

VILNIAUS UNIVERSITETAS
MATEMATIKOS IR INFORMATIKOS FAKULTETAS

Magistro darbas

HLM modeliai TIMSS duomenims

HLM models for TIMSS data

Eglė Kaleckaitė

VILNIUS 2013

MATEMATIKOS IR INFORMATIKOS FAKULTETAS
EKONOMETRINĖS ANALIZĖS KATEDRA
MATEMATINĖS ANALIZĖS KATEDRA

Darbo vadovas prof. Vydas Čekanavičius

Darbo recenzentas _____

Darbas apgintas _____

Darbas įvertintas _____

Registravimo NR. _____

Atidavimo į katedrą data _____

Turinys

| | |
|---|-----------|
| Turinys | 1 |
| 1 ĮVADAS | 3 |
| 2 HIERARCHINIAI TIESINIAI MODELIAI | 5 |
| 2.1 Hierarchinio tiesinio modelio aprašymas | 5 |
| 2.2 Modelio parinkimas | 7 |
| 3 PARAMETRŲ VERTINIMO METODAI | 9 |
| 3.1 Efektų vertinimas | 9 |
| 3.2 REML metodas | 10 |
| 3.3 MINQUE metodas | 10 |
| 3.4 Saviranka | 12 |
| 4 EMPIRINĖS SIMULIACIJOS | 15 |
| 4.1 Monte Carlo simuliacijos | 15 |
| 4.2 Simuliacijų rezultatai | 16 |
| 5 HLM MODELIS TIMSS DUOMENIMS | 20 |
| 5.1 TIMSS tyrimo aprašymas | 20 |
| 5.2 TIMSS tyrimų apžvalga | 21 |
| 5.3 TIMSS duomenys | 22 |
| 5.4 HLM modelis 2011 metų TIMSS duomenims | 24 |
| 6 IŠVADOS | 27 |
| LITERATŪRA IR ŠALTINIAI | 28 |
| LITERATŪRA IR ŠALTINIAI | 28 |
| Priedas | 31 |
| Lentelių sąrašas | 37 |
| Lentelių sąrašas | 37 |
| Iliustracijų sąrašas | 37 |
| Iliustracijų sąrašas | 37 |

HLM modeliai TIMSS duomenims

Santrauka

Šio magistro darbo tikslas sudaryti hierarchinį tiesinį modelį TIMSS tyrimo Lietuvos matematinio raštingumo testo rezultatams bei nustatyti geriausią vertinimo metodologiją šio tipo duomenims.

Raktiniai žodžiai: TIMSS, PWIGLS, HLM

HLM model for TIMSS data

Abstract

The main objective of this paper is to construct a hierarchical linear model for TIMSS survey Lithuanian mathematical literacy results and to identify the best method for this kind of data estimation.

Key words: TIMSS, PWIGLS, HLM

1 ĮVADAS

Hierarchiniai tiesiniai modeliai (*angl. Hierarchical Linear Models*) naudojami tirti tokias struktūras, kuomet vieni individai yra suskirstyti į kelias nesikertančias grupes, o šios yra dar didesnių grupių pogrupiai. Tipinis pavyzdys: mokyklų duomenys, kai mokiniai yra klasėse, o klasės mokyklose; arba organizcijos, kurių darbuotojai dirba skirtinguose padaliniuose. Šie modeliai naudojami ekonomikoje, biologijoje ir dar daug kur.

HLM modeliai yra platesnės mišrių efektų modelių klasės dalis. Dažniausiai sutinkama vertinimo procedūra yra REML - apribotas didžiausio tikėtinumo metodas (*Restricted Maximum Likelihood*), kuris reikalauja paklaidų normalumo prielaidos išpildymo. 1970-1972 metais Rao išleido straipsnių ciklą, kur pristatė naują mišrių efektų modelių vertinimo procedūrą, nereikalaujančią anksčiau minėtos normalumo prielaidos MINQUE - nepaslinktą mažiausios kvadratinės normos įvertinį (*Minimum Norm Quadratic Unbiased Estimator*). 1992 metais Bagakas pritaikė MINQUE dviejų lygių HLM modeliams su atsitiktiniu postūmiu vertinti. Kadangi MINQUE procedūrai nebuvo apibrėžtas standartinių paklaidų skaičiavimas, Bagakas tam panaudojo savirankos metodą (*angl. bootstrap*). 2006 metais Delpish savo daktaro disertacijoje simuliacijų būdu nustatė, jog MINQUE su saviranka duoda tikslesnius rezultatus nei REML. Vienas iš šio magistro darbo tikslų yra Monte Carlo simuliacijų būdu nustatyti, ar TIMSS duomenims MINQUE metodas gali potencialiai duoti geresnius rezultatus.

HLM neretai naudojamas ir TIMSS - tarptautinis matematikos ir gamtos mokslų tyrimas (*Trends in International Mathematics and Science Study*) - duomenims tirti. Šie duomenys ypatingi tuo, jog regresoriai yra indeksai, kategoriniai kintamieji bei sveikieji skaičiai (pvz.: mokinių skaičius mokykloje). Todėl normalumas negali būti garantuotas. Kita problema, jog TIMSS imtys sudaromos informatyviai ir yra pateikiami imties svoriai kiekvienam lygiui. HLM su imties svoriais vertinti naudojamas PWIGLS - iteratyvus tikimybėmis pasvertas apibendrintasis mažiausių kvadratų metodas (*angl. Possibility Weighted Iterative General Least Squares*), kuris kaip ir REML reikalauja paklaidų normalumo. Vertinimui dažniausiai naudojami paketai yra *HLM7* bei *MLwiN*. Tačiau šie paketai yra mokami ir galima legaliai gauti tik trumpalaikes apribotas versijas. Mokslinėje visuomenėje yra paplitęs nemokamas ir laisvai prieinamas statistinės analizės paketas \mathbb{R} . Nors hierarchinių modelių vertinimo su svoriais metodas nėra realizuotas, šiame darbe naudotas \mathbb{R} statistinės analizės paketas.

Lietuva dalyvavo visose TIMSS tyrimo cikluose. Tačiau HLM modeliai nebuvo plačiai analizuoti. Todėl kitas šio darbo tikslas yra sudaryti sudėtingesnę hierarchinę tiesinę modelį Lietuvos TIMSS (2011 metų) duomenims ir jį įvertinti pasirenkant tinkamą parametrų vertinimo metodą.

Šis magistro darbo struktūra išvardinta žemiau:

- 2 skyriuje aprašoma hierarchinių tiesinių modelių teorija.
- 3 skyrius skirtas hierarchinių tiesinių modelių parametrų vertinimo metodams. Aptariamas MINQUE (su saviranka) metodas ir jo realizacija \mathbb{R} aplinkoje.
- 4 skyrius skiriamas empirinėms parametrų vertinimo skirtingais metodais simuliacijoms.

- 5 skyriuje aprašomas TIMSS tyrimas, sudaromas bei įvertinamas modelis Lietuvos TIMSS duomenims atsižvelgiant į 4 skyriaus rezultatus.
- Pagrindiniai šio magistro darbo rezultatai dar kartą trumpai aptariami išvadose.

2 HIERARCHINIAI TIESINIAI MODELIAI

Šiame skyriuje pateikiamas HLM aprašymas, trumpai apžvelgiamos modelio parinkimo statistikos bei indeksai.

2.1 Hierarchinio tiesinio modelio aprašymas

Hierarchinis tiesinis modelis (toliau HLM) - metodas, kuris atsižvelgia į hierarchinę duomenų struktūrą ir galimą paklaidų heteroskedastiškumą. Literatūroje šis modelis dar vadinamas atsitiktinių efektų arba kovariacijų (dispersijos) komponentų modeliu ir yra apibendrintų mišrių tiesinių modelių klasės dalis[10].

HLM modeliai gali būti 2 ir daugiau lygių, tačiau kuo daugiau lygių, tuo vertinimo specifikacija ir vertinimo procedūra sudėtingesnė. Nors TIMSS (plačiau 5.1 skyrelyje) pateikia duomenis šalių, mokyklų, klasių (mokytojų) bei mokinių lygiuose, šiame darbe bus nagrinėjami tik dviejų (mokyklos ir mokinių) lygių HLM modeliai. Plačiau apie HLM modelius galima rasti Čekanavičiaus ir Murausko statistikos vadovėlyje [4].

Sakykime, jog turime J mokyklų, kuriose mokosi po n_j , $j = 1, \dots, J$ mokinių. Tuomet dviejų lygių modelio pavidalas:

$$\begin{cases} Y_{ij} = \beta_{0j} + \sum_{p=1}^P \beta_{pj} \times X_{pij} + \varepsilon_{ij}; \\ \beta_{pj} = \gamma_{p0} + \sum_{q=1}^{Q_p} \gamma_{pq} \times W_{pqj} + u_{pj}; \quad \forall p = 0, \dots, P, \end{cases} \quad (1)$$

kur

j - j -otoji mokykla;

ij - i -tasis mokinys j -tojoje mokykloje;

β_{0j} - atsitiktinis postūmis;

β_{pj} - atsitiktinis posvyris;

γ_{pq} - fiksuoti efektai;

X_{pij} - mokinio lygio kintamieji;

W_{pqj} - mokyklos lygio kintamieji;

Y_{ij} - aiškinamasis kintamasis - mokinio matematinio raštingumo rezultatas.

Ištačius antrą modelio (1) lygtį į pirmąją, gaunama jungtinė lygtis:

$$Y_{ij} = \sum_{p=1}^P \sum_{q=1}^{Q_p} \gamma_{pq} \times X_{pij} \times W_{pqj} + \sum_{p=1}^P X_{pij} \times u_{pj} + \epsilon_{ij}. \quad (2)$$

Tegu turime N mokinių, kurie yra natūraliai suskaidyti į J mokyklų ir $\sum_{j=1}^J n_j = N$. Tuomet (1) modelį galime užrašyti

$$\mathbf{Y}_j = \mathbf{Z}_j \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j; \quad (3)$$

$$\boldsymbol{\beta}_j = \mathbf{W}_j \boldsymbol{\gamma}_j + \mathbf{u}_j, \quad (4)$$

čia

$$\mathbf{Y}_j = \begin{pmatrix} Y_{1j} \\ Y_{2j} \\ \vdots \\ Y_{n_j j} \end{pmatrix}; \mathbf{Z}_j = \begin{pmatrix} 1 & X_{11j} & \cdots & X_{P1j} \\ 1 & X_{12j} & \cdots & X_{P2j} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n_j j} & \cdots & X_{Pn_j j} \end{pmatrix}; \boldsymbol{\beta}_j = \begin{pmatrix} \beta_{0j} \\ \beta_{1j} \\ \vdots \\ \beta_{Pj} \end{pmatrix}; \boldsymbol{\varepsilon}_j = \begin{pmatrix} \varepsilon_{1j} \\ \varepsilon_{2j} \\ \vdots \\ \varepsilon_{n_j j} \end{pmatrix};$$

$$\mathbf{W}_j = \begin{pmatrix} 1 & W_{1j} & \cdots & W_{Q_P j} & 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & W_{1j} & \cdots & W_{Q_P j} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}; \boldsymbol{\gamma} = \begin{pmatrix} \gamma_{00} \\ \vdots \\ \gamma_{0Q_P} \\ \gamma_{10} \\ \vdots \\ \gamma_{1Q_P} \\ \vdots \\ \gamma_{PQ_P} \end{pmatrix}; \mathbf{u}_j = \begin{pmatrix} u_{0j} \\ u_{1j} \\ \vdots \\ u_{Pj} \end{pmatrix}.$$

Tegu $X_{ij} = Z_{ij} \mathbf{W}_j$ ir $\mathbf{X}_j = (X'_{1j}, \dots, X'_{n_j j})'$, tada matricinis bendrojo HLM modelio pavidas yra

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}, \quad (5)$$

kur

$$\mathbf{Z} = \begin{pmatrix} \mathbf{Z}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{Z}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{Z}_J \end{pmatrix}.$$

Modelio prielaidos:

- $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_{n_j})$;
- $\mathbb{C}ov(X_{ij}, \varepsilon_{ij}) = 0, \forall i$;
- $\mathbb{C}ov(\varepsilon_j, \varepsilon'_j) = 0, \forall j \neq j'$;
- $\mathbf{u}_j = (u_{0j}, u_{1j}, \dots, u_{Pj})' \stackrel{iid}{\sim} \mathcal{N}(0, T)$;
- $\mathbb{C}ov(\mathbf{W}_j, \mathbf{u}_j) = 0$;
- $\mathbb{C}ov(\boldsymbol{\varepsilon}_j, \mathbf{u}_j) = 0$;
- $\mathbb{C}ov(\mathbf{X}_j, \mathbf{u}_j) = 0$;
- $\mathbb{C}ov(\mathbf{W}_j, \boldsymbol{\varepsilon}_j) = 0$.

Visos modelio paklaidos u_{ij} , ε_{ij} yra nepriklausomos ir vienodai pasiskirstę pagal normalųjį skirstinį, t.y.

$$\mathbb{E} \begin{pmatrix} \mathbf{u} \\ \boldsymbol{\varepsilon} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}; \mathbb{C}ov \begin{pmatrix} \mathbf{u} \\ \boldsymbol{\varepsilon} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \mathbf{T} & 0 & \cdots & 0 \\ 0 & \mathbf{T} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{T} \end{pmatrix},$$

kur \mathbf{R} vektoriaus $\boldsymbol{\varepsilon}$ kovariacijų matrica, dažniausiai $\mathbf{R} = \sigma^2 \mathbf{I}$. Nesunkiai galima parodyti, jog $\mathbb{E}(\mathbf{Y}) = \mathbf{X}\boldsymbol{\gamma}$ ir

$$\text{Var}(\mathbf{Y}) = \mathbf{V} = \mathbf{Z}\mathbf{Z}' + \mathbf{R} \quad (6)$$

arba

$$\text{Var}(\mathbf{Y}_j) = \mathbf{V}_j = \mathbf{Z}_j \mathbf{T} \mathbf{Z}_j' + \mathbf{R}_j \quad (7)$$

2.2 Modelio parinkimas

Hierarchinių tiesinių modelių sudarymas yra sudėtinga procedūra ir tai galima padaryti įvairiais būdais bei gauti labai skirtingus rezultatus.

