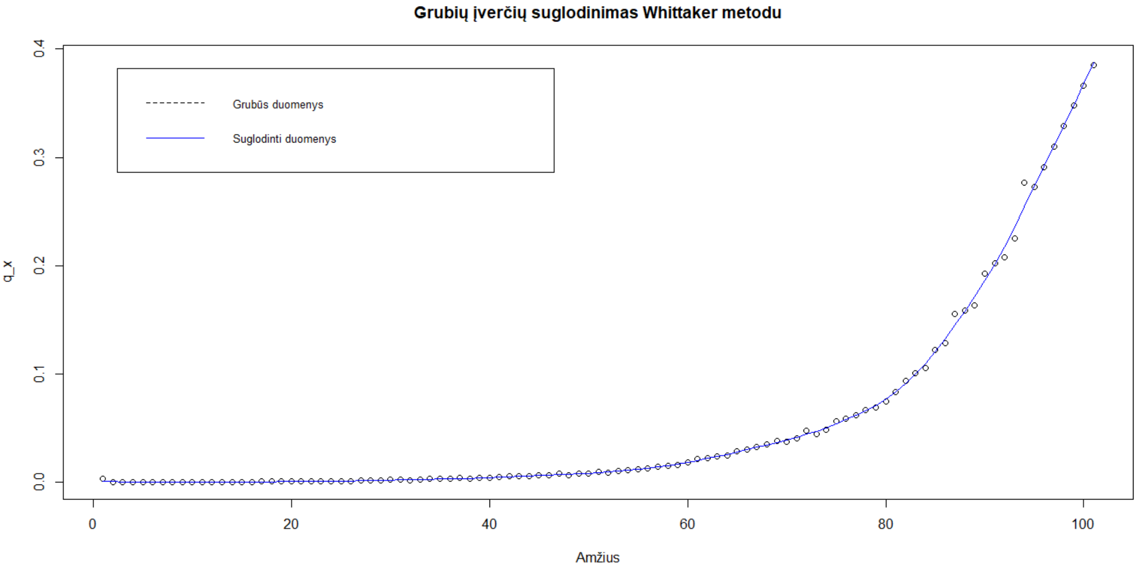
4.3 Skaičiavimų rezultatai

Neparametriniam glodinimui taip pat bus naudojami Lietuvos duomenys. Iš pradžių apžvelgsime Whittaker modelį. Glodinimui naudosime (4) formulę. Tiek vyrų, tiek moterų atveju laikysime, kad $h = 10$, o $z = 2$, nes keičiant parametrus, nesimato žymių pasikeitimų. Taip yra todėl, nes maža populiacija, miršta nedaug žmonių, lyginant su mažiau išsivysčiusiomis ar didesnę žmonių populiaciją turinčiomis šalimis. Modeliui naudosime pirminius mirtingumo tikimybių įverčius. Žemiau pateikti grubūs moterų mirtingumo tikimybių įverčiai ir jų suglodintos reikšmės (5 pav.).



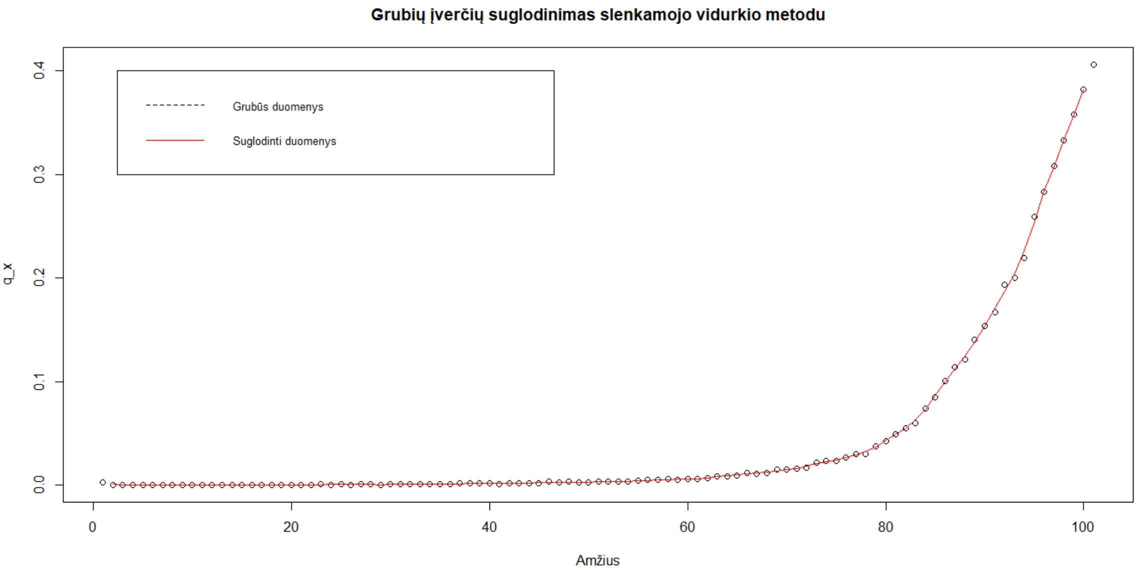
5 pav.: Moterų grubūs ir suglodinti įverčiai, h = 10

Toliau glodinami vyrų tikimybių įverčiai (6 pav.).

