

TARTU ÜLIKOOL
LOODUS- JA TÄPPISTEADUSTE VALDKOND
MATEMAATIKA JA STATISTIKA INSTITUUT

Darja Rohi
**Matemaatikavõistlustel osalemist ja edukust
mõjutavad tegurid**

Matemaatiline statistika
Bakalaureusetöö (9 EAP)

Juhendaja: PhD Krista Fischer

TARTU 2026

MATEMAATIKAVÕISTLUSTEL OSALEMIST JA EDUKUST MÕJUTAVAD
TEGURID

Bakalaureusetöö

Darja Rohi

Lühikokkuvõte

Käesoleva bakalaureusetöö eesmärk on uurida, kas algklassides matemaatikavõistlustel edukad õpilased jõuavad TOP 200 hulka ka gümnaasiumiastmes ning millised tunnused on selle tulemusega seotud. Töös analüüsitakse rahvusvahelise matemaatikavõistluse „Känguru“ andmeid aastatest 2012–2024, et hinnata varajase edu püsimist ning sellega seotud tegureid.

Analüüsis käsitletakse õpilaste edukust ning selle seost soo, koolitüübi ja kooli asukohaga. Edukuse kirjeldamiseks ja selgitamiseks kasutatakse kolme lähenemist: kirjeldavat analüüsi, statistilise tõenäosuse hindamist ning logistilist regressiooni. Statistilise tõenäosuse abil hinnatakse, kui suur osa algklassides edukatest õpilastest jõuab TOP 200 hulka ka gümnaasiumiastmes. Logistilise regressiooni abil hinnatakse soo, koolitüübi ja kooli asukoha samaaegset seost selle tõenäosusega.

Töö tulemustest selgus, et varajane edu kandub edasi vaid vähestel juhtudel. Samuti ilmnas, et nii sugu kui ka koolitüüp on seotud TOP 200 hulka jõudmisega gümnaasiumiastmes. Poistel ja eliitkoolide õpilastel on suurem tõenäosus saavutada sarnane tulemus ka hiljem. Kooli asukoha seos ei olnud kõigi rühmade puhul statistiliselt oluline.

CERCS teaduseriala: P160 Statistika, operatsioonianalüüs, programmeerimine, finants- ja kindlustusmatemaatika.

Märksõnad: Matemaatikavõistlused, osalemine, edukus, logistiline regressioon, haridusstatistika.

FACTORS INFLUENCING PARTICIPATION AND SUCCESS IN MATHEMATICS COMPETITIONS

Bachelor thesis

Darja Rohi

Abstract

The aim of this bachelor's thesis is to examine whether students who perform well in mathematics competitions in early school years also reach the TOP 200 at the upper secondary level, and which characteristics are associated with this outcome. The analysis is based on data from the international mathematics competition "Kangaroo" covering the years 2012–2024.

The study focuses on students' success and its association with gender, school type, and school location. Three approaches are used to describe and explain success: descriptive analysis, statistical probability estimation, and logistic regression. Statistical probability is used to estimate how many students who were successful in early school years also reach the TOP 200 later. Logistic regression is applied to assess the simultaneous association between gender, school type, school location, and this probability.

The results show that early success is carried forward only in a small number of cases. Gender and school type are significantly associated with reaching the TOP 200 at the upper secondary level. Boys and students from elite schools have a higher probability of achieving similar success later. The association with school location was not statistically significant for all groups.

CERCS research specialisation: P160 Statistics, operations research, programming, financial and actuarial mathematics.

Key Words: mathematics competitions, participation, success, logistic regression, educational statistics.

Sisukord

| | |
|---|-----------|
| Sissejuhatus | 4 |
| 1 Andmestik ja andmete töötlemine | 7 |
| 1.1 Andmestiku kirjeldus | 7 |
| 1.2 Andmete puhastamine ja tunnuste loomine | 8 |
| 2 Metoodika | 11 |
| 2.1 Uuritav tunnus ja selgitavad muutujad | 11 |
| 2.2 Statistilise tõenäosuse hindamine | 12 |
| 2.3 Logistiline regressioon | 13 |
| 2.4 Statistiline olulisus | 13 |
| 2.5 Analüüsi piirangud | 14 |
| 3 Andmeanalüüs | 15 |
| 3.1 Kirjeldav analüüs | 15 |
| 3.2 Edukuse püsimise tõenäosus | 20 |
| 3.3 Logistiline regressioonanalüüs | 25 |
| 3.4 Tulemuste arutelu | 28 |
| Kokkuvõte | 30 |
| Kasutatud allikad | 31 |

Sissejuhatus

Matemaatikaalased saavutused koolis on pikka aega olnud haridusteaduslike uurimuste keskmes. Varasemad uuringud on näidanud, et varased matemaatilised oskused on olulised hilisema akadeemilise edukuse ennustajad (Duncan *et al.*, 2007). Üheks võimaluseks matemaatilisi teadmisi ja oskusi arendada on osalemine ainevõistlustel. Näiteks rahvusvaheline matemaatikavõistlus "Känguru" pakub õpilastele võimalust rakendada oma teadmisi ning testida neid võistlusolukorras. Sellised võistlused võivad toetada õpilaste huvi matemaatika vastu ning arendada probleemilahendusoskusi (98th Percentile, 2023).