Pradžioje visuomet skaičiuojamas nulinis (besąlyginis) modelis:

$$\begin{cases} Y_{ij} = \beta_{0j} + \varepsilon_{ij}, \varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2); \\ \beta_{0j} = \gamma_{00} + u_{0j}, u_{0j} \sim \mathcal{N}(0, \tau_{00}). \end{cases} \quad (8)$$

Gavus (8) įverčius $\hat{\gamma}_{00}$, $\hat{\sigma}^2$, $\hat{\tau}_{00}$, nustatoma, ar turima hierarchinė struktūra, ar visgi užtenka paprasto tiesinio modelio. Realiai yra tikrinamos hipotezės, ar gautas įvertis nelygus 0. Tai pat, kiek bendrą skirtumą paaiškina individų, o kiek grupių skirtumai - skaičiuojamas ICC koeficientas. Toliau toks modelis naudojamas sudėtingesnių modelių parinkime kaip atspirties taškas - skaičiuojami R^2 . (Remiamasi [4])

ICC - tarpklasinės koreliacijos koeficientas (*angl. Intraclass Correlation Coefficient*), kuris parodo grupių rezultatų ir rezultatų grupėse skirtumą [4]. Dar gali būti interpretuojamas kaip dviejų individų toje pačioje grupėje rezultatų koreliaciją. Šis koeficientas skaičiuojamas taip:

$$ICC = \frac{\tau_{00}}{\sigma^2 + \tau_{00}}. \quad (9)$$

Shackman savo 2001 metų darbe [47] pabrėžė: kuo mažesnis ICC koeficientas, tuo patikimesni bus modelio įverčiai. Tačiau, pasak jo, ICC nėra toks svarbus daugialygiuose modeliuose kaip DEFF - dizaino efektas (*ang. Design Effect*), kuris parodo kiek efektyvumo prarandama renkant imtis iš lizdų, o ne visiškai atsitiktinai parenkant individus. Šio koeficiento ryšys su ICC:

$$DEFF \approx [1 + (n - 1) \times ICC], \quad (10)$$

kur n yra vidutinis imties grupės dydis. Simuliacijų būdu, buvo nustatyta, jog kol $DEFF < 2$, tol rezultatai nėra netikri (*angl. spurious*).

Statistika R^2 dažnai naudojama ir vieno, ir kelių lygių modeliuose. Ji naudojama nustatyti ar įtrauktas kintamasis (įtraukti kintamieji) pagerina modeliuojamo kintamojo įverčio tikslumą. Hierarchiniuose modeliuose skaičiuojami atskiri R^2 kiekvienam lygiui. Pavyzdžiui, įtraukus pirmo lygio kintamąjį, imamas santykinis skirtumas $\frac{\hat{\sigma}_{(nul)}^2 - \hat{\sigma}_1^2}{\hat{\sigma}_{(nul)}^2}$ ($\hat{\sigma}_{(nul)}^2$ - besąlyginio modelio, $\hat{\sigma}_1^2$ - sąlyginio modelio) ir žiūrima, kiek skirtumo buvo paaiškinta. Įtraukus antro lygio kintamąjį analogiškai skaičiuojama $\frac{\hat{\tau}_{00(nul)} - \hat{\tau}_{00}}{\hat{\tau}_{00(nul)}}$. Šie koeficientai dar vadinami pirmo ir antro lygio R^2 .

Daugialygiuose modeliuose naudojamos ir kitokios aiškinamojo kintamojo įverčio tikslumo statistikos. Apie jas daugiau galima rasti 2008 metų Orien ir Edwards[33] straipsnyje. Autoriai simuliacijų būdu nustatė geriausią fiksuotų efektų įtraukimo į modelį (arba atitikimo *angl. goodness-of-fit*) matą, kurį pavadino R_1^2 . Šią statistiką sukūrė Vonesh ir Chinchilli[51]. Ji atrodo taip:

$$R_1^2 = 1 - \frac{\sum_{j=1}^J (Y_j - \hat{Y}_j)' (Y_j - \hat{Y}_j)}{\sum_{j=1}^J (Y_j - \bar{Y} \times \mathbb{1}_{n_j})' (Y_j - \bar{Y} \times \mathbb{1}_{n_j})}, \quad (11)$$

čia Y_j yra $n_j \times 1$ vektorius su j -tosios grupės stebėjimais, \bar{Y} - bendras visų stebėjimų vidurkis, $\hat{Y}_j = X\hat{\beta} + Z\hat{u}_j$, o $\mathbb{1}_{n_j}$ - vienetukų stulpelis.

Informaciniai indeksai. Kaip jau buvo minėta, modelio parinkimas susideda iš dviejų modelių palyginimo (t.y. imamas paprastesnis ir sudėtingesnis modeliai), tam yra sukurti informaciniai indeksai, kurie paremti didžiausio tikėtinumo f-jos logaritmo maksimumu l . Daugiau apie juos Čekanavičiaus ir Murausko statistikos vadovėlyje [4]. Šiame magistro darbe remiamasi $AIC = -2l + 2d$ kriterijumi, kur d - atsitiktinio poveikio parametrų modelyje skaičius. Čia verta paminėti, jog informaciniai indeksai gali būti skaičiuojami tik vertinant ML metodu.

3 PARAMETRŲ VERTINIMO METODAI

Praktikoje populiariausias hierarchinių tiesinių modelių parametrų vertinimo metodas - REML. Literatūroje nurodoma (Delphish 2006 [11]), jog šis metodas duomenims, kurie netenkina normalumo prielaidos, duoda tikslesnius parametrų (dispersijos komponentų) įverčius. Po ilgų paieškų R statistinės analizės programoje nepavyko rasti norimo MINQUE metodo hierarchiniams tiesiniams modeliams vertinti. Paketo *varComp* funkcija *minque* vertina tik centruotą aiškinamąjį kintamąjį. Paketo *maanova* funkcija *fitmaanova* vertina tik ANOVA mikro vektoriams. O Demidenko knygos [12] priede pateikta funkcija *lmevarMINQUE* vertina tik labai atskirą atvejį ir negalima pasirinkti *a priori* reikšmių. Dėl šios priežasties žemiau aprašytas MINQUE vertinimo metodas buvo realizuotas R aplinkoje pačios šio magistro baigiamojo darbo autorės.

3.1 Efektų vertinimas

Fiksuoti efektai dažniausiai vertinami taikant apibendrintąjį mažiausių kvadratų metodą (GLS):

$$\hat{\gamma} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}. \quad (12)$$

Taip pat ir atsitiktinio poveikio parametrų įverčiai:

$$\hat{\beta}_j = (\mathbf{Z}_j'\mathbf{R}_j^{-1}\mathbf{Z}_j)^{-1} \mathbf{Z}_j'\mathbf{R}_j^{-1}\mathbf{Y}_j; \quad \mathbf{R}_j = \sigma^2\mathbf{I}. \quad (13)$$

Atsitiktinės paklaidos \mathbf{u}_j įvertis gaunamas (pagal [29]) taip:

$$\hat{\mathbf{u}}_j = \mathbf{T}\mathbf{Z}_j'\hat{\mathbf{V}}_j^{-1} (\mathbf{Y}_j - \mathbf{X}_j\hat{\gamma}). \quad (14)$$

β_j įvertinys (13) turės didelę dispersiją, kai j -toji grupė maža, todėl įvertis gali būti visai nepanašus į tikrąją reikšmę. Dėl šios priežasties praktikoje dažnai naudojamas empirinis Bajaso įvertinys (EBS) dar vadinamas sutraukiančiuoju (*angl. shrinkage*) įvertiniu (pagal [30]):

$$\hat{\beta}_j^* = \Theta_j\hat{\beta}_j + (\mathbf{I} - \Theta_j)\mathbf{W}_j\hat{\gamma}, \quad (15)$$

čia $\Theta_j = T / (T + \sigma^2(\mathbf{Z}_j'\mathbf{Z}_j)^{-1})$. Visi aukščiau minimi įvertiniai remiasi tuo, jog T ir σ^2 yra žinomi, tačiau praktikoje juos reikia įsivertinti. Dispersijos komponentams vertinti yra ne vienas metodas, žemiau išvardinti keli iš jų.

Dar skaičiuojamos ir standartinės paklaidos, kurios gaunamos iš $(\text{Cov}(\hat{\gamma}))^{\frac{1}{2}}$ ir $(\text{Cov}(\hat{\mathbf{u}}_j))^{\frac{1}{2}}$:

$$\text{Cov}(\hat{\gamma}) = (\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1} \quad (16)$$

arba sumuštinio principo įvertinys (pagal [31]):

$$\text{Cov}(\hat{\gamma}) = (\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1} \frac{N}{N-1} \left(\sum_{j=1}^J \mathbf{X}_j'\hat{\mathbf{V}}_j^{-1}\hat{\mathbf{e}}_j\hat{\mathbf{e}}_j'\hat{\mathbf{V}}_j^{-1}\mathbf{X}_j \right) (\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}, \quad (17)$$

kur $\hat{\mathbf{e}}_j = \mathbf{Y}_j - \mathbf{X}_j\hat{\gamma}$, o $\text{Cov}(\hat{\mathbf{u}}_j)$ gaunama taip:

$$\text{Cov}(\hat{\mathbf{u}}_j) = \left(\sum_{i=1}^{n_j} \mathbf{Z}_j'\hat{\mathbf{V}}_j^{-1} \right) [\mathbf{V}_j - \mathbf{X}_j\text{Cov}(\hat{\gamma})\mathbf{X}_j']^{-1} \left(\sum_{i=1}^{n_j} \mathbf{Z}_j'\hat{\mathbf{V}}_j^{-1} \right)'. \quad (18)$$

Turint aukščiau išvardintus įverčius, galima gauti aiškinamojo kintamojo įvertį $\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\gamma}} + \mathbf{Z}\hat{\mathbf{u}}$ bei liekanas $\hat{\boldsymbol{\varepsilon}} = \mathbf{Y} - \hat{\mathbf{Y}}$.

3.2 REML metodas

Didžiausio tikėtinumo metodas (ML) (*angl. Maximum Likelihood*) [22] vienas dažniausiai naudojamų metodų vertinti HLM modelius. Modelio (5) j-tosios grupės su dispersijų matrica (7) didžiausio tikėtinumo f-jos logaritmas:

$$L_j(\sigma^2, \mathbf{T}, \boldsymbol{\gamma}) = -\frac{n_j}{2} \log(2\pi) - \frac{1}{2} \log |\mathbf{V}_j| - \frac{1}{2} (\mathbf{Y}_j - \mathbf{X}_j \boldsymbol{\gamma})' \mathbf{V}_j^{-1} (\mathbf{Y}_j - \mathbf{X}_j \boldsymbol{\gamma}) \quad (19)$$

Viso modelio log-tikėtinumas gali būti užrašytas kaip atskirų grupių suma $L(\sigma^2, \mathbf{T}, \boldsymbol{\gamma}) = \sum_{j=1}^J L_j(\sigma^2, \mathbf{T}, \boldsymbol{\gamma})$. Taip pat naudojamas ir apribotasis didžiausio tikėtinumo metodas (REML), kuris nuo ML skiriasi tuo, jog atskiria fiksuotų efektų dalį ir optimizuoja tik atsitiktinių efektų dalį:

$$L_j^R(\sigma^2, \mathbf{T}, \boldsymbol{\gamma}) = L_j(\sigma^2, \mathbf{T}, \boldsymbol{\gamma}) - \frac{1}{2} \log |\mathbf{X}_j' \mathbf{V}_j^{-1} \mathbf{X}_j| \quad (20)$$

Nors šios procedūros dažnai naudojamos vertinant HLM modelius, jos visgi reikalauja normalumo prielaidos, kuri negali būti garantuota, kai turime daug kategorinių aiškinančiųjų kintamųjų.

3.3 MINQUE metodas

Savo straipsnių serijoje (1970[36], 1971a[37], 1971b[38], 1972[39]) Rao pristatė dispersijos - kovariacijos matricių mišriuose tiesiniuose modeliuose vertinimo procedūrą, kurią pavadino mažiausios kvadratinės normos nepaslinktu įvertiniu - MINQUE (*angl. Minimum Norm Quadratic Unbiased Estimator*). Šis įvertinys išskirtinis tuo, jog yra nepaslinktas, invariantiškas ir nereikalauja normalumo prielaidos ar žinių apie paklaidų ar atsitiktinių efektų pasiskirstymą.. Tačiau įvertis priklauso nuo *a priori* reikšmių ir keli žmonės turėdami tuos pačius duomenis, bet skirtingas pradines reikšmes gali gauti visiškai skirtingus parametrų įverčius ir visi jie bus mažiausi kvadratinės normos įvertiniai, tačiau tik prie pasirinktų *a priori* reikšmių. Rao MINQUE procedūra bei jos pritaikymas HLM bus trumpai aprašyti šiame skyrelyje. Daugiau informacijos galima rasti minėtoje Rao straipsnių serijoje ([36], [37], [38], [39]).

Rao teorija remiasi Euklidinės normos minimizavimu. Pagal Rao, turime modelį $\mathbf{Y} = \mathbf{X}\boldsymbol{\gamma} + \mathbf{e}$, kur $\mathbf{e} = \mathbf{H}\boldsymbol{\xi} = \mathbf{H}_1\boldsymbol{\xi}_1 + \dots + \mathbf{H}_k\boldsymbol{\xi}_k$, $\boldsymbol{\xi}_i$ -atsitiktiniai vektoriai, \mathbf{H}_i -lydinčiosios matricos. Imama kvadratinė forma $\mathbf{Y}'\mathbf{A}\mathbf{Y}$ ir sakoma, kad ji yra mažiausias tiesinės f-jos $\sum_{r=1}^l g_r \theta_r$ (čia θ_r - dispersijos komponentės, o g_r - lydintysis parametras) įvertinys su sąlygomis:

1. $\mathbf{A}\mathbf{X} = \mathbf{0}$ - invariantiškumas;
2. $\mathbb{E}(\mathbf{Y}'\mathbf{A}\mathbf{Y}) = \sum_{r=1}^l g_r \theta_r$ - nepaslinktumasis;
3. $\text{Var}(\mathbf{Y}'\mathbf{A}\mathbf{Y}) = \min$ pagal $\theta_1, \dots, \theta_l$.

Jei žinotume visus ξ_i , tai natūralus $\sum_{r=1}^l g_r \theta_r$ įvertinys būtų $\boldsymbol{\xi}' \mathbf{R} \boldsymbol{\xi}$, kur \mathbf{R} - $k \times k$ matrica. Iš kitos pusės, pagal invariantiškumą, $\mathbf{Y}' \mathbf{A} \mathbf{Y} = \boldsymbol{\xi}' \mathbf{H}' \mathbf{A} \mathbf{H} \boldsymbol{\xi}$. Skirtumas tarp šių dviejų įvertinių yra $\boldsymbol{\xi}' (\mathbf{H}' \mathbf{A} \mathbf{H} - \mathbf{R}) \boldsymbol{\xi}$. Norima gauti šį skirtumą kuo mažesnį, tam sudaroma Euklidinė norma $\|\mathbf{H}' \mathbf{A} \mathbf{H} - \mathbf{R}\|^2$ ir ji yra minimizuojama.