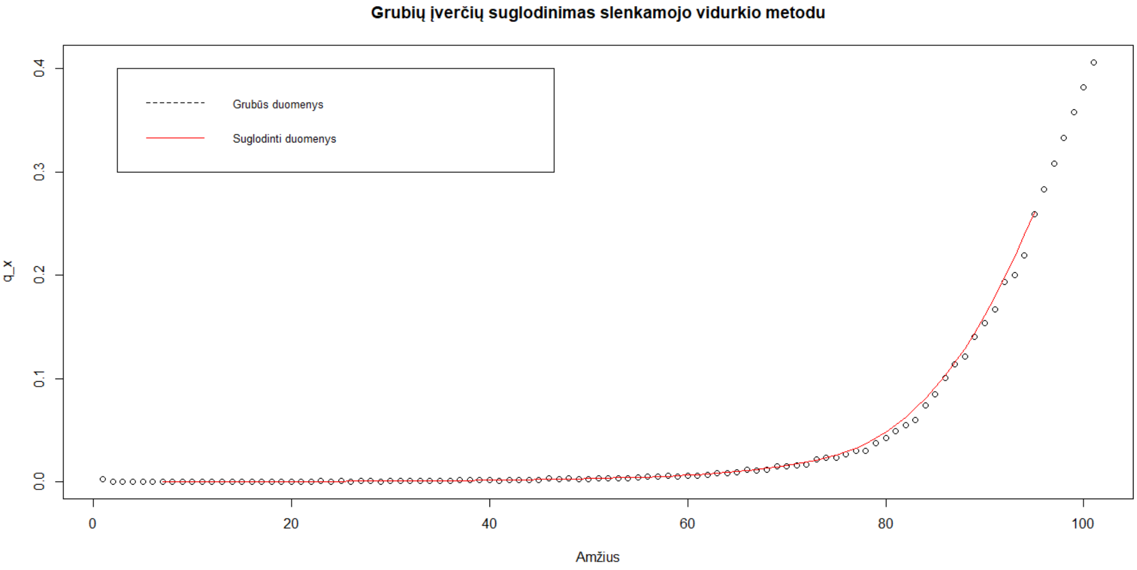


6 pav.: Vyrų grubūs ir suglodinti įverčiai, h = 10

Toliau pateikiama moterų (7 pav.) ir (8 pav.) bei vyrų (9 pav.), (10 pav.) grubių ir suglodintų reikšmių grafikai, kur buvo taikoma slenkamojo vidurkio metodas, (5) formulė. Pirmu atveju imama n=3, o kitu n=12, galime matyti, jog kuo n aukštesnis, tuo geriau suglodina pradinius įverčius.

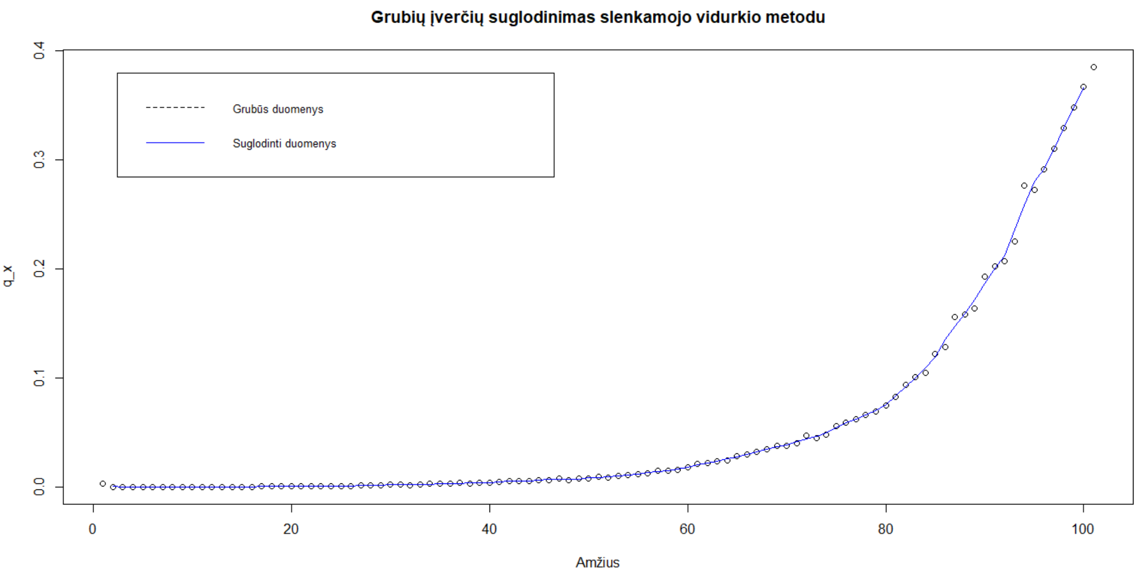


7 pav.: Moterų grubūs ir suglodinti įverčiai, n=3.