Tuomet tiesinės kombinacijos $g_1 \theta_1 + \dots + g_l \theta_l$ MINQUE yra $\mathbf{Y}' \mathbf{A} \mathbf{Y}$, jei matrica \mathbf{A} gaunama išsprendus:

$$tr(\mathbf{A} \mathbf{V} \mathbf{A} \mathbf{V}) \rightarrow \min \text{ s.t. } \mathbf{A} \mathbf{X} = 0; \quad tr(\mathbf{A} \mathbf{Q}_r) = g_r$$

Tada $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \dots, \hat{\theta}_l)' = S^{-1} \mathbf{q}$, kur $S = \{s_{ij}\}$, $s_{ij} = tr(\mathbf{C} \mathbf{Q}_i \mathbf{C} \mathbf{Q}_j)$, o $\mathbf{q} = \{q_i\}$, $q_i = \mathbf{Y}' \mathbf{C} \mathbf{Q}_i \mathbf{C} \mathbf{Y}$, $\mathbf{C} = \hat{\mathbf{V}}^{-1} \left(\mathbf{I} - \mathbf{X} \left(\mathbf{X}' \hat{\mathbf{V}}^{-1} \mathbf{X} \right)^{-1} \mathbf{X}' \hat{\mathbf{V}}^{-1} \right)$, $\hat{\mathbf{V}} = \sum_{r=1}^l \alpha_r \mathbf{Q}_r$, α_r - *a priori* reikšmės (gaunamos iš išorės, pvz.: tvirtai žinant apytiksles dispersijos komponentų proporcijas). Fiksuotų parametrų įverčiai gaunami pagal GLS[39]:

$$\hat{\boldsymbol{\gamma}} = \left(\mathbf{X}' \hat{\mathbf{V}}^{-1} \mathbf{X} \right)^{-1} \mathbf{X}' \hat{\mathbf{V}}^{-1} \mathbf{Y}.$$

MINQUE pritaikymas HLM modeliams. Literatūros apie MINQUE pritaikymą hierarchiniams tiesiniams modeliams nėra daug arba sunku tokią gauti. Delpish[11] 2006 metų savo daktaro disertacijoje remiasi Bagakas[6] darbu, kuriame pristatomas minėtasis MINQUE metodas ir jo pritaikymas dviejų lygių hierarchiniams modeliams su atsitiktiniu postūmiu. Pasak autorių, MINQUE neturi procedūros skaičiuoti standartinės paklaidoms ir pasikliautinumo intervalams, tad šiam tikslui buvo pasitelktas savirankos metodas. Kita problema, MINQUE labai priklauso nuo *a priori* reikšmių, tad praktikoje dažniau taikomas iteratyvus metodas, kuris aprašytas toliau. Tačiau Bagakas darbo nepavyko gauti, o Delpish savo darbe tik labai abstrakčiai pristato vertinimo metodą. Be to, pristatomas metodas tik atsitiktinio postūmio atveju, tačiau vertinami modeliai su atsitiktiniu postūmiu ir posvyriu. Delpish modelius vertina su SAS programos PROC MOXED paketu, apie šį paketą plačios informacijos taip pat nepavyko gauti. Dėl šių priežasčių teko MINQUE metodą dviejų lygių HLM modeliams su atsitiktiniu postūmiu ir posvyriu išsivesti pačiai.

Tegul turime bendrą HLM modelio išraišką pavidalu (5) su \mathbf{Y}' ko kovariacijų matrica $\mathbf{V} = \mathbf{Z} \mathbf{T} \mathbf{Z}' + \sigma^2 \mathbf{I}$. Kadangi $\mathbf{T} = \{\tau_{ij}\}_{i,j=0}^P$ galima išskaidyti į $l = \frac{P(P+1)}{2}$ elementų sumą, t.y. $\mathbf{T} = \sum_{r=1}^l \theta_r \mathbf{T}_r$, kur $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_l)' = (\tau_{00}, \tau_{01}, \dots, \tau_{0q}, \tau_{11}, \dots, \tau_{q-1,q})'$ ir atitinkamai

$$\mathbf{T}_1 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}; \mathbf{T}_2 = \begin{pmatrix} 0 & 1 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}; \dots; \mathbf{T}_l = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Pažymėkime $\mathbf{Q}_r := \mathbf{Z} \mathbf{T}_r \mathbf{Z}'$. Tuomet galima užrašyti $\mathbf{V} = \sum_{r=1}^l \theta_r \mathbf{Q}_r + \sigma^2 \mathbf{I}$. Tegul $\theta_0 = \sigma^2$ ir $\mathbf{Q}_0 = \mathbf{I}$ ir galiausiai gauname

$$\mathbf{V} = \sum_{r=0}^l \theta_r \mathbf{Q}_r. \quad (21)$$

Tokiu būdu gavome dispersijos išskaidymą kaip ir Rao MINQUE atveju. Pagrindinį vaidmenį šioje vietoje atlieka *a priori* reikšmių nustatymas, norint gauti $\hat{\mathbf{V}}$. Kai nagrinėjame tik atsitiktinio

postūmio atvejį, galime pasitelkti tarpklausinės koreliacijos koeficientą ICC ir imti $\alpha_0 = 1 - \frac{\tau_{00}}{\tau_{00} + \sigma^2}$, o $\alpha_1 = \frac{\tau_{00}}{\tau_{00} + \sigma^2}$. Šis būdas veiks gerai, kai turėsime pakankamai tvirtas pradines žinias apie ICC.[11]

Tačiau praktikoje dažnai nežinome pradinių ICC reikšmių. Dėl šios priežasties dažnai taikomas iteratyvus MINQUE metodas, kuris aprašytas žemiau.

I-MINQUE metodas yra iteratyvi MINQUE procedūra, kuri yra nejautri *a priori* reikšmėms. Pavyzdžiui, paimkime $\alpha_0 = \dots = \alpha_l = 1$. Gauname $\hat{\mathbf{V}}_{(0)} = \sum_{r=0}^l \mathbf{Q}_r$, įsivertiname pagal MINQUE metodą $\hat{\boldsymbol{\theta}}_{(0)} = (\hat{\theta}_0, \hat{\theta}_1, \dots, \hat{\theta}_l)'_{(0)}$. Su naujomis θ reikšmėmis gaunamas $\hat{\mathbf{V}}_{(1)}$, o su juo įsivertiname $\hat{\boldsymbol{\theta}}_{(1)} = (\hat{\theta}_0, \hat{\theta}_1, \dots, \hat{\theta}_l)'_{(1)}$. Šioje vietoje procedūrą galime nutraukti arba tęsti tol, kol $|\hat{\boldsymbol{\theta}}_{(i)} - \hat{\boldsymbol{\theta}}_{(i+1)}| < \delta$, kur δ yra pasirinktas konvergavimo tolerancijos parametras.

Delpish[11] savo darbe naudoja tik du pirmuosius I-MINQUE žingsnius. Toks metodas bus naudojamas ir šiame darbe. Skyriaus pradžioje buvo minėta, jog \mathbb{R} statistinės analizės programoje nepavyko rasti reikiamos funkcijos minėtam metodui vertinti, todėl reikiamos funkcijos buvo parašytos pačios autorės. Kodas pateiktas šio darbo priede. Rašant kodą buvo susidurta su viena (pagrindine) problema. Kai turimas mažas duomenų kiekis, aprašytas MINQUE metodas veikia gan efektyviai. Tačiau TIMSS atveju, kai turime apie 4 500 stebėjimų, tenka apversti ir sudauginti labai dideles matricas ir tai reikalauja daug kompiuterio resursų ir laiko. Pati \mathbb{R} programa nėra pritaikyta tokio tipo skaičiavimas, tačiau egzistuoja paketai *Matrix* ir *matrixcalc*, kurie efektyviau vykdo matricų algebrą. Taip pat yra sukurtas \mathbb{R} papildymas BLAS, kuris įdiegiamas rankiniu būdu ir paspartina didelius skaičiavimus išskaidydamas vykdomą procesą per kelis procesoriaus branduolius. Plačiau apie tai internetiniame tinklapyje <http://www.r-bloggers.com/faster-r-through-better-blas>.

Pastabos. Savo straipsnių serijoje Rao pristatė ir kitokius mažiausios kvadratinės normos įvertinius, pats bendriausias - MINQE, kuris neapriboja invariantiškumu ir nepaslinktumu. Daugiau MINQE(I) - su invariantikumo apribojimu, MINQE(U) - tik su nepaslinktumo reikalavimu, MINQE(U,I) - tapatus nagrinėjamam MINQUE, MINVQUE - sudarytas prie normalumo prielaidos ir minimizuoja dispersiją (MIVQUE prie normalumo = MINQUE). Teorijoje ir praktikoje naudojami įvairūs MINQUE metodai, daugiau apie juos - 1979 metų Rao straipsnyje [40]. Kiekvienas iš šių metodų turi savo pliusų ir savo minusų. Rao, aukščiau minėtame straipsnyje, taip pat tiria ir MINQUE savybes ir parodo, jog REML pirmoji iteracija yra MINQUE, o REML įvertinys yra identiškas I-MINQUE, kai *apriori* (pradinės) reikšmės sutampa. Todėl MINQUE yra dažnai naudojamas gauti pradinės REML reikšmės.

3.4 Saviranka

Savirankos metodas pristatytas 1979 metais Efron [16] ir yra plačiai paplitęs modeliavime. Šio metodo esmė, perrinkti turimą imtį be galo daug kartų, kad būtų atspindėta populiacija, jos skirstinys. Kitaip: sakome, kad turima imtis yra mini populiacija ir iš jos traukiame be galo daug imčių, taip siekdami gauti stabilesnius įverčius. Tad šis metodas naudojamas poslinkio koregavimui, standartinėms paklaidoms bei pasikliautiniams intervalams gauti. Ir reikalauja iš turimos imties tik tiek, jog ji būtų atsitiktinai parinkta ir kuo geriau atspindėtų populiaciją. Standartiškai, iš turimos imties parenkama tokio paties dydžio imtis su grąžinimu ir skaičiuojami įverčiai ar kitos charakteristikos. Tai kartojama kažkokį didelį B kartų skaičių ir atliekami papildomi skaičiavimai, kaip kad randamos standartinės paklaidos. Tačiau turint kelių lygių arba hierarchinę struktūrą (tuo labiau lizdinę imtį), savirankos metodą reikia pritaikyti turimam atvejui, kadangi matavimai nėra nepriklausomi ir vienodai pasiskirstę.

Hierarchiniams duomenims (kaip ir paprastai) taikomi parametriniai, neparametriniai ir pusiau parametriniai, atvejų ir paklaidų savirankos metodai. Leeden ir kiti [28] savo 1997 metų straipsnyje Monte Carlo simuliacijų būdu tyrė tris metodus taikytus dviejų lygių duomenims: atvejų saviranką, paklaidų saviranką ir parametrinę saviranką, kai modeliai vertinami didžiausio tikėtino metodo (ML). Jų rezultatai parodė, jog fiksuotų parametrų įverčiai yra nepaslinkti ir saviranka neatneša jokios naudos, tačiau standartinės paklaidos, įvertintos ML metodu yra paslinktos į mažąją pusę ir atvejų saviranka čia duoda geriausius rezultatus. Kiek kitaip yra su atsitiktiniais efektais, čia mažiausią poslinkį turi, autorių vadinama, sutraukianti saviranka. Bet standartinių paklaidų atžvilgiu geriausia atvejų saviranka. Nors atvejų savirankos metodas nebuvo pripažintas geriausiu, autoriai rekomenduoja šį metodą, nes jis reikalauja mažiausiai prielaidų apie modelį ar turimų duomenų struktūrą. Taigi, neparametrinė atvejų saviranka bus naudojama šiame magistro darbe.

Dėl hierarchinės struktūros sudėtingumo, neparametrinės atvejų savirankos imtys gali būti sudaromos įvairiai: perrenkant tik vieną iš lygių, kelis arba visus, perrenkant su grąžinimu ir be jo. Ren ir kiti [41] straipsnyje apie neparametrinę saviranką parodė, jog geriausia perrinkti su grąžinimu aukščiausiam lygyje, o visuose kituose lygiuose perrinkti be grąžinimo. Taip ir bus daroma šiame darbe. Toliau bus pateiktas naudojamas neparametrinės atvejų savirankos metodas.

Neparametrinė atvejų saviranka. Sakykime, jog turime H lizdų, kiekviename iš jų turime po n_h mokyklų, o j - tojoje mokykloje iš h - tojo lizdo mokosi n_{hj} mokinių. Tuomet saviranka atliekama taip:

1. Iš kiekvieno lizdo h ($h = 1, \dots, H$) parenkama $n_h - 1$ mokyklų su grąžinimu. Kiekvienai parinktai mokyklai j iš h - tojo lizdo paimame atitinkamą aiškinančiųjų kintamųjų rinkinį \mathbf{W}_{jh}^* .
2. Tuomet kiekvienoje parinktoje mokykloje j (jei mokykla kartojasi, kiekvieną kart parenkama nepriklausomai) parenkama $n_{hj} - 1$ mokinių be grąžinimo. Ir kiekvienam parinktam i - tajam mokiniui iš j - tosios mokyklos iš h - jojo lizdo paimamas kintamųjų rinkinys $(\mathbf{Y}_{ijh}^*, \mathbf{Z}_{ijh}^*)$. Kiekvieną pirmo lygio kintamąjį atitinkamai sujungus su antro lygio kintamuoju, gaunama savirankos imtis.
3. Gautai savirankos imčiai įvertinami norimi modelio parametrai.
4. Žingsniai 1-3 kartojami be galo daug kartų B.

Taip gaunama B savirankos parametrų įverčių, kurie apdorojami toliau aprašytu būdu.

Savirankos metodas. Sakykime, kad $\chi = \{X_1, \dots, X_n\} \overset{vnp}{\sim} F$, kur F yra nežinoma pasiskirstymo f-ja. Statistikai $\hat{\theta}$ pasirinktai įvertinti pasiskirstymo F parametrai θ gaunamas savirankos pasiskirstymas, perrenkant su grąžinimu iš χ . Praktikoje pasirenkamas pakankamai didelis skaičius B savirankos imčių ir gaunama $\hat{\theta}_b^*$ ($b = 1, \dots, B$) savirankos įverčių, kurie atspindi $\hat{\theta}$ pasiskirstymą. Tuomet parametro θ savirankos įvertis gaunamas suvidurkinus visas $\hat{\theta}_b^*$, o standartinės paklaidos (*s.e.*) taip:

$$\hat{\theta}_{(.)}^* = \frac{1}{B} \sum_{b=1}^B \hat{\theta}_b^*; \text{ s.e.}(\hat{\theta}) = \left[\frac{1}{B-1} \sum_{b=1}^B (\hat{\theta}_b^* - \hat{\theta}_{(.)}^*)^2 \right]^{\frac{1}{2}}.$$

Dar skaičiuojamas įverčio poslinkis $\widehat{bias}(\hat{\theta}) = \hat{\theta}_{(.)}^* - \hat{\theta}$ ir gaunama parametro poslinkio korekcija $\hat{\theta}_{boot} = \hat{\theta} - \widehat{bias}(\hat{\theta}) = 2\hat{\theta} - \hat{\theta}_{(.)}^*$. Skaičiuojami ir savirankos pasikliautiniai intervalai, čia bus aprašytas ir naudojamas tik percentilių metodas, kuris nereikalauja žinių apie θ pasiskirstymą.

Apskaičiuojami $d_b^* = \hat{\theta}_b^* - \hat{\theta}$ ir parenkamas d_c toks, kad $(1 - \alpha)100\%$ visų d_b^* reikšmių patektų į intervalą $\pm d_c$. Tuomet $(1 - \alpha)100\%$ savirankos pasikliautiniai intervalai gaunami iš $\hat{\theta} \pm d_c$, čia α - pasirinktas reikšmingumo lygmuo.

4 EMPIRINĖS SIMULIACIJOS

Šiame skyriuje aprašomos vykdytos Monte Carlo simuliacijos, pateikiami generuotų modelių įverčiai bei empirinėmis simuliacijomis grįstos išvados.

4.1 Monte Carlo simuliacijos

Šiame skyrelyje nusakoma, kaip buvo sudaryta populiacija, imčių ėmimas ir svorių sudarymas.

Stengtasi, kad sukurta populiacija kuo labiau atspindėtų TIMSS populiaciją bei imtį Lietuvos atveju. Lietuvos mokyklų 2011 metais aštuntokų populiacija siekia ko ne 40 000, o TIMSS imtis apima iki 5000 aštuntokų (Lietuvos statistikos departamento duomenimis). Ji apima vos 12% visos aštuntokų populiacijos. Tokią procentinę dalį stengtasi atspindėti. Deja, didelė populiacija ir didelės imtys reikalauja ir daug kompiuterio resursų, todėl populiacijos dydis pasirinktas 300 mokyklų, o imamos imties - 35 mokyklos. Mokyklų parinkime lieka šiokia tokia problema dėl lizdų, šiame darbe lizdai nebuvo sudaryti. Taigi, iš pradžių sugeneruojama 300 mokyklų dydžio populiacija pagal paprastą hierarchinį modelį:

$$\begin{cases} Y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \varepsilon_{ij}, \varepsilon_{ij} \sim (0, \sigma^2); \\ \beta_{0j} = \gamma_{00} + \gamma_{01}W_j + u_{0j}, u_{0j} \sim (0, \tau_{00}); \\ \beta_{1j} = \gamma_{10}; \\ \beta_{2j} = \gamma_{20}. \end{cases} \quad (22)$$

Kiekvienoje iš mokyklų, sugeneruojamas jos dydis, panašiai kaip daryta Vakilian daktaro disertacijoje [49]. j - tosios mokyklos dydis gaunamas $N_j = [50 \times \exp(\tilde{u}_{0j})]$, čia laužtiniai skliaustai žymi sveikąją dalį, o \tilde{u}_{0j} sugeneruotas iš $\mathcal{N}(0; 0, 2)$ skirstinio apriboto tarp $-1, 5\sqrt{0, 2}$ ir $1, 5\sqrt{0, 2}$. Ir kiekvienai mokyklai sugeneruojamas aiškinamasis kintamasis W_j , kuris įgyja reikšmę 1 su tikimybe 0,2. Tai galėtų būti fiktyvus kintamasis, kuris parodo, ar tiriama mokykla yra Vilniaus mokykla. Galiausiai antrajame lygyje pridedami atsitiktiniai efektai u_{0j} . Čia išskiriami du atvejai, $u_{0j} \sim \mathcal{N}(0; 1300)$ arba $u_{0j} \sim \chi^2(0; 1300)$. Toks χ^2 pasiskirstymas gaunamas sugeneravus $\tilde{u}_{0j} \sim \chi_1^2$, jį standartizavus ir padauginus iš $\sqrt{1300}$. Čia pasirinkta $\tau_{00} = 1300$, $\gamma_{00} = 500$, $\gamma_{01} = 20$, $\gamma_{10} = 50$ ir $\gamma_{20} = -4$. Toliau kuriami mokinio (pirmojo) lygio kintamieji.

TIMSS tyrime parinktoje mokykloje parenkama klasė ir tiriami visi jos mokiniai. Čia turime per mažai mokinių, kad juos skaidyti į klases, todėl tiesiog sugeneruojami mokyklos mokiniai (tuo labiau, jog šiame darbe modeliuojamas tik mokyklos lygis). Tuomet $X_{1ij} \sim \mathcal{N}(90, 400)$ ir X_{2ij} parenkamas kaip pseudo kintamasis, kuris įgyja 1 reikšmę su tikimybe 0,5. Jis gali būti interpretuojamas kaip mokinio lytį žymintis kintamasis, kur 1 žymi mergaitę. Ir galiausiai pridedami atsitiktiniai svyravimai. Kaip ir anksčiau, išskiriami du atvejai: $\varepsilon_{ij} \sim \mathcal{N}(0; 2000)$ arba $\varepsilon_{ij} \sim \chi^2(0; 2000)$ (gautas kaip anksčiau). Vadinasi, iš viso, turimos keturios kombinacijos: abejos paklaidos pasiskirsčiusios pagal normalųjį skirstinį, pirmo lygio paklaidos pagal normalųjį ir antro pagal χ^2 , pirmo pagal χ^2 ir antro pagal normalųjį ir abejos paklaidos pagal χ^2 skirstinį.

Sudarius populiaciją, parenkama imtis. Kaip jau minėta iš 300 mokyklų populiacijos atsitiktinai parenkamos 35 mokyklos su tikimybėmis atvirkščiai proporcingomis mokyklos mokinių skaičiui. Tuomet nustatomi mokinių imties dydžiai kiekvienoje mokykloje n_j . Jei parinktoje mokykloje yra 30 ir mažiau mokinių, imami visi mokiniai. Jei tarp 30 ir 60 mokinių, tuomet parenkama pusė mokinių. O jei mokykla didelė (daugiau nei 60 mokinių), imamas trečdalis. Gautas skaičius suapvalinamas iki sveikosios dalies. Nustačius mokinių imties dydį, tos mokyklos mokiniai skeliami į dvi grupes pagal normaliasias paklaidas, t.y. imamos $\varepsilon_{ij} \sim \mathcal{N}(0, 2000)$ sugeneruotos sudarant populiaciją ir žiūrima, ar $\varepsilon_{ij} \leq 0$, ar atvirkščiai. Pirmu atveju, mokiniai papuola į pirmą grupę,

kitu, į antrą. Iš pirmosios grupės su lygiomis tikimybėmis $\left(\frac{1}{N_{1j}}\right)$ parenkama $n_{1j} = \left[\frac{3}{5}n_j\right]$ ir iš antrosios $n_{2j} = \left[\frac{2}{5}n_j\right]$ taip, kad $n_{1j} + n_{2j} = n_j$. Jei pagal paklaidas sudarytų grupių neužtenka parinkti imtims, parenkama n_j mokinių bendrai iš visos mokyklos. Taip iš dalies imituojama TIMSS klasių struktūra, tikimybės bei svoriai. Mokyklų svoriai $w_j = \frac{\sum_{i=1}^J N_{ij}}{N_j \times m}$, čia m - parinktų mokyklų skaičius (šiuo atveju $m = 35$). Mokinio svoriai sudaromi atitinkamai pagal imties sudarymą. Jei mokins yra parinktoje j - tojoje mokykloje ir papuola pagal paklaidas į pirmą grupę, tai jo svoris $w_{i|j} = \frac{N_{1j}}{n_{1j}}$. Panašiai gaunami svoriai ir kitais atvejais.

Šiame darbe atliekama 200 Monte Carlo simuliacijų. Kiekvienos iš jų metu sugeneruojama populiacija aukščiau aprašytu būdu, tuomet traukiama imtis pagal anksčiau aprašytą metodą. Ir vertinamas modelis (22) visais metodais (REML, REML su saviranka, MINQUE su saviranka). Gauti parametrų įverčiai, jų standartinės paklaidos bei pasikliautiniai intervalai analizuojami toliau pateiktais būdais.

Santykinis poslinkis (*angl. Relative Bias*) skaičiuojamas $\frac{\hat{\theta}_r - \theta}{\theta}$, kur θ - tikroji reikšmė, o $\hat{\theta}_r$ - r - tosios simuliacijos θ įvertis. Šis indeksas dažnai žymimas *RBIAS*. Jei šis indeksas yra mažesnis nei 5%, tai sakome, jog poslinkio nėra. Jei tarp 5% ir 20%, tai sakome, jog turime vidutinį poslinkį. Ir jei daugiau nei 20%, tai sakome, jog poslinkis didelis.

Santykinė vidutinė kvadratinė paklaida (*angl. Relative Mean Squared Error*) skaičiuojama $RMSE = \frac{1}{R} \left(\sum_{r=1}^R \frac{(\hat{\theta}_r - \theta)^2}{\theta} \right)^{\frac{1}{2}}$, kur R - Monte Carlo simuliacijų skaičius, $\hat{\theta}_r$ - r - tasis simuliacijos įvertis.

Pasikliautinių intervalų padengiamumas (*angl. Coverage*) skaičiuojamas labai paprastai. Tiesiog imamas procentas, kiek kartų tikroji reikšmė papuolė į simuliacijų metu gautus pasikliautinuosius intervalus.

Saviranka reikalauja daug kompiuterio resursų, ypač esant didesniai duomenų skaičiui, todėl buvo pasirinkta $B = 500$. Net turint tik 200 Monte Carlo simuliacijų ir 500 savirankos iteracijų, modelių vertinimas kiekvienai iš galimų paklaidų kombinacijų ir kiekvienam metodui užtruko apie tris savaites (naudotas keturių branduolių procesorius *Intel CORE i7*). Gauti rezultatai pateikti kitame skyrelyje.

4.2 Simuliacijų rezultatai

Šiame skyrelyje pateikiami anksčiau aprašytų Monte Carlo simuliacijų rezultatai. Žemiau esančiose lentelėse pateikti simuliacijų rezultatai. Įverčiai pateikti visiems trimis metodams. REML reiškia, jog pateiktas įvertinys gautas vertinant paprastu REML metodu; bREML - REML su saviranka; bMINQUE - MINQUE su saviranka. Gauti rezultatai yra kitokie nei tikėtasi. Šios simuliacijos atveju, kitaip nei Delpish, parametrų įverčiai praktiškai sutampa, kai abiejų lygių paklaidos nepriklausomai vienodai pasiskirsčiusios pagal normalųjį skirstinį (lentelė 4.2). Tiesa, lentelėse pateikti skaičiai yra suapvalinti, o skirtumas tarp metodų yra labai mažas, todėl pateiktos reikšmės atrodo identiškos daugeliui atvejų.

Lentelėje 4.2 pateiktas pasikliautinių intervalų padengiamumas. Tai rodo, kiek kartų tikroji reikšmė pateko į gautą intervalą. Matyti, jog nėra ryškaus skirtumo tarp visų trijų metodų.

Įverčiai, kuomet turime skirtingai pasiskirsčiusias paklaidas per lygius, nėra pateikiami. Jų rezultatai nesiskiria nuo gautųjų su abiem χ^2 skirstinio paklaidomis. Rezultatai pateikti 4.2 ir 4.2 lentelėse. REML vertinimo metodas yra pakankamai atsparus netenkinamai normalumo prielaidai. Tačiau dispersijos komponentų standartinės paklaidos ir, tuo pačiu, pasikliautiniai intervalai yra paslinkti. Palyginus MINQUE su saviranka ir REML su saviranka, didelio skirtumo nematyti, tad negalime teigti, jog MINQUE su saviranka yra geresnis metodas nei REML. Tai prieštarauja Delphish gautiems rezultatams. To priežastys gali būti:

- Delphish neatsižvelgė į tai, jog I-MINQUE ir REML metodai sutampa, kuomet sutampa *a priori* reikšmės. Šios autorės daktaro disertacijoje netiriamas REML metodas su saviranka.
- Šiame magistro darbe pateikiamas išvestas MINQUE metodas HLM modeliams, kuris nebūtinai sutampa su Delphish ar Bagakas nagrinėtu vertinimo metodu.
- Pasirinktas savirankos metodas nėra tinkamas nagrinėjamai simuliacijai. Viena iš priežasčių gali būti - per mažas iteracijų skaičius.

Kadangi empirinių simuliacijų būdu buvo parodyta, jog REML su saviranka nėra akivaizdžiai blogesnis už MINQUE su saviranka, Lietuvos TIMSS duomenų modelis bus kuriamas REML su savirankos būdu gautais pasikliautiniais intervalais ir standartinėmis paklaidomis.

| Fiksuoti efektai | | | | | | | | | | |
|-----------------------|---------------------|--------|-------|------|------|--------------------|--------|-------|------|------|
| Metodas | $\gamma_{00} = 500$ | | | | | $\gamma_{01} = 20$ | | | | |
| | MBias | SDBias | Min | Max | RMSE | MBias | SDBias | Min | Max | RMSE |
| REML | -0.01 | 0.02 | -0.07 | 0.03 | 0.02 | 0.06 | 0.88 | -1.88 | 2.57 | 0.88 |
| bREML | -0.01 | 0.02 | -0.07 | 0.03 | 0.02 | 0.06 | 0.88 | -1.88 | 2.57 | 0.88 |
| bMINQUE | -0.01 | 0.02 | -0.06 | 0.10 | 0.03 | 0.06 | 0.88 | -1.88 | 2.55 | 0.88 |
| Metodas | $\gamma_{10} = 50$ | | | | | $\gamma_{20} = -4$ | | | | |
| | MBias | SDBias | Min | Max | RMSE | MBias | SDBias | Min | Max | RMSE |
| REML | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.06 | 0.77 | -2.30 | 2.17 | 0.77 |
| bREML | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.06 | 0.77 | -2.30 | 2.17 | 0.77 |
| bMINQUE | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.06 | 0.77 | -2.30 | 2.17 | 0.77 |
| Atsitiktiniai efektai | | | | | | | | | | |
| Metodas | $\sigma^2 = 2000$ | | | | | $\tau_{00} = 1300$ | | | | |
| | MBias | SDBias | Min | Max | RMSE | MBias | SDBias | Min | Max | RMSE |
| REML | 0.01 | 0.23 | -0.60 | 0.62 | 0.23 | -0.00 | 0.05 | -0.11 | 0.13 | 0.05 |
| bREML | 0.01 | 0.23 | -0.60 | 0.62 | 0.23 | -0.00 | 0.05 | -0.11 | 0.13 | 0.05 |
| bMINQUE | 0.01 | 0.23 | -0.60 | 0.62 | 0.23 | -0.00 | 0.05 | -0.11 | 0.13 | 0.05 |

1 lentelė: Lentelėje pateiktos santykinio poslinkio statistikos (MBias - santykinio poslinkio vidurkis, SDBias - santykinio poslinkio standartinis nuokrypis, Min ir Max- atitinkamai mažiausia ir didžiausia įgyta santykinio poslinkio reikšmė, RMSE - santykinė vidutinė kvadratinė paklaida) visiems metodams, kai abiejų lygių paklaidos normaliosios.