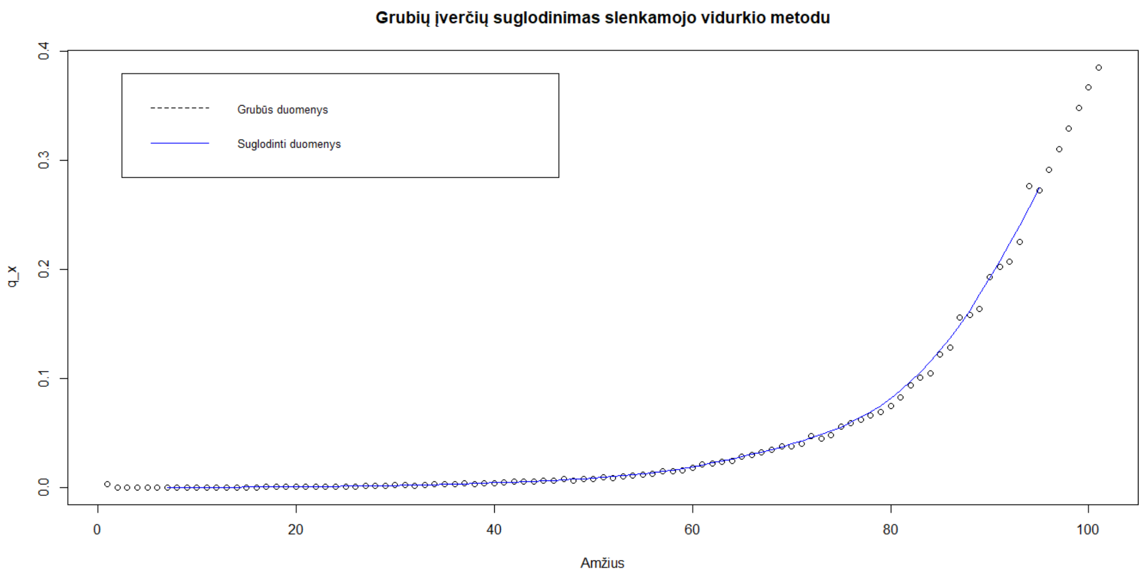


8 pav.: Moterų grubūs ir suglodinti įverčiai, n=12.

Panašiai ir vyrų atveju, kai imamos n=3 ir n=12 reikšmės.



9 pav.: Vyrų grubūs ir suglodinti įverčiai, n=3.



10 pav.: Vyrų grubūs ir suglodinti įverčiai, n=12.

Whittaker metodas geriau suglodina grubius įverčius nei slenkamojo vidurkio, nes (kai imamas n=3 atvejis) mažiau skiriasi nuo pirminių įverčių, todėl kreivė mažiau glodi. Be to, Whittaker metodo suglodinimas priklauso nuo pasirinktų parametrų reikšmių.

Žemiau (3 lentelė, 4 lentelė) taip pat pateikiama apskaičiuota lauktinoji likusio gyvenimo trukmė.

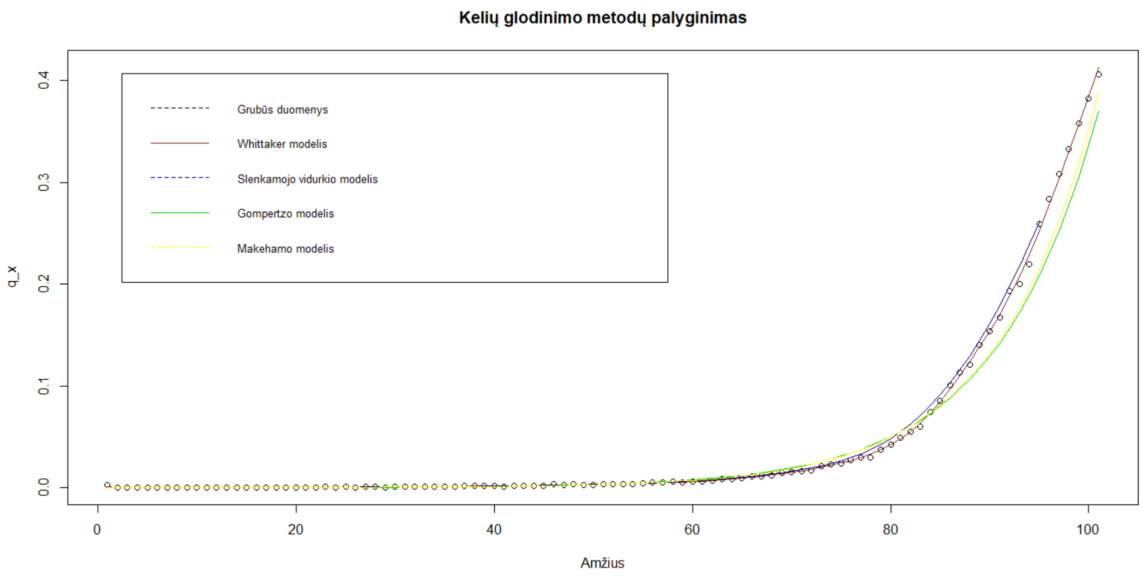
	Neglodinti duomenys	Whittaker metodas	Slenkamojo vidurkio metodas (n=3)
e_{65}	18.8077	18.8203	18.7854

3 lentelė: 65 metų moters likusio gyvenimo trukmė grubių įverčių atveju ir taikant metodus.

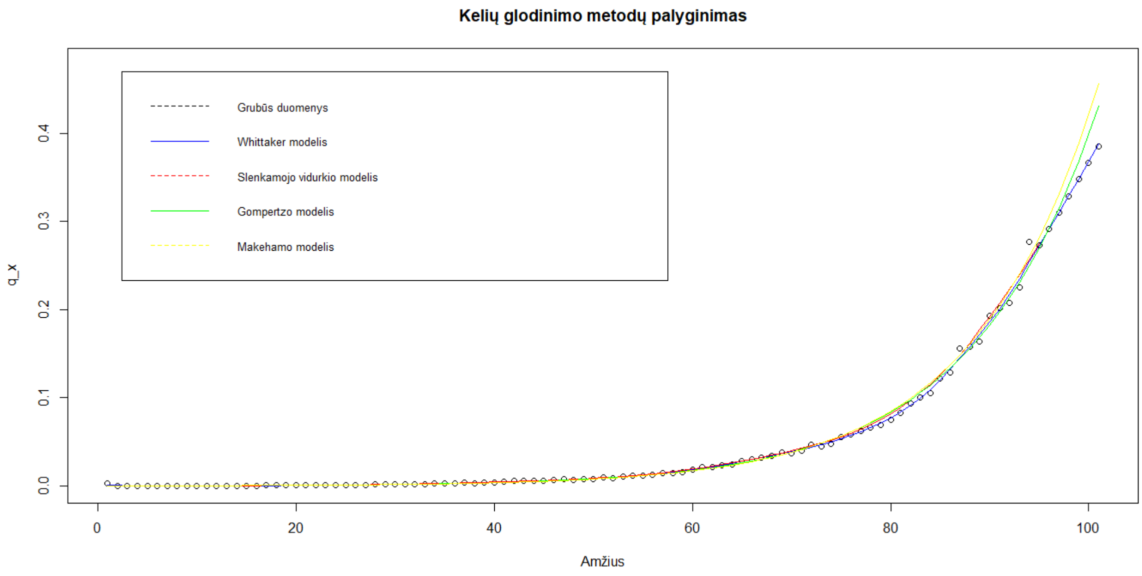
	Neglodinti duomenys	Whittaker metodas	Slenkamojo vidurkio metodas (n=3)
e_{65}	13.8748	13.8795	13.8582

4 lentelė: 65 metų vyro likusio gyvenimo trukmė grubių įverčių atveju ir taikant metodus.