| Parametras | REML | bREML | bMINQUE |
|---------------|-------|-------|---------|
| γ_{00} | 0.938 | 0.942 | 0.943 |
| γ_{01} | 0.931 | 0.938 | 0.937 |
| γ_{10} | 0.950 | 0.953 | 0.952 |
| γ_{20} | 0.932 | 0.941 | 0.941 |
| σ^2 | 0.949 | 0.951 | 0.948 |
| τ_{00} | 0.944 | 0.954 | 0.945 |

2 lentelė: Lentelėje pateikiamas pasikliautinių intervalų padengiamumas, kai vertinami duomenys pasiskirstę pagal normalųjį skirstinį (t.y. pirmo ir antro lygio paklaidos normaliosios.), tikroji reikšmė 0,95.

| Fiksuoti efektai | | | | | | | | | | |
|-----------------------|---------------------|--------|-------|------|------|--------------------|--------|-------|------|------|
| Metodas | $\gamma_{00} = 500$ | | | | | $\gamma_{01} = 20$ | | | | |
| | MBias | SDBias | Min | Max | RMSE | MBias | SDBias | Min | Max | RMSE |
| REML | -0.01 | 0.02 | -0.06 | 0.06 | 0.03 | 0.07 | 0.82 | -1.54 | 2.71 | 0.82 |
| bREML | -0.01 | 0.02 | -0.06 | 0.06 | 0.03 | 0.07 | 0.82 | -1.54 | 2.70 | 0.82 |
| bMINQUE | -0.01 | 0.02 | -0.06 | 0.07 | 0.03 | 0.07 | 0.82 | -1.54 | 2.72 | 0.82 |
| Metodas | $\gamma_{10} = 50$ | | | | | $\gamma_{20} = -4$ | | | | |
| | MBias | SDBias | Min | Max | RMSE | MBias | SDBias | Min | Max | RMSE |
| REML | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | -0.05 | 0.73 | -1.89 | 2.11 | 0.73 |
| bREML | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | -0.05 | 0.73 | -1.88 | 2.12 | 0.73 |
| bMINQUE | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | -0.05 | 0.73 | -1.89 | 2.11 | 0.73 |
| Atsitiktiniai efektai | | | | | | | | | | |
| Metodas | $\tau_{00} = 1300$ | | | | | $\sigma^2 = 2000$ | | | | |
| | MBias | SDBias | Min | Max | RMSE | MBias | SDBias | Min | Max | RMSE |
| REML | -0.04 | 0.31 | -0.63 | 0.98 | 0.31 | -0.00 | 0.05 | -0.10 | 0.14 | 0.05 |
| bREML | -0.04 | 0.31 | -0.63 | 0.99 | 0.32 | -0.00 | 0.05 | -0.10 | 0.14 | 0.05 |
| bMINQUE | -0.04 | 0.31 | -0.63 | 0.98 | 0.31 | -0.00 | 0.05 | -0.10 | 0.14 | 0.05 |

3 lentelė: Lentelėje pateiktos santykinio poslinkio statistikos (MBias - santykinio poslinkio vidurkis, SDBias - santykinio poslinkio standartinis nuokrypis, Min ir Max- atitinkamai mažiausia ir didžiausia įgyta santykinio poslinkio reikšmė, RMSE - santykinė vidutinė kvadratinė paklaida) visiems metodams, kai abiejų lygių paklaidos pasiskirsčiusios pagal χ^2 skirstinį.

| Parametras | REML | bREML | bMINQUE |
|---------------|-------|-------|---------|
| γ_{00} | 0.928 | 0.943 | 0.942 |
| γ_{01} | 0.925 | 0.936 | 0.938 |
| γ_{10} | 0.944 | 0.954 | 0.951 |
| γ_{20} | 0.912 | 0.923 | 0.926 |
| σ^2 | 0.920 | 0.925 | 0.931 |
| τ_{00} | 0.567 | 0.921 | 0.915 |

4 lentelė: Lentelėje pateikiamas pasikliautinių intervalų padengiamumas, kai vertinami duomenys pasiskirstę pagal χ^2 skirstinį (t.y. pirmo ir antro lygio paklaidos iš χ^2).

5 HLM MODELIS TIMSS DUOMENIMS

Šiame skyriuje trumpai aprašomas TIMSS tyrimas bei hierarchinių tiesinių modelių TIMSS duomenims apžvalga. Atsižvelgiant į 4 skyriaus rezultatus sudaromas ir įvertinamas modelis Lietuvos TIMSS duomenims.

5.1 TIMSS tyrimo aprašymas

Tyrimas TIMSS (*angl. Trends in International Mathematics and Science Study*) – tai tarptautinis matematikos ir gamtos mokslų gebėjimų tyrimas. Šį tyrimą inicijuoja IEA (*angl. International Association of the Evaluation of Educational Achievement*) – Tarptautinė švietimo pasiekimų vertinimo asociacija. Tyrime dalyvauja apie 70 pasaulio šalių. IEA vykdo ir kitus švietimo tyrimus: PIRLS, ICCS ir SITES, TEDS-M ir ICILS. TIMSS tyrimas pradėtas 1995 metais ir tęsiamas kas 4 metus. Tiriami ketvirtokų ir aštuntokų matematikos ir gamtos mokslų gebėjimai. Remiamasi tuo, jog tai yra paskutinės pradinės ir pagrindinės mokyklų pakopos. Šiuo metu jau yra įvykdyti 5 tokie tyrimai, paskutinis vyko 2011 metais. Lietuva dalyvavo visuose etapuose. Šiame darbe pasirinkti nagrinėti 2011 metų tyrimo aštuntos klasės Lietuvos duomenys. (Pagal [32])

TIMSS tyrimas vykdomas šalies matematiniame raštingumui bei jį lemiantiems veiksniams nustatyti. Sudaromi 28 testo užduočių blokai. Vienam mokiniui skiriamas vienas testo sąsiuvinis, kurį sudaro 4 blokai, po du matematikos ir gamtos mokslų. Sąsiuviniai atrinktiems mokiniams paskirstomi pagal iš anksto numatytą tvarką (neatsitiktinai). Mokinys testą turi atlikti per tam tikrą laiką ir gali nespėti atlikti visų užduočių. Norėdama įvertinti bendrą rezultatą, TIMSS komanda sudaro matematinio testo rezultatus pagal moderniąją testų teoriją, apie tai plačiau galima rasti 1995 metų TIMSS vartotojo vadovo [20] 5 skyriuje. Kad rezultatai būtų palyginami, jie yra pernuoformuojami taip, kad bendras visų tiriamųjų rezultatų vidurkis būtų 500, o standartinis nuokrypis 100. Kartu yra renkama informacija apie mokinį, jo klasę ir mokyklą. Kiekvienas lygmuo gauna po klausimyną, kurį užpildo. (Pagal [32])

TIMSS tyrimas vykdomas ne visai populiariai, o atrinktai jos daliai (imčiai). TIMSS imtys yra dviejų pakopų sluoksninės lizdinės imtys, kur pirmos pakopos elementai išrenkami su tikimybėmis proporcingomis dydžiui, o antros pakopos elementai išrenkami su lygiomis tikimybėmis. Pirmoje pakopoje išrenkamos mokyklos pagal aštuntokų skaičių mokykloje. Antroje pakopoje su lygiomis tikimybėmis parenkamos klasės. Išrinktoje klasėje apklausiami visi jos mokiniai. Stengiamasi imtį sudaryti taip, kad ji kuo labiau atspindėtų populiaciją. Daugiau apie tai TIMSS internetinėje nuorodoje [43]. Dėl nelygių patekimo į imtį tikimybių sudaromi kiekvieno mokinio bei mokyklos svoriai, kurie vėliau naudojami tyrimuose.

Kiekviena šalis sudaro imtį pagal savo poreikius ir TIMSS standartus. Lietuvoje mokyklų lizdai sudaromi pagal mokyklos vietovę (t.y. didmiestis, miestelis ar kaimas ir pan.), o paskui skirstoma pagal mokyklos tipą (t.y. pagrindinė, vidurinė ar gimnazija). Tuomet iš kiekvieno lizdo imama negrąžintinė imtis su tikimybėmis proporcingomis aštuntokų skaičiui mokykloje. Tegu turime H lizdų. Tuomet mokyklos svoriai h -tajame lizde ($h = 1, \dots, H$) gaunami pagal:

$$W_{hj}^{sc} = \frac{M_h}{m_{hj}n_h}; M_h = \sum_{j=1}^{N_h} m_{hj}, \quad (23)$$

kur M_h – bendras aštuntokų skaičius h -tajame lizde, m_{hj} – aštuntokų skaičius j -tojoje h -tojo lizdo mokykloje, n_h – mokyklų, parinktų iš h -tojo lizdo skaičius ir N_h – mokyklų skaičius h -tajame lizde. Iš kiekvienos parinktos mokyklos, atitinkamai pagal jos dydį, parenkamos klasės su lygiomis

tikimybėmis. Klasės svoriai gaunami:

$$W_{k|hj}^{cl} = \frac{C_{hj}}{c_{hj}}, \quad (24)$$

čia C_{hj} - klasių skaičius j -tojoje mokykloje, c_{hj} - parinktų klasių skaičius j -tojoje mokykloje. Parinktose klasėse apklausiami visi mokiniai, todėl sąlyginis svoris yra $W_{i|hjk}^{st} = 1$. Galiausiai sudaromas bendras kiekvieno parinkto mokinio svoris:

$$W_{ikjh}^{st} = W_{hj}^{sc} \times W_{k|hj}^{cl} \times W_{i|hjk}^{st}. \quad (25)$$

Čia pateikti svoriai nėra visiškai tikslūs, nes dar taikomi nedalyvavimo korekcijos faktoriai. Tačiau apie tai plačiau galima rasti [43].

5.2 TIMSS tyrimų apžvalga

TIMSS duomenys labai plačiai naudojami lyginamajai (bei kitoms) analizei atlikti. Kiekviena TIMSS tyrimo šalis turi instituciją, kuri atsakinga už ataskaitų paruošimą. Lietuvoje tai atlieka *Nacionalinis egzaminų centras (NEC)*. Taip pat ir IEA organizacija išleidžia keletą leidinių, kuriuose galima rasti pagrindines ir platesnes TIMSS rezultatų statistikas bei išvadas. Visa informacija patalpinta oficialiame TIMSS tinklapyje (<http://timss.bc.edu>).

Dažniausiai HLM modeliai TIMSS duomenims sudaromi, nustatyti mokyklos, mokytojo ar klasės bei mokinio aplinkos įtaką mokinio pasiekimams (pvz.: Fullarton [19], Chen [5] arba Afana ir Lietz [1]). Taip pat sudaromi modeliai kelioms šalims ar regionams palyginti (pvz.: Gustafsson [21], Kim ir kitų [26], Phan ir kitų [35] arba Wiberb ir Andersson [52]). Buvo atlikta analizė net šalių lygio mastu (pvz.: Kyriakides ir Charalambous [25]). Analizuota šalių politikos įtaka mokinio matematikos pasiekimams (Stidd [45]).

HLM modeliai Lietuvos duomenims nebuvo labai plačiai kurti. Dažniausiai Lietuvos TIMSS duomenys modeliuoti bendrai su kitomis šalimis ar regionais kaip Akyuz ir Berberoglu [2]. Autoriai sudarė modelį 1999 metų TIMSS tyrimo duomenims. Jų tikslas buvo palyginti Turkijos ir Europos sąjungos šalių mokytojo ir klasės įtaką mokinio matematinio raštingumo rezultatams. Buvo sudarytas bendras modelis visoms jų nagrinėjamoms šalims su vienu mokinio lygio aiškinančiuoju kintamuoju, tačiau į modelį buvo įtrauktas ir atsitiktinis posvyris. Deja, Lietuvos atveju atsitiktinis posvyris nebuvo įtrauktas. Žemiau pateikta minimo modelio išraiška:

$$\begin{cases} Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + \varepsilon_{ij}; \\ \beta_{0j} = \gamma_{00} + \gamma_{01} W_{1j} + \dots + \gamma_{0m} W_{mj} + u_{0j}; \\ \beta_{1j} = \gamma_{10} + u_{1j}; \end{cases}$$

čia i - žymi i -tąjį mokinį,

j - j -tąją mokyklą/klasę/mokytoją,

Y_{ij} - i -tojo mokinio iš j -tosios mokyklos/klasės/mokytojo matematinio raštingumo rezultatas,

β_{0j} - j -tosios mokyklos/klasės/mokytojo įnašas,

β_{1j} - atsitiktinis posvyris,

X_{ij} - pirmo lygio kintamasis, šiuo atveju, namų edukaciniai ištekliai,

ε_{ij} - pirmo lygio paklaidos,

γ_{00} - vidutinis mokyklų/klasių/ mokytojų įnašas,

γ_{01} - vidutinis mokyklų/klasių/ mokytojų nuolydis,

W_{mj} - antro lygio kintamasis,

u_{0j} - atsitiktinis efektas,

u_{1j} - atsitiktinis efektas.

Šis modelis vertintas su *HLM6* programa visoms 5 galimoms matematinio raštingumo reikšmėms ir su imties svoriais.

Kitas pavyzdys, modelis sudarytas Robitaille ir Beaton [42] 1995 metų TIMSS duomenims. Autoriai sudarė net trijų lygių modelį su atsitiktiniais postūmiais. Kaip antro ir trečio lygio kintamieji buvo naudoti suvidurkinti (atitinkamai pagal klases ir pagal mokyklas) mokinio lygio kintamieji. Čia vietoje anksčiau minėto modelio namų mokymosi išteklių imti atskiri kategoriniai kintamieji - tėvų išsilavinimas, mokinio turimi mokymosi reikmenys ir jo užklasinė veikla. Tačiau atsitiktiniai posvyriai neįtraukti, be to, nagrinėta visa šalių grupė, o ne atskiros šalys. Modeliai vertinti taip pat su *HLM6* programa visoms 5 galimoms matematinio raštingumo reikšmėms ir su imties svoriais.

Lietuvoje Lietuvos TIMSS duomenims sudarytų HLM modelių beveik nerasta. Galbūt dėl to, jog tokie modeliai sudaromi studijų baigiamuosiuose darbuose, kurie nėra labai viešinami. Vienintelis darbas, kuriame sudaromi HLM modeliai Lietuvos TIMSS duomenims, kurį pavyko gauti, tai Dudaitės daktaro laipsnio disertacija [13]. Savo darbe Dudaitė plačiai nagrinėja 1995, 1999 ir 2003 metų TIMSS tyrimo duomenis įvairiais aspektais, tačiau pagrindinis tikslas - nustatyti matematinio raštingumo kaitą, keičiantis Lietuvos mokymo aplinkai. HLM modeliai sudaryti 2003 metų duomenims ir tik su labai mažu kintamųjų kiekiu. Galima sakyti, jog beveik kiekvienam aiškinančiajam kintamajam sudarytas atskiras modelis. Modelio pavyzdys:

$$\begin{cases} Y_{ij} = \beta_{0j} + \varepsilon_{ij}; \\ \beta_{0j} = \gamma_{00} + \sum_{q=1}^Q \gamma_{0q} W_{qj} + u_{0j}; \end{cases} \text{ arba } \begin{cases} Y_{ij} = \beta_{0j} + \sum_{p=1}^P \beta_{pj} X_{pij} + \varepsilon_{ij}; \\ \beta_{0j} = \gamma_{00} + u_{0j}; \\ \beta_{pj} = \gamma_{p0}; \quad \forall q = 1, \dots, P. \end{cases}$$

čia i - žymi i -tąjį mokinį,

j - j -tąją mokyklą/klasę/mokytoją,

Y_{ij} - i -tojo mokinio iš j -tosios mokyklos/klasės/mokytojo matematinio raštingumo rezultatas,

β_{0j} - j -tosios mokyklos/klasės/mokytojo įnašas,

ε_{ij} - pirmo lygio paklaidos,

γ_{00} - vidutinis mokyklų/klasių/ mokytojų įnašas,

γ_{01} - vidutinis mokyklų/klasių/ mokytojų posvyris,

X_{pij} - p - tasis i - tojo mokinio aiškinamasis kintamasis,

W_{qj} - antro lygio kintamasis,

u_{0j} - antro lygio paklaidos.