Žemiau pateikti du grafikai: moterų (11 pav.) ir vyrų (12 pav.), kur pavaizduoti grubūs įverčiai ir suglodinti pagal visus keturis metodus. Geriausiai suglodina Whittaker metodas tiek moterų, tiek vyrų atveju.



11 pav.: Moterų grubūs ir suglodinti įverčiai.



12 pav.: Vyrų grubūs ir suglodinti įverčiai.

5 Išvados

Šiame darbe nagrinėjome du pagrindinius mirtingumo lentelės sudarymo etapus: gubių įverčių skaičiavimą ir jų suglodinimą. Apskaičiavus grubius įverčius jie yra suglodinami, naudojant parametrinius modelius, kaip Gompertzo, Makehamo, ir neparametrinius metodus, kaip Whittaker, slenkamojo vidurkio. Skaičiavimams naudojami 2017 metų Lietuvos gyventojų duomenys, buvo imami pirminiai mirtingumo tikimybių įverčiai, kurie vėliau buvo suglodinami.

Gompertzo ir Makehamo modeliai nelabai tiko glodinti moterų pradinius įverčius, nes gauti įverčiai buvo mažesni nei pradiniai. Geriausiai moterų ir vyrų grubius įverčius suglodino neparametrinis Whittaker metodas.

Literatūra

[1] Debon A., Montes F., Sala R. *A comparison of parametric models for mortality graduation. Application to mortality data for the Valencia Region (Spain)* SORT 29 (2) July-December 2005, 269-288

[2] Lapinskas Remigijus *Ekonometrija su kompiuteriu II* Vilnius, 2008.

[3] Leipus Remigijus. *Laiko eilutės* Vilnius, 2019.

[4] London Dick. *Graduation: The revision of estimates* Winsted, CT: Actex Publications, 1985.

[5] Marielynn E. Chanco *Mortality Rates Estimation Using Whittaker-Henderson Graduation Technique* Journal of the Mathematical Society of the Philippines, Vol. 39 Special Issue (2016) pp. 7-16

[6] Missov Trifon I., Lenart Adam,Nemeth Laszlo, Canudas-Romo Vladimir, Vaupel James W. *The Gompertz force of mortality in terms of the modal age at death* VOLUME 32, Article 36, pp. 1031–1048, published 20 May 2015.

[7] The Human Mortality Database,
`https://www.mortality.org`

[8] Pitacco Ermanno, Denuit Michel, Haberman Steven, Olivieri Annamaria. *Modelling longevity dynamics for pensions and annuity business* Oxford university press, New York 2009.

[9] Ramonat Stefan J., Kaufhold Kai F. Kaufhold. *A Practitioner's Guide to Statistical Mortality Graduation* Society of Actuaries, 2018.

[10] Skučaitė Aldona. *Išgyvenamumo - demografiniai modeliai. Paskaitų konspektas* Vilnius, 2018 m. ruduo.

A Priedai

A.1 Žymėjimų lentelė.

Žymėjimas	Reikšmė
x	Amžius
t_x	"Tikroji" reikšmė.
n_x	Imties dydis
u_x	Grubūs (pirminiai) įverčiai
v_x	Suglodinti įverčiai
w_x	Stebėjimų svoris, naudojamas glodinimui
e_x	Įverčio paklaidos.

A.2 Whittaker metodas, funkcijos M minimizavimas [4]

M yra funkcija, sudaryta iš n nežinomų v_x įverčių reikšmių. Tada v_x įverčiai, kuris minimizuoja M, randami, sudarant n lygčių sistemą, gaunama iš M dalinių išvestinių ir jas prilyginant nuliui, t.y. (10)

$$\frac{\partial M}{\partial v_r} = 0 \tag{7}$$

Čia $r = 1, 2, \ldots, n$.

Nors šis standartinis skaičiavimo metodas nustato globalų minimumo tašką M, aprašysime kitą būdą minimizavimo problemai spręsti, t.y. naudosime matricinį vektorinį pavidalą. Vectorius ir matricas žymėsime pusjuodžiu šriftu.

Tarkime, kad $\mathbf{u}$ yra pradinio įverčio vektoriaus stulpelis, o $\mathbf{v}$ suglodintos reikšmės. Transponuotą $\mathbf{b}$ matricą žymėsime $\mathbf{b}'$. Turime

$\mathbf{u}' = [u_1, \ldots, u_n]$ ir $\mathbf{v}' = [v_1, \ldots, v_n]$

Tarkime

$$\mathbf{w} = \begin{bmatrix} w_1 & & \bigcirc \\ & \ddots & \\ \bigcirc & & w_n \end{bmatrix}$$

turime $n \times n$ diagonalinę matricą. Laikykime, kad $\boldsymbol{k_z v}$ yra tam tikra matrica, laipsnio z. Jei v_x turi n reikšmių, tada vektorius $\mathbf{v}$ yra $n \times 1$ eilės, o $\boldsymbol{k_z}$ bus $(n-z) \times n$ eilės, o $\boldsymbol{k_z v}$ bus $(n-z) \times 1$ ilgio vektorius. Pavyzdžiui, jei turime $z = 2$ ir $n = 6$ tada

$$\boldsymbol{k_2} = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{bmatrix}$$