Pasak autorės, platesnių modelių sudaryti nepavyko. Šio magistro darbo autorė išvelgia dvi galimas priežastis. Pirmoji - modeliams vertinti buvo naudotasi specialiu HLM modelių vertinimo paketu *HLM6*, ši programa mokama, tačiau akademinei visuomenei suteikiama nemokama versija, kuri vertina tik labai mažos apimties modelius. Antra - TIMSS duomenys (aiškinantieji kintamieji) yra labai koreliuoti, pavyzdžiui mokyklos vietovė ir mokyklos mokinių finansinė padėtis arba mokyklos dydis. Todėl norint pažiūrėti kiekvieno iš tokių kintamųjų įtaką, juos reiktų modeliuoti atskirai.

5.3 TIMSS duomenys

Šiame skyrelyje pateikti modeliavime naudoti duomenys. Magistro darbe nagrinėjami 2011 metų Lietuvos duomenys, kurie pateikti oficialiame TIMSS tinklapyje (<http://timss.bc.edu>). Pavadinimai: BCGLTUM5 - mokyklos failas, BSGLTUM5 - mokinių failas, BSTLTUM5 - mokytojų ir mokinių ryšių failas ir BTMLTUM5 - matematikos mokytojų failas. Nors pateiktos visos 5

galimos mokinio testo atlikimo reikšmės, toliau analizėje bus naudojama tik pirmoji - BSMMAT01. Lentelėje 5 pateiktas originalus kintamojo pavadinimas, jo aprašymas bei pavadinimas naudojamas šiame darbe. Kai kurie kategoriniai kintamieji per ranguoti, o indeksai centruoti mokyklos vidurkiu bei suvidurkinti kiekvienai mokyklai. Simbolis "*" reiškia, jog kintamasis modifikuotas ir nuo originalaus skiriasi. Atliekant vidurkinimą naudoti svoriai. Mokyklos, klasės ir mokiniai identifikuojami pagal IDSCHOOL, IDCLASS ir IDSTUD atitinkamai. Mokyklos, klasės ir mokinio svoriai pateikti kaip WGTFAC3, WGTFAC2 ir WGTFAC1. Taip pat pateikiami ir neatsakymo į klausymus koregavimo faktoriai WGTADJ3, WGTADJ2 ir WGTADJ1. Toliau naudojamų pavadinimų pradžioje pridėta raidė C reiškia, jog kintamasis centruotas pagal mokyklą, o M - mokyklos vidurkis. Centruojant buvo naudoti imties svoriai.

| Trumpinys | Aprašymas | Tipas/Skalė | Originalus pavadinimas |
|---------------------------|---|--|------------------------|
| Mokinio lygio kintamieji | | | |
| MatRes | Mokinio matematinio raštingumo testo rezultatas | Tolydus kintamasis, originalus | BSMMAT01 |
| SSEX* | Mokinio lytis | 0 - Mergaitė; 1 - Berniukas. | BSBG01 |
| STHMWRK* | Mokinio skiriamas laikas namų darbams | 0 - 45 min ir mažiau; 1 - daugiau nei 45 min. | BSDMWKHW |
| SMENG | Mokinio išitraukimas į matematikos pamokas | Indeksas. Didesnė reikšmė - stipresnis išitraukimas. | BSBGEML |
| SMCONF | Mokinio pasitikėjimas savimi atliekant matematikos užduotis | Indeksas. Didesnė reikšmė - stipresnis pasitikėjimas. | BSBGSCM |
| SMVALUE | Mokinys vertina matematikos žinias (pvz.: jos jam padės gauti trokštamą darbą) | Indeksas. Didesnė reikšmė - stipresnis vertinimas. | BSBG SVM |
| SMLIKE | Mokiniui patinka mokytis matematiką | Indeksas. Didesnė reikšmė - mėgsta labiau. | BSBG SLM |
| SBULLED | Mokinys yra engiamas mokykloje | Indeksas. Didesnė reikšmė - skriaudžiamas dažniau. | BSBG SBS |
| SHEDRES | Mokinio namų mokymosi išteklių. Anksčiau vadinta namų socioekonominiu statusu. Įeina tėvų išsilavinimas, ar mokinys turi kompiuterį ir atskirą kambarį. | Indeksas. Didesnė reikšmė - daugiau išteklių. | BSBGHER |
| Mokyklos lygio kintamieji | | | |
| MDYDIS | Mokyklos mokinių skaičius | Natūralusis skaičius. | BCBG01 |
| MSUCCESS | Mokytojų teigimu mokykla skiria dėmesį akademinei sėkmei. | Indeksas. Didesnė reikšmė - labiau tiki. | BCBG EAS |
| MSUDET* MSUDET.1-3 | Mokyklos mokinių sudėtis pagal finansinę padėtį | 1 - Daugiau vargšų; 2 - Apylygiai; 3 - Daugiau turtingų. | BCDG03 |

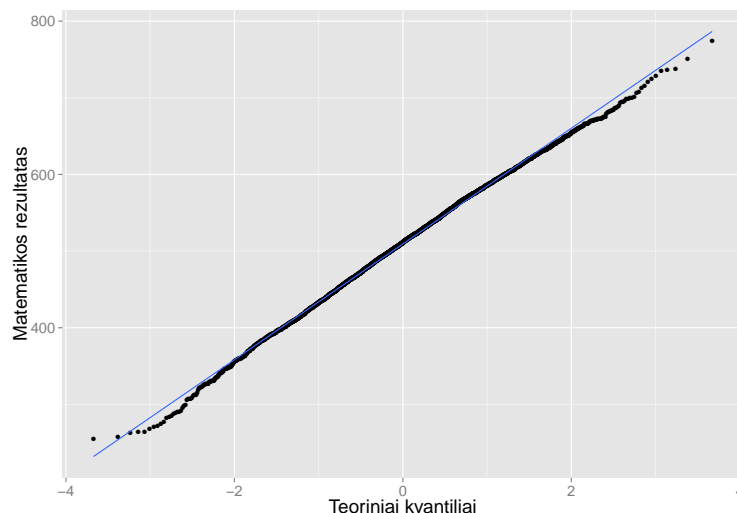
| | | | |
|-------------------------------|---|--|----------|
| MDISSAFE | Saugi ir disciplinuota mokykla. | Indeksas. Didesnė reikšmė - daugiau saugumo ir disciplinos. | BCBGDAS |
| MINSTRHWR | Valandos, skiriamos instruktažui mokykloje per metus. | Valandos. | BCDG06HY |
| MMATSHORT | Mokytojai skundžiasi mokyklos išteklių trūkumu. | Indeksas. Didesnė reikšmė - daugiau skundžiasi. | BCBGMRS |
| MAINCOME* MAINCOME.1 -3 | Mokyklos aplinkos pajamų dydis. | 1 - Žemas; 2 - Vidutinis; 3 - Aukštas. | BCBG05C |
| MVIETA* MVIETA.1-5 | Mokyklos vietovė. | 1 - Atokesnis kaimas; 2 - Mažas miestelis; 3 - Vidutinis miestas; 4 - Didmiestis; 5 - Vilnius, | BCBG05B |
| MKITKALB* MKITKALB.1 -3 | Kitakalbių dalis mokykloje. | 1 - 10% ir mažiau; 2 - 10% - 50%; 3 - daugiau nei 50%. | BCBG05C |

5 lentelė: Magistro darbe naudotų TIMSS duomenų detalizavimas

5.4 HLM modelis 2011 metų TIMSS duomenims

Šiame skyrelyje pateikiama detalesnė turimų ir nagrinėjamų TIMSS duomenų analizė, sudaromas nulinis modelis bei pateikiamas galutinis modelis ir jo statistikos. Nulinio ir galutinio modelio įverčiai pateikiami įvertinti visais nagrinėjamais metodais. Galiausiai aptariami rezultatai ir padaromos išvados.

Pirmiausia, domina aiškinamasis kintamasis - matematinio raštingumo testo rezultatai. Kaip jau buvo minėta, TIMSS pateikia 5 galimas testo atlikimo reikšmes, tačiau šiame darbe naudojama tik viena - pirmoji. Taip daroma, dėl skaičiavimų sudėtingumo (tai nėra implementuota R statistinės analizės pakete) ir dėl to, jog vėliau rezultatus galima pritaikyti ir likusioms galimoms testo atlikimo reikšmėms. Lieka klausimas, ar galime laikyti aiškinamąjį kintamąjį pasiskirsčiusiu pagal normalųjį skirstinį. Paveikslėlyje 1 pateiktas pirmosios galimos matematinio raštingumo testo atlikimo reikšmės teorinių kvantilių grafikas. Jame galime įžvelgti, jog turime sunkias uodegas. Aiškinančiajam kintamajam buvo atliktas Shapiro-Wilk normalumo testas. gauta testo statistika $W = 0.9977$ ir $p\text{-value} < 0,01$. Todėl hipotezė apie aiškinamojo kintamojo pasiskirstymą pagal normalųjį skirstinį atmetama.



1 pav.: Mokinio matematinio raštingumo pirmos galimos testo atlikimo reikšmės teorinių kvantilių grafikas.

Aiškinantieji kintamieji šiame darbe negali būti laikomi pasiskirstę pagal normalųjį skirstinį, tačiau kategoriniai kintamieji nėra modeliuojami. Dudaitė savo daktaro disertacijoje [13] ne tik sudarinėjo labai menkus modelius, bet ir naudojo kategorinius kintamuosius (kurie įgyja iki 5 reikšmių) kaip regresorius. Šiame darbe kategoriniai kintamieji išskaidyti į fiktyvius kintamuosius. Tokių yra du mokinio lygyje ir keturi mokyklos lygyje. Pačiuose pateiktuose TIMSS duomenyse kategoriniai kintamieji suvesti į indeksus, kurie įgyja daugiau nei 5 reikšmes. Nors jie ir negali būti laikomi pasiskirstę pagal normalųjį dėsnį, modeliuojant nėra stipriai nusižengiama. Šeši indeksai yra imami kaip mokinio lygio aiškinantieji kintamieji ir 3 mokyklos lygio kintamieji. Detalesni kintamųjų aprašymai pateikti 5 lentelėje 5.3 skyrelyje. Modeliuoti ir suvidurkinti mokytojo bei klasės kintamieji, tačiau jie nepasirodė statistiškai reikšmingi, todėl apie šiuos kintamuosius šiame darbe nepasakojama.

HLM modelio sudarymas yra sudėtingas procesas. Kad modeliai būtų palyginami, jie turi būti sudaryti tiems patiems duomenims. Čia susiduriama su problema, jog TIMSS duomenys turi daug neatsakytų reikšmių, t.y. trūkstamų duomenų. Modeliuojant atsižvelgiama ne tik į tai, kiek kintamasis paaiškina dispersijos, bet ir į tai, ar jis neturi daug praleistų reikšmių. Šiame darbe modeliai sudaryti įtraukiant po vieną pirmo po to antro lygio kintamąjį ir lyginant su nuliniu ir prieš tai gautu modeliais, kiekvieną kartą perskaičiuojant pašalinus mokinį ar mokyklą su trūkstamais duomenimis. Gavus galutinį modelį, visi modelių įverčiai buvo pervertinti jau galutiniam duomenų kiekiui. Taip išvengiama per didelio duomenų pašalinimo pirmame etape. Modeliai parinkti vadovaujantis logika ir 2.2 skyrelyje pateiktais kriterijais. Toliau bus pateikti tik nulinis ir galutinis modeliai, o tarpiniai modeliai praleisti.

Pagrindines 2011 metų TIMSS tyrimo suvestines galima rasti jau minėtoje Nacionalinio egzaminų cento ataskaitoje [32]. Todėl šiame darbe bus pateikti trumpi rezultatai po duomenų pašalinimo. Pašalinus mokyklas ir mokinius su trūkstamais duomenimis iš 141 tyrime dalyvavusios mokyklos liko 128 mokyklos, o iš 4747 mokinių liko 4167 mokiniai. Tai optimalus skaičius siekiant paaiškinti kuo daugiau kaitos su kuo mažesniu mokinių praradimu. Svertinis pirmos galimos matematinio raštingumo reikšmės vidurkis likusiems duomenims yra 504,57. Toliau pateikiama nulinio modelio išraiška ir REML su saviranka metodu gauti įverčiai.

Nulinis modelis aprašytas 2.2 skyrelyje sudarytas jau mažesnės apimties duomenims. Pakartota (8) modelio išraiška:

$$\begin{cases} MatRes_{ij} = \beta_{0j} + \varepsilon_{ij}, \varepsilon_{ij} \sim (0, \sigma^2); \\ \beta_{0j} = \gamma_{00} + u_{0j}, u_{0j} \sim (0, \tau_{00}). \end{cases} \quad (26)$$

Tegul Y_{ij} žymi i - tojo mokinio iš j - tosios mokyklos matematinio raštingumo testo rezultato pirmąją galimą reikšmę. Jungtinė modelio lygtis:

$$MatRes_{ij} = \gamma_{00} + u_{0j} + \varepsilon_{ij}.$$

Šis modelis įvertintas REML su saviranka metdu. Rezultatai pateikti lentelėje 5.4. Šio modelio $ICC \approx 0,2$. Tai rodo, jog apie 20% mokinių rezultatų skirtumų lemia mokyklos. Skyrelyje 2.2 minėta R_1^2 statistika yra apie 0,22.

| Parametras | Įvertis | St. paklaida | Pasikliautinis intervalas | | Ilgis |
|---------------|---------|--------------|---------------------------|---------|--------|
| γ_{00} | 504.7 | 3.15 | 498.32 | 511.07 | 12.75 |
| σ^2 | 4531.98 | 102.682 | 4339.86 | 4735.34 | 395.48 |
| τ_{00} | 1165.54 | 141.75 | 878.04 | 1540.50 | 662.46 |

6 lentelė: Nulinio modelio parametų įverčiai; REML su saviranka metodas.