Su nelyginėmis z reikšmėmis, tarkime $z = 3$ ir $z = 7$, turime

$$\boldsymbol{k_3} = \begin{bmatrix} -1 & 3 & -3 & 1 & 0 & 0 & 0 \\ 0 & -1 & 3 & -3 & 1 & 0 & 0 \\ 0 & 0 & -1 & 3 & -3 & 1 & 0 \\ 0 & 0 & 0 & -1 & 3 & -3 & 1 \end{bmatrix}$$

Naudojant šiuos $\mathbf{v}$, $\mathbf{u}$, $\mathbf{w}$ bei $\boldsymbol{k_z}$, galime pamatyti, kad funkcija M (6), gali būti užrašyta kaip (7)

$$(\mathbf{v}-\mathbf{u})'\mathbf{w}(\mathbf{v}-\mathbf{u}) + h(\boldsymbol{k_z v})'\boldsymbol{k_z v} = (\mathbf{v}-\mathbf{u})'\mathbf{w}(\mathbf{v}-\mathbf{u}) + h\mathbf{v}'(\boldsymbol{k_z}'\boldsymbol{k_z})\mathbf{v} \tag{8}$$

kai $\mathbf{y}$'$\mathbf{y}$ pateikia vektoriaus $\mathbf{y}$ stulpelio elementų kvadratų sumą.

Tą patį galima padaryti ir su (7) lygtimi ir gauname matrcios eilutę:

$$(\mathbf{w} + h\boldsymbol{k_z}'\boldsymbol{k_z})\mathbf{v} = \mathbf{wu} \tag{9}$$

A.3 2017 metų lietuvių moterų mirtingumo duomenys

Metai	Amžius	mx	qx	lx	dx
2017	0	0.0028	0.00279	100000	279
2017	1	0.00013	0.00013	99721	13
2017	2	0.00007	0.00007	99708	7
2017	3	0.00028	0.00028	99701	28
2017	4	0	0	99673	0
2017	5	0.00014	0.00014	99673	14
2017	6	0.00007	0.00007	99659	7
2017	7	0.00014	0.00014	99652	14
2017	8	0.00014	0.00014	99638	14
2017	9	0	0	99624	0
2017	10	0	0	99624	0
2017	11	0.00008	0.00008	99624	8
2017	12	0.00008	0.00008	99616	8
2017	13	0.00016	0.00016	99608	16
2017	14	0.00024	0.00024	99592	24
2017	15	0.00016	0.00016	99568	16
2017	16	0.00007	0.00007	99552	7
2017	17	0.00014	0.00014	99545	13
2017	18	0.00026	0.00026	99532	26
2017	19	0.00033	0.00033	99506	33
2017	20	0.00013	0.00013	99473	13
2017	21	0.00025	0.00025	99461	25
2017	22	0.00048	0.00048	99436	48
2017	23	0.00018	0.00018	99388	17
2017	24	0.00079	0.00079	99371	79
2017	25	0.00031	0.00031	99292	31
2017	26	0.00074	0.00074	99261	74
2017	27	0.00055	0.00055	99187	55
2017	28	0.00023	0.00023	99132	23
2017	29	0.00044	0.00044	99110	44
2017	30	0.00065	0.00065	99066	64
2017	31	0.00049	0.00049	99002	49
2017	32	0.00067	0.00067	98953	67
2017	33	0.00109	0.00108	98887	107
2017	34	0.00077	0.00077	98779	76
2017	35	0.00094	0.00093	98703	92
2017	36	0.00137	0.00137	98611	135
2017	37	0.00141	0.00141	98476	139
2017	38	0.00143	0.00143	98337	141
2017	39	0.00156	0.00156	98197	153
2017	40	0.00118	0.00118	98044	116

2017	41	0.0017	0.0017	97928	167
2017	42	0.00165	0.00165	97761	161
2017	43	0.00205	0.00204	97600	200
2017	44	0.00208	0.00207	97400	202
2017	45	0.00303	0.00302	97198	294
2017	46	0.00251	0.0025	96905	243
2017	47	0.00297	0.00297	96662	287
2017	48	0.00251	0.0025	96375	241
2017	49	0.00253	0.00252	96134	243
2017	50	0.0037	0.0037	95891	354
2017	51	0.00351	0.00351	95537	335
2017	52	0.00372	0.00371	95202	353
2017	53	0.00342	0.00342	94848	324
2017	54	0.00388	0.00387	94524	366
2017	55	0.00484	0.00483	94159	455
2017	56	0.00496	0.00495	93704	464
2017	57	0.00567	0.00565	93240	527
2017	58	0.00511	0.00509	92713	472
2017	59	0.00567	0.00565	92241	521
2017	60	0.00568	0.00567	91720	520
2017	61	0.00691	0.00689	91200	628
2017	62	0.00857	0.00853	90572	773
2017	63	0.0087	0.00866	89799	778
2017	64	0.0092	0.00916	89021	815
2017	65	0.01139	0.01133	88206	999
2017	66	0.01097	0.01091	87207	951
2017	67	0.01156	0.01149	86255	991
2017	68	0.01476	0.01465	85264	1249
2017	69	0.01503	0.01491	84015	1253
2017	70	0.01622	0.01609	82762	1331
2017	71	0.01683	0.01669	81430	1359
2017	72	0.0215	0.02127	80071	1703
2017	73	0.02343	0.02316	78368	1815
2017	74	0.02371	0.02343	76553	1794
2017	75	0.02716	0.02679	74759	2003
2017	76	0.03024	0.02979	72756	2168
2017	77	0.03017	0.02973	70588	2098
2017	78	0.03768	0.03699	68490	2533
2017	79	0.04328	0.04236	65957	2794
2017	80	0.05035	0.04911	63163	3102

2017	81	0.05662	0.05506	60061	3307
2017	82	0.06208	0.06021	56753	3417
2017	83	0.07678	0.07394	53336	3944
2017	84	0.08873	0.08496	49392	4196
2017	85	0.1062	0.10084	45196	4558
2017	86	0.12035	0.11352	40638	4613
2017	87	0.12894	0.12113	36025	4364
2017	88	0.1506	0.14005	31661	4434
2017	89	0.16623	0.15347	27227	4179
2017	90	0.18203	0.16685	23048	3846
2017	91	0.21431	0.19356	19203	3717
2017	92	0.22248	0.20021	15486	3100
2017	93	0.24676	0.21966	12386	2721
2017	94	0.29822	0.25952	9665	2508
2017	95	0.33036	0.28352	7157	2029
2017	96	0.36416	0.30806	5128	1580
2017	97	0.39935	0.33288	3548	1181
2017	98	0.43562	0.35771	2367	847
2017	99	0.47259	0.38226	1520	581
2017	100	0.50986	0.40629	939	382

A.4 2017 metų lietuvių vyrų mirtingumo duomenys

Metai	Amžius	mx	qx	lx	dx
2017	0	0.00295	0.00294	100000	294
2017	1	0.00025	0.00025	99706	25
2017	2	0.00013	0.00013	99681	13
2017	3	0.00033	0.00033	99668	33
2017	4	0.00027	0.00027	99635	26
2017	5	0.00034	0.00034	99609	34
2017	6	0.00013	0.00013	99575	13
2017	7	0.00026	0.00026	99562	26
2017	8	0.0002	0.0002	99535	20
2017	9	0.00022	0.00022	99515	22
2017	10	0.00015	0.00015	99493	15
2017	11	0.00038	0.00038	99478	38
2017	12	0.00023	0.00023	99440	23
2017	13	0	0	99417	0
2017	14	0.0003	0.0003	99417	30
2017	15	0.00015	0.00015	99387	15
2017	16	0.00062	0.00062	99372	62
2017	17	0.00051	0.00051	99310	51
2017	18	0.00091	0.00091	99259	90
2017	19	0.00104	0.00104	99169	103
2017	20	0.00066	0.00066	99066	66
2017	21	0.00094	0.00094	99001	93
2017	22	0.00092	0.00092	98907	91
2017	23	0.00094	0.00094	98816	93
2017	24	0.00095	0.00095	98724	93
2017	25	0.00108	0.00108	98630	106
2017	26	0.00123	0.00123	98524	121
2017	27	0.00174	0.00174	98403	171
2017	28	0.00175	0.00175	98232	172
2017	29	0.00232	0.00231	98060	227
2017	30	0.00234	0.00233	97833	228
2017	31	0.00193	0.00193	97605	188
2017	32	0.00217	0.00217	97417	211
2017	33	0.00301	0.00301	97205	292
2017	34	0.00292	0.00291	96913	282
2017	35	0.00331	0.0033	96631	319
2017	36	0.00369	0.00368	96312	354
2017	37	0.00297	0.00297	95957	285
2017	38	0.00404	0.00403	95672	385
2017	39	0.00372	0.00371	95287	354
2017	40	0.00456	0.00455	94933	432

2017	41	0.00556	0.00555	94501	524
2017	42	0.00562	0.00561	93977	527
2017	43	0.0056	0.00559	93450	522
2017	44	0.00631	0.00629	92928	585
2017	45	0.00664	0.00662	92343	612
2017	46	0.00776	0.00773	91732	709
2017	47	0.00643	0.00641	91023	583
2017	48	0.0082	0.00816	90439	738
2017	49	0.00785	0.00782	89701	701
2017	50	0.00982	0.00977	89000	870
2017	51	0.0085	0.00846	88130	746
2017	52	0.01031	0.01026	87385	896
2017	53	0.0114	0.01134	86488	981
2017	54	0.01201	0.01194	85508	1021
2017	55	0.01271	0.01263	84487	1067
2017	56	0.01469	0.01458	83419	1216
2017	57	0.01518	0.01507	82203	1239
2017	58	0.01586	0.01573	80964	1274
2017	59	0.01851	0.01834	79690	1462
2017	60	0.02138	0.02115	78229	1655
2017	61	0.02213	0.02189	76574	1676
2017	62	0.02416	0.02387	74898	1788
2017	63	0.02511	0.02479	73110	1813
2017	64	0.02887	0.02846	71298	2029
2017	65	0.0303	0.02985	69268	2068
2017	66	0.03303	0.03249	67201	2183
2017	67	0.03516	0.03456	65017	2247
2017	68	0.03875	0.03801	62771	2386
2017	69	0.03815	0.03744	60384	2261
2017	70	0.04124	0.0404	58124	2348
2017	71	0.04845	0.0473	55775	2638
2017	72	0.04561	0.0446	53137	2370
2017	73	0.04923	0.04805	50767	2439
2017	74	0.05756	0.05595	48328	2704
2017	75	0.06054	0.05876	45624	2681
2017	76	0.06393	0.06195	42943	2660
2017	77	0.06859	0.06632	40283	2671
2017	78	0.07183	0.06934	37612	2608
2017	79	0.07772	0.07481	35003	2619
2017	80	0.0866	0.083	32385	2688