Galutinis modelis sudarytas atsižvelgiant į kintamųjų multikoliniarumą, pasitelkiant logiką bei skyrelyje 2.2 pateiktais indeksus ir statistikas. Mokinio lygio kintamieji - indeksai SMCONF, SMLIKE, SMVALUE ir SMENG - sudaryti iš skirtingų mokinio klausimyno klausimų, tačiau tarpusavyje labai susiję. Į modelį įtraukus visus kintamuosius, koeficientų įverčiai prie šių kintamųjų tampa nebe paaiškinami. Atsižvelgiant į tai jog SMCONF paaiškino labai daug (apie 1000) pirmo lygio dispersijos, tik šis kintamasis buvo įtrauktas į galutinį modelį.

Panaši situacija ir su mokyklos lygio fiktyviais kintamaisiais. Tad į galutinį modelį įtraukti tik keli iš jų. Galutinio modelio išraiška pateikta toliau:

$$\begin{cases} MatRes_{ij} = \beta_{0j} + \beta_{1j} \times SSEX_{ij} + \beta_{2j} \times STHMWRK_{ij} + \beta_{3j} \times CSMCONF_{ij} + \\ \quad + \beta_{4j} \times CSBEDRES_{ij} + \beta_{5j} \times CSBULLED_{ij} + \varepsilon_{ij}; \\ \beta_{0j} = \gamma_{00} + \gamma_{01} \times MDYDIS_j + \gamma_{02} \times MSUCCESS_j \times MVIETA.4_j + \\ \quad + \gamma_{03} \times MSMLIKE_j \times MVIETA.4_j + \gamma_{05} \times MSBEDRES_j + \\ \quad + \gamma_{06} \times MAINCOME.1_j + \gamma_{07} \times MVIETA.4_j + u_{0j}; \\ \beta_{1j} = \gamma_{10}; \\ \beta_{2j} = \gamma_{21} \times MVIETA.5_j; \\ \beta_{3j} = \gamma_{30} + u_{3j}; \\ \beta_{4j} = \gamma_{40} + \gamma_{42} \times MVIETA.5_j + u_{4j}; \\ \beta_{5j} = \gamma_{51} \times MVIETA.3_j. \end{cases} \quad (27)$$

Jungtinė 27 modelio lygtis yra:

$$\begin{aligned} MatRes_{ij} = & \gamma_{00} + \gamma_{01} \times MDYDIS_j + \gamma_{02} \times MSUCCESS_j \times MVIETA.4_j + \\ & + \gamma_{03} \times MSMLIKE_j \times MVIETA.4_j + \\ & + \gamma_{05} \times MSBEDRES_j + \gamma_{06} \times MAINCOME.1_j + \gamma_{07} \times MVIETA.4_j + \\ & + \gamma_{10} \times SSEX_{ij} + \gamma_{21} \times MVIETA.5_j \times STHMWRK_{ij} + \gamma_{30} \times CSMCONF_{ij} + \\ & + \gamma_{40} \times CSBEDRES_{ij} + \\ & + \gamma_{42} \times MVIETA.5_j \times CSBEDRES_{ij} + \gamma_{51} \times MVIETA.3_j \times CSBULLED_{ij} + \\ & + [u_{0j} + u_{3j} \times CSMCONF_{ij} + u_{4j} \times CSBEDRES_{ij} + \varepsilon_{ij}] \end{aligned} \quad (28)$$

| Parametras | Įvertis | St. paklaida | Pasikliautinis intervalas | | Ilgis |
|---------------|---------|--------------|---------------------------|---------|--------|
| γ_{00} | 146.31 | 33.7 | 89.19 | 219.72 | 130.53 |
| γ_{01} | 0.01 | 0.007 | -0.01 | 0.02 | 0.03 |
| γ_{02} | 12.75 | 4.87 | 3.47 | 23.55 | 20.08 |
| γ_{03} | 26.69 | 8.52 | 9.31 | 43.00 | 33.69 |
| γ_{05} | 32.97 | 3.16 | 25.45 | 38.65 | 13.20 |
| γ_{06} | 62.26 | 21.60 | 17.04 | 110.18 | 93.14 |
| γ_{07} | -367.12 | 91.20 | -555.59 | -186.85 | 368.74 |
| γ_{10} | 10.93 | 1.86 | 7.41 | 14.32 | 6.91 |
| γ_{21} | 12.40 | 2.86 | 6.37 | 18.19 | 11.82 |
| γ_{30} | 18.73 | 0.48 | 17.82 | 19.71 | 1.89 |
| γ_{40} | 15.34 | 1.81 | 12.08 | 18.57 | 6.49 |
| γ_{42} | -3.65 | | | | |
| γ_{51} | -1.74 | | | | |
| σ^2 | | | | | |
| τ_{00} | | | | | |
| τ_{01} | | | | | |
| τ_{02} | | | | | |
| τ_{11} | | | | | |
| τ_{12} | | | | | |
| τ_{22} | | | | | |

7 lentelė: Nulinio modelio parametų įverčiai; REML su saviranka metodas.

6 IŠVADOS

Pagrindiniai šio magistro darbo rezultatai:

- Parašytos \mathbb{R} funkcijos MINQUE su saviranka metodui, kuris skirtas HLM modelių parametų bei jų pasikliautinių intervalų vertinimui.
- Atliktos empirinės simuliacijos, kurios parodė, jog Delphish tirtas MINQUE metodas su saviranka nėra geresnis nei REML metodas su saviranka. Tikslūs pasikliautiniai intervalai savirankos būdu gaunami ir turint duomenis pasiskirsčiusius pagal normalųjį skirstinį, ir pagal Puasono (ar bet koki kitą ???) skirstinį. To priežastys gali būti tai, jog Delphish netyrė REML su saviranka arba šiame darbe pristatytas MINQUE metodas nėra tapatus Bagakas ir Delphish metodui.
- Sudarytas sudėtingesnis hierarchinis tiesinis modelis (nei anksčiau buvę) Lietuvos matematinio raštingumo testo rezultatams.

Tolimesnis tyrimas galėtų apimti daugiau metodų fiksuotiems ir atsitiktiniams efektams vertinti (pvz.: Hamilton'o III metodas). Sudėtingesni neparimetrinės savirankos metodai taip pat galėtų pagerinti įvertinių tikslumą.

LITERATŪRA IR ŠALTINIAI

- [1] Afana, Y., Lietz, P., *The relationship between school resources and mathematics achievement at grade 8: A comparison of Israeli and Palestinian schools in TIMSS 2007*, Journal for Educational Research Online 5.1, 2013, 59-89.
- [2] Akyuz, G., Berberoglu, G., *Teacher and Classroom Characteristics and Their Relations to Mathematics Achievement of the Students in the TIMSS*, New Horizons in Education, 2010.
- [3] Campbell, J., Verna, M., Petruso, S., *Gender Paradigms*, IRC, 2004.
- [4] V. Čekanavičius, G. Murauskas, *Statistika ir jos taikymai*, 3 dalis, TEV, 2009.
- [5] Chen, Q., *A Multilevel Analysis of Mathematically Low-Achieving Students in Singapore*, IRC, 2013.
- [6] Bagakas, J. G., *Two level nested hierarchical linear model with random intercepts via the bootstrap*, Unpublished doctoral dissertation, Michigan State University, 1992.
- [7] Berkhof, J., Snijders, T. A., *Variance Component Testing in Multilevel Models*, Journal of Educational and Behavioral Statistics, Summer 2001, Vol. 26, Nr. 2, 133-152.
- [8] Birenbaum, M., Tatsuoka, C., Yamada, T., *Diagnostic Assessment in TIMSS-R: Comparison of Eight Graders Mathematics Knowledge States in The United States, Japan, Israel*, IRC, 2004.
- [9] Brečko, B., *How Family Background Influences Student Achievement*, IRC, 2004.
- [10] Davidian, M., and Giltinan, D., *Nonlinear Models for Repeated Measurement Data*, New York: Chapman & Hall/CRC, 1995.
- [11] Delpish, A., *Comparison of Estimators in Hierarchical Linear Modeling: Restricted Maximum Likelihood Versus Bootstrap via Minimum Norm Quadratic Unbiased Estimators*, Electronic Theses, Treatises and Dissertations. Paper 771, 2006.
- [12] Demidenko, E., *Mixed Models– Theory and Applications with R*, 2nd Edition, Wiley Series in Probability and Statistics, John Wiley & Sons, 2004.
- [13] J. Dudaitė, A. Eljio, Ž. Urbienė, A. Zabulionis, *Tarptautinis matematikos ir gamtos mokslų tyrimas. TIMSS 2003. Ataskaita*, NEC, 2004.
- [14] J. Dudaitė, *Tarptautinis matematikos ir gamtos mokslų tyrimas TIMSS 2003. Rezultatų Analizė, matematika, viii klasė*. Nacionalinis egzaminų centras, 2004.
- [15] J. Dudaitė, A. Eljio, *Change in Lithuanian Basic School Students' Mathematics Achievement through 1995-2003*, Tarptautinė matematikos dėstymo konferencija – 2005, 57-62.
- [16] Efron, B., *Bootstrap Methods: Another Look at the Jackknife*, The Annals of Statistics 7 (1): 1–26, 1979.
- [17] A. Eljio, J. Dudaitė, I. Mackevičienė, V. Trublenkovaitė, *Tarptautinis matematikos ir gamtos mokslų tyrimas TIMSS 2007. Ataskaita. Matematika*, NEC, 2008.
- [18] A. Eljio, J. Dudaitė, *Social, Economical, and Educational Factors in Relation to Mathematics Achievement*, Tarptautinė matematikos dėstymo konferencija – 2005, 62-67.

- [19] Fullarton, S., *Closing the Gaps Between Schools: Accounting for Variation in Mathematics Achievement in Australian Schools Using TIMSS 95 and TIMSS 99*, IRC, 2004.
- [20] Gonzalez, E., Smith, T., *User Guide for the TIMSS International Database. Final Year of Secondary School*, IEA, 1998.
- [21] Gustafsson, A., M., *School production modelling to strengthen government monitoring programmes in developing countries*, Thesis, University of Stellenbosch, 2006.
- [22] Harley, H. O., Rao, J., *Maximum-likelihood estimation for the mixed analysis of variance model*, Biometrika, 54, 1967, 93.
- [23] Jarošova, E., *Estimation with the Linear Mixed Effects Model*, Journal of Applied Mathematics, volume 3, number 3, 2010.
- [24] Kiamanesh, A., *Items Affecting Iranian Students' Achievement in Mathematics*, IRC, 2004.
- [25] Kyriakides, L., Charalambous, C., *Extending the Scope of Analysing Data of IEA Studies: Applying Multilevel Modeling Techniques to Analyse TIMSS Data*, IRC, 2004.
- [26] Kim, S. J., Park, J. H., Park, S. W., Kim, S. S., *The Effects of School and Students' Educational Contexts in Korea, Singapore, and Finland using TIMSS 2011*, IRC, 2013.
- [27] Kunter, M., Baumert, J., *Constructivist Approches in the Secondary School Mathematics Classroom and Their Effects on Students' Learning, Interest and Sense of Challenge: a Re-analysis of the German TIMSS Data*, IRC, 2004.
- [28] Van der Leeden, R., Busing, F. M. T. A., and Meijer, E., *Bootstrap methods for two-level models*, Technical Report PRM 97-04, Leiden University, Department of Psychology, Leiden, 1997.
- [29] Leeuw, J., Meijer, E., *Handbook of Multilevel Analysis*, Springer, 2007.
- [30] Leeuw, J., Kreft, I., *Questioning Multilevel Models*, Journal of Educational and Behavioral Statistics, Summer 1995, Vol. 20, No. 2, 171-189.
- [31] Maas, C., Hox, J., *The influence of violations of assumptions on multilevel parameter estimates and their standard errors*, Computational Statistics & Data Analysis 46, 2004, 427 – 440.
- [32] Nacionalinio egzaminų centro Mokinių pasiekimų tyrimų ir analizės skyrius. *TIMSS 2011 Ataskaita, Matematika, 8 klasė*, 2011
- [33] Orelie, J., Edwards, L., *Fixed-effect variable selection in linear mixed models using R2 statistics*, Computational Statistics & Data Analysis, 52, 2008, 896 – 1907.
- [34] Pfeiffermann, Skinner, D., C.J., Holmes, D.J., Goldstein, H., Rasbash, J., *Weighting for unequal selection probabilities in multilevel models*, Journal of Royal Statistical Society, Series B, 1998, 60(1):23-40.
- [35] Phan, H., Sentovich, C., Kromrey, J., Dedrick, R., Ferron, J., *Correlates of Mathematics Achievement in Developed and Developing Countries: An HLM Analysis of TIMSS 2003 Eighth-grade Mathematics Scores*, pristatyma metiniame American Educational Research Association susitikime, Denver, Colorado, 2010.

- [36] Rao, C. R., *Estimation of heteroscedastic variances in linear models*, Journal of the American Statistical Association, 65, 1970, 161–172.
- [37] Rao, C. R., *Estimation of variance and covariance components–MINQUE theory*, Journal of Multivariate Analysis, 1, 1971a, 257–275.
- [38] Rao, C. R., *Minimum variance quadratic unbiased estimation of variance components*, Journal of Multivariate Analysis, 1, 1971b, 445–456.
- [39] Rao, C. R., *Estimation of variance and covariance components in linear models*, Journal of the American Statistical Association, 67, 1972, 112–115.
- [40] Rao, C.R., *MINQUE theory and its relation to ML and MML estimation of variance components*, Sankhya 41, 1979, 138-153.
- [41] Ren, S., Lai, H., Tong, W., Aminzadeh M., *Nonparametric bootstrapping for hierarchical data*, Journal of Applied Statistics, 37:9, 1487-1498, 2010.
- [42] Robitaille, D., Beaton, A., *Secondary Analysis of the TIMSS Data*, Kluwer Academic Publishers, 2002.
- [43] *Sample Design and Implementation*, [žiūrėta 2013-11-23], prieiga per internetą : <http://timssandpirls.bc.edu/methods/t-sample-design.html>.
- [44] Searle, S. R., Casella, G., McCulloch, Ch.E., *Variance Components*, John Wiley & Sons, Inc., New York, 1992.
- [45] Stidd, M., *A Comparative Analysis of Freedom and Educational Outcomes Using TIMSS and PIRLS*, University of Colorado Colorado Springs, 2012.
- [46] Snijders, T., Bosker, R., *Multilevel Analysis: An introduction to basic and advanced multilevel modeling*, SAGE Publications, 1999.
- [47] Shackman, G., *Developing Sample Sizes for New Surveys: Estimating the Design Effect Adjustment*, What's Hot / What's Next VI. Albany, 2001.
- [48] Švietimo ir mokslo ministerijos bei Atviros Lietuvos fondo spaudos konferencija, *Tarptautinis matematikos ir gamtos mokslų tyrimas TIMSS – R*, 2000.
- [49] Vakilian, S., *Simulation Studies on Estimation of Variance Components for Multilevel Models*, Theses and Dissertations (Comprehensive), paper 922, 2009.
- [50] Tong, H., Kumar, K., Huang, Y., *Developing Econometrics*, John Wiley & Sons, 2011.
- [51] Vonesh, E.F., Chinchilli, V.M., *Linear and Nonlinear Models for the Analysis of Repeated Measurements*, Marcel Dekker, 1997, 419–424.
- [52] Wiberg, M., Andersson, E., *School-Effectiveness in Mathematics in Sweden compared with countries in Europe and Asia-Pacific*, IRC, 2010.
- [53] Zuzovsky, R., *Grouping and Its effects on 8th Graders' Science and Mathematics Achievements*, IRC, 2004.