2017	81	0.09824	0.09364	29697	2781
2017	82	0.1062	0.10085	26916	2714
2017	83	0.11096	0.10512	24201	2544
2017	84	0.13023	0.12227	21657	2648
2017	85	0.13737	0.12855	19009	2444
2017	86	0.16887	0.15572	16566	2580
2017	87	0.17178	0.15819	13986	2213
2017	88	0.17826	0.16367	11774	1927
2017	89	0.21364	0.19302	9847	1901
2017	90	0.22509	0.20232	7946	1608
2017	91	0.23134	0.20735	6338	1314
2017	92	0.25381	0.22522	5024	1132
2017	93	0.32121	0.27676	3893	1077
2017	94	0.31584	0.27277	2815	768
2017	95	0.34088	0.29124	2047	596
2017	96	0.36684	0.30998	1451	450
2017	97	0.39359	0.32887	1001	329
2017	98	0.421	0.34779	672	234
2017	99	0.4489	0.36661	438	161
2017	100	0.47713	0.38523	278	107

A.5 R kodas, Gompertzo metodas.

```
library(MortalityLaws)

DataM <- read.csv2("Life tableM.csv", dec = ".")
x <- 0:100
DTH_M <- as.numeric(DataM$dx)
ETR_M <- as.numeric(DataM$lx)
QX_M <- as.numeric(DataM$qx)

GompertzM <- MortalityLaw(x = x, qx = QX_M, law = 'gompertz')
GompertzMale <- GompertzM[["fitted.values"]]

plot(QX_M, type = "p", ylim = c(0, max(GompertzMale, QX_M)), xlab="Amžius",
     ylab="q_x", main = "Grubių įverčių suglodinimas Gompertzo modeliu")
lines(GompertzMale, type = "l", lty = 7, col = "blue")
legend("topleft", inset = 0.05, legend=c("Grubūs duomenys", "Suglodinti duomenys"),
       col=c("black", "blue"), lty=2:1, cex=0.8)

############# Moterys (F)
DataF <- read.csv2("Life tableF.csv", dec = ".")
x <- 0:100
DTH_F <- as.numeric(DataF$dx)
ETR_F <- as.numeric(DataF$lx)
MX_F <- as.numeric(DataF$mx)
QX_F <- as.numeric(DataF$qx)

GompertzF <- MortalityLaw(x = x, Dx = DTH_F, Ex = ETR_F, law = 'gompertz')
GompertzFemale <- GompertzF[["fitted.values"]]

plot(QX_F, type = "p", ylim = c(0, max(GompertzFemale, QX_F)), xlab="Amžius",
     ylab="q_x", main = "Grubių įverčių suglodinimas Gompertzo modeliu")
lines(GompertzFemale, type = "l", lty = 7, col = "red")
legend("topleft", inset = 0.05, legend=c("Grubūs duomenys", "Suglodinti duomenys"),
       col=c("black", "red"), lty=2:1, cex=0.8)
```

A.6 R kodas, Makehamo modelis.

```
library(MortalityLaws)

DataM <- read.csv2("Life tableM.csv", dec = ".")
x <- 0:100
DTH_M <- as.numeric(DataM$dx)
ETR_M <- as.numeric(DataM$lx)
QX_M <- as.numeric(DataM$qx)


MakehamM <- MortalityLaw(x = x, qx = QX_M, law = 'makeham')
MakehamMale <- MakehamM[["fitted.values"]]

plot(QX_M, type = "p", ylim = c(0, max(MakehamMale, QX_M)), xlab="Amžius",
     ylab="q_x", main = "Grubių įverčių suglodinimas Makehamo modeliu")
lines(MakehamMale, type = "l", lty = 7, col = "blue")
legend("topleft", inset = 0.05, legend=c("Grubūs duomenys", "Suglodinti duomenys"),
       col=c("black", "blue"), lty=2:1, cex=0.8)



############# Moterys (F)
DataF <- read.csv2("Life tableF.csv", dec = ".")
x <- 0:100
DTH_F <- as.numeric(DataF$dx)
ETR_F <- as.numeric(DataF$lx)
MX_F <- as.numeric(DataF$mx)
QX_F <- as.numeric(DataF$qx)

MakehamF <- MortalityLaw(x = x, qx = QX_F, law = 'makeham')
MakehamFemale <- MakehamF[["fitted.values"]]

plot(QX_F, type = "p", ylim = c(0, max(MakehamFemale, QX_F)), xlab="Amžius",
     ylab="q_x", main = "Grubių įverčių suglodinimas Makehamo modeliu")
lines(MakehamFemale, type = "l", lty = 7, col = "red")
legend("topleft", inset = 0.05, legend=c("Grubūs duomenys", "Suglodinti duomenys"),
       col=c("black", "red"), lty=2:1, cex=0.8)
```

A.7 R kodas, Whittaker metodas.

```
library(MortalityTables)
library(pracma)
dev.off()

Whittaker_Smooth <- function(a, d, e, h){
  Obs <- log(d/e)
  n <- length(a)
  E <- diag(n)
  D <- diff(E, diff = 2)
  Smooth = solve(E + h*t(D) %*% D, Obs)
  plot(a + 0.5, Obs, xlab = "Age", ylab = "log(mortality)")
  lines(a + 0.5, Smooth, col = "red", lwd = 2)
  return(list(Smooth = Smooth))
}

### Duomenys iš failo, moterys.
DataF <- read.csv2("Life tableF.csv", dec = ".")
Age <- 0:100
Dth <- DataF$dx
Exp <- DataF$lx
SmoothF <- DataF$WFsmooth
qx <- as.numeric(DataF$qx)



### Grafikas, h = 1000
## Keičiant h, gaune kitokį grafiką.
OutputF <- Whittaker_Smooth(Age, Dth, Exp, 10)
title(main = "Whittaker method")
Smooth <- OutputF$Smooth
as.matrix(OutputF[["Smooth"]])

plot(qx, type = "p", ylim = c(0, max(qx, SmoothF)), xlab="Amžius",
     ylab="q_x", main = "Grubių įverčių suglodinimas Whittaker metodu")
lines(SmoothF, type = "l", lty = 7, col = "red")
legend("topleft", inset = 0.05, legend=c("Grubūs duomenys", "Suglodinti duomenys"),
       col=c("black", "red"), lty=2:1, cex=0.8)




###############Vyrai

library(MortalityTables)
library(pracma)
```

```
Whittaker_Smooth <- function(a, d, e, h){
  Obs <- log(d/e)
  n <- length(a)
  E <- diag(n)
  D <- diff(E, diff = 2)
  Smooth = solve(E + h*t(D) %*% D, Obs)
  plot(a + 0.5, Obs, xlab = "Age", ylab = "log(mortality)")
  lines(a + 0.5, Smooth, col = "blue", lwd = 2)
  return(list(Smooth = Smooth))
}

### Duomenys iš failo, vyrai..
DataM <- read.csv2("Life tableM.csv", dec = ".")
Age <- 0:100
Dth <- as.numeric(DataM$dx)
Exp <- as.numeric(DataM$lx)
SmoothM <- as.numeric(DataM$Wmsmooth)


OutputM <- Whittaker_Smooth(Age, Dth, Exp, 10)
title(main = "Whittaker method")
Smooth <- OutputM$Smooth
as.matrix(OutputM[["Smooth"]])

plot(DataM[,4], type = "p", ylim = c(0, max(DataM[,4], SmoothM)), xlab="Amžius",
     ylab="q_x", main = "Grubių įverčių suglodinimas Whittaker metodu")
lines(SmoothM, type = "l", lty = 7, col = "blue")
legend("topleft", inset = 0.05, legend=c("Grubūs duomenys", "Suglodinti duomenys"),
       col=c("black", "blue"), lty=2:1, cex=0.8)
```

A.8 R kodas, slenkamojo vidurkio metodas.

```
library(MortalityLaws)
library(forecast)
library(fpp2)

#### Vyrai (M)
DataMa <- as.ts(read.csv2("Life tableM.csv", dec = "."))

plot(DataMa[,4], type="p", xlab="Amžius",
     ylab="q_x", main = "Grubių įverčių suglodinimas slenkamojo vidurkio metodu")
lines(ma(DataMa[,4], 3), col="blue")
legend("topleft", inset = 0.05, legend=c("Grubūs duomenys", "Suglodinti duomenys"),
       col=c("black", "blue"), lty=2:1, cex=0.8)

##### Moterys (F)
DataFe <- as.ts(read.csv2("Life tableF.csv", dec = "."))

plot(DataFe[,4], type="p", xlab="Amžius",
     ylab="q_x", main = "Grubių įverčių suglodinimas slenkamojo vidurkio metodu")
lines(ma(DataFe[,4], 3), col="red")
legend("topleft", inset = 0.05, legend=c("Grubūs duomenys", "Suglodinti duomenys"),
       col=c("black", "red"), lty=2:1, cex=0.8)
```

A.9 R kodas, visi metodai.

```
library(MortalityLaws)
library(MortalityTables)

DataM <- read.csv2("Life tableM.csv", dec = ".")
DataMa <- as.ts(read.csv2("Life tableM.csv", dec = "."))
Age <- 0:100
DTH_M <- as.numeric(DataM$dx)
ETR_M <- as.numeric(DataM$lx)
QX_M <- as.numeric(DataM$qx)
Dth <- as.numeric(DataM$dx)
Exp <- as.numeric(DataM$lx)
SmoothM <- as.numeric(DataM$Wmsmooth)

plot(DataM[,4], type = "p", ylim = c(0, max(DataM[,3], SmoothM, GompertzMale,
      MakehamMale)), xlab="Amžius", ylab="q_x", main = "Kelių glodinimo metodų palygin
lines(SmoothM, type = "l", lty = 7, col = "blue")
lines(ma(DataMa[,4], 12), col="red")
lines(GompertzMale, type = "l", lty = 7, col = "green")
lines(MakehamMale, type = "l", lty = 7, col = "yellow")
legend("topleft", inset = 0.05, legend=c("Grubūs duomenys", "Whittaker modelis",
        "Slenkamojo vidurkio modelis", "Gompertzo modelis", "Makehamo modelis"),
       col=c("black", "blue", "red", "green", "yellow"), lty=2:1, cex=0.8)

################
DataF <- read.csv2("Life tableF.csv", dec = ".")
DataFe <- as.ts(read.csv2("Life tableF.csv", dec = "."))
Age <- 0:100
Dth <- as.numeric(DataF$dx)
Exp <- as.numeric(DataF$lx)
SmoothF <- DataF$WFsmooth
qx <- as.numeric(DataF$qx)
DTH_F <- as.numeric(DataF$dx)
ETR_F <- as.numeric(DataF$lx)
QX_F <- as.numeric(DataF$qx)

plot(qx, type = "p", ylim = c(0, max(qx, SmoothF)), xlab="Amžius",
     ylab="q_x", main = "Kelių glodinimo metodų palyginimas")
lines(SmoothF, type = "l", lty = 7, col = "brown4")
lines(ma(DataFe[,4], 12), col="darkblue")
lines(GompertzFemale, type = "l", lty = 7, col = "green3")
lines(MakehamFemale, type = "l", lty = 7, col = "yellow")
legend("topleft", inset = 0.05, legend=c("Grubūs duomenys", "Whittaker modelis",
"Slenkamojo vidurkio modelis", "Gompertzo modelis", "Makehamo modelis"),
       col=c("black", "brown4", "darkblue", "green3", "yellow"),
lty=2:1, cex=0.8)
```