PRIEDAS

```
rm(list = ls())
library(nlme)
library(plyr)
library(gdata)
library(lme4)
library(mitools)
library(lmerTest)
library(robustlmm)
library(foreach)
library(Matrix)
library(gtools)
library(snow)
library(matrixcalc)
library(mitools)
library(MASS)
library(snow)
library(boot)
library(plyr)
library(mvtnorm)
library(expm)
library(simFrame)
library(foreach)
library(msm)
library(gtools)
library(MASS)
library(lme4)
library(gdata)
library(Matrix)
#####
# MINQUE funkcija HLM (arba misriu efektu) modeliams vertinti
#####
myMINQUE <- function(dt, fixed, random1 = NULL,
                     apriori = NULL) {
  # Kintamieji:
  #   dt - duomenys pateikti data.frame klaes objekte;
  #   fixed - fiksuotu efektu fomule pateikta character formatu;
  #   random1 - atsitiktiniu efektu formule pateikta character formatu;
  #   apriori - pradines dispersijos komponenciu reiksmes.

  N <- nrow(dt)
  # Suformuojam Y ir X matricas
  ff <- model.matrix(as.formula(fixed), model.frame(fixed, dt))
  Y <- as(as(model.frame(fixed, dt)[,1, drop = F], "matrix"), "Matrix")
  if (grepl("-1",fixed)) {
    X <- as(as(ff[,-1, drop = F], "matrix"), "Matrix")
  } else {
    X <- as(as(ff, "matrix"), "Matrix")
  }

  # Suformuojamos atsitiktiniu efektu matricos
  if (!is.null(random1)) {
    id1 <- strsplit(random1, "\\|")
  }
}
```

```

random1 <- unlist(id1)[1]
nmid1 <- unlist(id1)[2]
rr1 <- model.frame(random1, dt)
id1 <- model.frame(paste("~", nmid1), dt)[,1]
if (grepl("-1", random1)) {
  Z1 <- as(as( rr1, "matrix"), "Matrix")
} else {
  Z1 <- as(as(cbind(matrix(1, ncol = 1, nrow = nrow(rr1),
                        dimnames = list(NULL, "(Intercept)")), rr1),
            "matrix"), "Matrix")
}
q <- ncol(Z1)
colnames(Z1) <- paste(nmid1, colnames(Z1), sep = ".")
clnms1 <- colnames(Z1)
Z1 <- foreach(i = unique(id1)) %do% {Z1[id1 %in% i, ,drop = F]}
Z <- bdiag(Z1)
n1 <- length(unique(id1))
l <- q*(q+1)/2
TT1 <- formTT(q)
QI <- llply(TT1, function(tt) llply(Z1, function(z) z%*%tt%*%t(z)))
} else {
  Z1 <- NULL
  clnms1 <- NULL
  n1 <- NULL
  QI <- NULL
}

l <- l+1
Qj <- c(list(llply(QI[[1]], function(zj) diag(nrow(zj)))), QI)
QI <- llply(Qj, function(x) bdiag(x))
Xj <- foreach(i = unique(id1)) %do% {X[id1 %in% i, ,drop = F]}
X <- as(foreach(ii = 1:length(Xj), .combine = rbind) %do%
      as.matrix(Xj[[ii]]), "Matrix")
Yj <- foreach(i = unique(id1)) %do% {Y[id1 %in% i, ,drop = F]}
Y <- as(foreach(ii = 1:length(Xj), .combine = rbind) %do%
      as.matrix(Yj[[ii]]), "Matrix")

if (is.null(apriori)) apriori <- rep(1, length(QI))

Vj <- foreach(i = 1:length(unique(id1))) %do%
  solve(Reduce("+", mapply(function(qj, th) th*qj[[i]], Qj,
    apriori)))
V <- bdiag(Vj)

iM <- iMINQUE(V, X, Y, QI, wgt2i)
gc()
theta <- iM$theta
print(theta)
iM <- iMINQUE(iM$V, X, Y, QI, wgt2i)

P <- ginv(as.matrix((t(X)%*%iM$V%*%X)))
iM$beta <- P%*%t(X)%*%iM$V%*%Y
gc()
sigma2 <- iM$theta[1]

```

```

names(sigma2) <- "sigma2"
TT <- fillT(iM$theta[-1])
dimnames(TT) <- list(clnms1, clnms1)
beta <- as.numeric(iM$beta)
names(beta) <- colnames(X)
return(list(sigma2 = sigma2, TT = TT, beta = beta, N = N, n1 = n1))
}

# S matricos uzpildymui
fillSMINQUE2 <- function(CQ) {
  # Kintamieji:
  #   CQ - matricu C ir Qi sandaugu sarasas (list).

  x <- matrix(0, ncol = length(CQ), nrow = length(CQ))
  ns <- nrow(x)
  cc <- combinations(ns,2,1:ns, repeats.allowed = T)
  foreach(ii = 1:nrow(cc)) %do% {
    k <- cc[ii, 1]
    l <- cc[ii, 2]
    s <- sum(diag(CQ[[k]]%*%CQ[[l]]))
    x[k,l] <- s
    x[l, k] <- s
    gc()
  }
  gc()
  return(x)
}

# MINQUE iteracijai
iMINQUE <- function(V, X, Y, QI) {
  XVX <- ginv(as.matrix((t(X)%*%V%*%X)))
  Cjj <- V-V%*%X%*%XVX%*%t(X)%*%V#(V*unlist(wgt2i))
  CQ <- llply(QI, function(qi) Cjj%*%qi)

  S <- fillSMINQUE2(CQ)
  WI <- laply(CQ, function(cq) sum(diag(t(Y)%*%cq%*%Cjj%*%Y)))

  theta0 <- ginv(S)%*%matrix(WI, ncol = 1)

  V <- ginv(as.matrix(Reduce("+", mapply(function(qi, th)
    {qi*th}, QI, theta0))))
  return(list(theta = theta0, V = V, P = XVX))
}

#####
# MINQUE su saviranka
#####
bootMINQUE <- function(data = data.2011, FUN = myMINQUE, B = 5,
  fixed = "BSMMAT01 ~ 1+SSEX",
  random = "~1|IDSCHOOL",
  idstrata = "IDSTRATE", idschool = "IDSCHOOL",
  idstud = "IDSTUD", apriori = NULL,
  BFUN = bootSample) {
  # Kintamieji:

```

```

# data - duomenys pateikti data.frame klaes objekte;
# FUN - MINQUE f-jos pavadinimas;
# B - savirankos iteraciju skaicius;
# fixed - fiksuotu efektu fomule pateikta character formatu;
# random - atsitiktiniu efektu formule pateikta character formatu;
# idstrata - data stulpelio pavadinimas, kuriame identifikuojami lizdai;
# idschool - stulpelio pavadinimas, kuriame identifikuojama klasė;
# idstud - stulpelio pavadinimas, kuriame identifikuojamas individas;
# apriori - pradines dispersijos komponenciu reiksmes;
# BFUN - f-jos pavadinimas savirankoms imtims.

print("Iteration:")
t0 <- FUN(data, fixed = fixed, random = random)

bootRes <- foreach(ii = 1:B, .combine = rbind) %do% {
  print(paste("boot", ii))
  smpldt <- try(BFUN(data, idstrata, idschool,
                    idstud, wgt))
  if(class(smpldt) == "try-error") return(NULL)
  res <- try(FUN(smpldt, fixed = fixed, random = random, apriori = apriori))
  if(class(res) == "try-error") return(NULL)
  cc <- c(res$beta, res$sigma2, unlist(res$TT))
  names(cc) <- c("beta", "sigma2", "")
  return(cc)
}
print("Collect")

# Rezultatų suvedimas

tt <- bootRes
# Fiksuoti efektai
beta0 <- t0$beta
beta <- tt[,1:length(beta0), drop = F]
colnames(beta) <- rownames(beta0)
Bcov <- cov(beta)
Bint <- t(t(beta)-as.numeric(beta0))
Bint <- apply(Bint, 2, function(x) quantile(x, probs=c(0.025,0.975)))
Bint <- rbind((t(beta0)-Bint[2,]), (t(beta0)-Bint[1,]))
rownames(Bint) <- c("2.5%", "97.5%")
beta <- apply(beta, 2, mean)
Bbias <- beta - beta0
beta <- beta0-Bbias
beta <- cbind(Coefficient = beta, St.err = sqrt(diag(Bcov)),
              Bias = Bbias, t(Bint))
colnames(beta)[1:3] <- c("Coeff", "St.err", "Bias")
attr(beta, "cov") <- Bcov

tt <- tt[,-1:-length(beta0), drop = F]

# Sigma
sigma0 <- t0$sigma2
sigma <- tt[,1, drop = F]
colnames(sigma) <- "sigma2"
Scov <- cov(sigma)

```

```

#Sint <- apply(sigma, 2, function(x) quantile(x, c(.025, .975)) )
Sint <- t(t(sigma)-(sigma0))
Sint <- apply(Sint, 2, function(x) quantile(x, probs=c(0.025,0.975)))
Sint <- rbind((t(sigma0)-Sint[2,]), (t(sigma0)-Sint[1,]))
rownames(Sint) <- c("2.5%", "97.5%")
sigma <- apply(sigma, 2, mean)
Sbias <- sigma - sigma0
sigma <- sigma0 - Sbias
sigma <- cbind(Coefficient = sigma, St.err = sqrt(diag(Scov)),
               Bias = Sbias, t(Sint))
attr(sigma, "cov") <- Scov

tt <- tt[,-1, drop = F]

# Atsitiktiniai efektai
TT0 <- as.numeric(t0$TT)
TT <- tt
TTcov <- cov(TT)
TTint <- t(t(TT)-(TT0))
TTint <- apply(TTint, 2, function(x) quantile(x, probs=c(0.025,0.975)))
TTint <- rbind((t(TT0)-TTint[2,]), (t(TT0)-TTint[1,]))
rownames(TTint) <- c("2.5%", "97.5%")

TT <- apply(TT, 2, mean)
TTbias <- TT - TT0
TT <- TT0 - TTbias
TT <- cbind(Coefficient = TT, St.err = sqrt(diag(TTcov)), Bias = TTbias,
            t(TTint))

return(list(beta = beta, sigma2 = sigma, TT = TT, t0 = t0, result = bootRes))
}

bootSample <- function(dt, idstrata = "IDSTRATE", idschool = "IDSCHOOL",
                       idstud = "IDSTUD")) {
  strata <- dt[, idstrata]
  sd <- foreach(ii = unique(strata), .combine = rbind) %do% {
    sdata <- dt[strata %in% ii, ,drop = F]
    sch <- unique(sdata[,idschool])
    nh <- length(sch)
    if (nh == 1) nh <- 2
    smSCH <- as.numeric(sample(as.character(sch), nh, replace = T))
    smSCH <- sch
    std <- foreach(jj = smSCH, .combine = rbind) %do% {
      studdt <- sdata[sch %in% jj,,drop = F]
      idst <- studdt[, idstud]
      nhc <- length(idst)
      smplStud <- sample(idst, nhc, replace = T)
      studdt1 <- foreach(kk = smplStud, .combine = rbind) %do%
        studdt[idst %in% kk,,drop = F]
      studdt1$nhc <- nhc
      return(studdt1)
    }
  }
  return(std)
}

```

```

    return(sd)
}

#####
# Naudojimosi pavyzdys
#####
n1 <- 10
beta0 <- 50+rmnorm(n1, 0, sqrt(0.25))
df <- data.frame(IDCLASS = 1:10, beta0 = beta0)
df <- foreach(ii = 1:10, .combine = rbind) %do% {data.frame(df[ii,],
    IDINDIV = paste(ii, 1:5, sep = ""))}
df$X <- rmnorm(nrow(df), 10, sqrt(50))
df$e <- rmnorm(nrow(df), 0, sqrt(0.5))
df$y <- df$beta0+4*df$X+df$e

min1 <- myMINQUE(dt = df,
    fixed = " y~ 1+X",
    random1 = "~1|IDCLASS")

min1$beta
min1$TT
min1$sigma

df$idstrata <- 1
bb <- bootMINQUE(fixed = "y ~ 1+X", idschool = "IDCLASS",
    idstud = "IDINDIV", random = "~1|IDCLASS", wgt = NULL,
    idstrata = "idstrata", R = 100, data = df)

bb$beta
bb$TT
bb$sigma

```

Lentelių sąrašas

| | | |
|---|---|----|
| 1 | Lentelėje pateiktos santykinio poslinkio statistikos (MBias - santykinio poslinkio vidurkis, SDBias - santykinio poslinkio standartinis nuokrypis, Min ir Max- atitinkamai mažiausia ir didžiausia įgyta santykinio poslinkio reikšmė, RMSE - santykinė vidutinė kvadratinė paklaida) visiems metodams, kai abiejų lygių paklaidos normalios. | 17 |
| 2 | Lentelėje pateikiamas pasikliautinių intervalų padengiamumas, kai vertinami duomenys pasiskirstę pagal normalųjį skirstinį (t.y. pirmo ir antro lygio paklaidos normalios), tikroji reikšmė 0,95. | 18 |
| 3 | Lentelėje pateiktos santykinio poslinkio statistikos (MBias - santykinio poslinkio vidurkis, SDBias - santykinio poslinkio standartinis nuokrypis, Min ir Max- atitinkamai mažiausia ir didžiausia įgyta santykinio poslinkio reikšmė, RMSE - santykinė vidutinė kvadratinė paklaida) visiems metodams, kai abiejų lygių paklaidos pasiskirsčiusios pagal χ^2 skirstinį. | 18 |
| 4 | Lentelėje pateikiamas pasikliautinių intervalų padengiamumas, kai vertinami duomenys pasiskirstę pagal χ^2 skirstinį (t.y. pirmo ir antro lygio paklaidos iš χ^2). | 19 |
| 5 | Magistro darbe naudotų TIMSS duomenų detalizavimas | 24 |
| 6 | Nulinio modelio parametų įverčiai; REML su saviranka metodas. | 26 |
| 7 | Nulinio modelio parametų įverčiai; REML su saviranka metodas. | 27 |

Iliustracijų sąrašas

| | | |
|---|--|----|
| 1 | Mokinio matematinio raštingumo pirmos galimos testo atlikimo reikšmės teorinių kvantilių grafikas. | 25 |
|---|--|----|