Käesoleva bakalaureusetöö eesmärk on uurida, kui suur osa algklassides TOP 200 hulka kuulunud õpilastest jõuab TOP 200 hulka ka gümnaasiumiastmes ning millised tunnused on selle tulemusega seotud.

Töö keskendub järgmistele uurimisküsimustele:

- Kas algklassides TOP 200 hulka kuulunud õpilased jätkavad osalemist matemaatikavõistlustel gümnaasiumiastmes ning kui suur on tõenäosus, et nad jõuavad ka seal TOP 200 hulka?
- Millised tunnused, näiteks sugu, koolitüüp ja kooli asukoht, on seotud sellega, kas algklassides TOP 200 hulka kuulunud õpilane jõuab TOP 200 hulka ka gümnaasiumiastmes?
- Kuidas jaotuvad TOP 200 hulka kuuluvad osalemised soo, koolitüübi ja kooli asukoha lõikes eri kooliastmetes?

Varasemad uuringud on näidanud, et soolised erinevused matemaatikasaavutustes esinevad mitmes riigis, kuid need erinevused on pigem kultuurilised

kui bioloogilised (Else-Quest, Hyde ja Linn, 2010). Samuti on leitud, et koolikeskkond, kooli sotsiaal-majanduslik profiil ja õppimisvõimalused võivad olla seotud õpilaste õpitulemustega (OECD, 2018). Seetõttu on oluline analüüsida, kas ja kuidas sellised tunnused on seotud matemaatikavõistlustel osalemise ja edukusega.

Uuringu olulisus seisneb selles, et selgitades, kas ja kuidas varajane edu säilib ning millised tunnused on selle püsimisega seotud, saab pakkuda teadmisi hariduspoliitika kujundamiseks. Võimalike erinevuste tuvastamine aitab mõista, millised õpilasarühmad vajavad rohkem tuge ning kuidas oleks võimalik soodustada kõigi võrdseid võimalusi hariduses. Eriti väärtuslik on see info olukorras, kus Eestis ja mujal maailmas otsitakse viise, kuidas tõsta õpilaste huvi reaalinete vastu ning vähendada soolist lõhet teadus- ja tehnoloogiavaldkonnas (Else-Quest, Hyde ja Linn, 2010).

Analüüs põhineb rahvusvahelise matemaatikavõistluse "Känguru" andmetel, milles osalevad igal aastal tuhanded Eesti õpilased. Kasutatud andmed hõlmavad mitme aasta tulemusi erinevatest klassiastmetest ning sisaldavad teavet õpilase kooli tüübi ja kooli asukoha kohta. Töös rakendatakse kvantitatiivseid statistilisi meetodeid, sealhulgas kirjeldavat analüüsi ja logistilist regressiooni, et hinnata erinevate tunnuste seost sellega, kas algklassides edukas õpilane jõuab TOP 200 hulka ka gümnaasiumiastmes. Analüüs keskendub TOP 200 hulka kuulunud õpilastele, mistõttu tuleb tulemuste tõlgendamisel arvestada, et need ei kirjelda kogu osalejate populatsiooni.

Töö on üles ehitatud järgmiselt. Esimeses peatükis kirjeldatakse kasutatud andmestikku ja andmete töötlemist. Teises peatükis tutvustatakse rakendatud meetodikat. Kolmandas peatükis esitatakse andmeanalüüsi tulemused ning arutletakse nende tõlgendamise üle.

Autor tänab bakalaureusetöö juhendajat Krista Fischerit väärtuslike soovitude ja asjatundliku juhendamise eest.

1 Andmestik ja andmete töötlemine

1.1 Andmestiku kirjeldus

Käesolevas töös analüüsitakse andmeid rahvusvahelise Känguru matemaatikavõistluse tulemuste põhjal, mis on pärit Tartu Ülikooli Teaduskoolist. Andmestik katab ajavahemiku 2012–2024 ning sisaldab kokku 15 166 kirjet (Tartu Ülikooli Teaduskool, [2024](#)).

Tuleb siiski märkida, et andmestikus puuduvad andmed 2020. aasta kohta, kuna sel aastal ei toimunud võistlus COVID-19 pandeemia tõttu. Samuti ei sisalda andmestik kõige hilisema aasta andmeid, kuna alates 2025. aastast on võistluse formaat muutunud ning osa osalejatest sooritab ülesandeid veebipõhiselt. Selline muutus võib mõjutada tulemuste võrreldavust varasemate aastatega, mistõttu neid andmeid käesolevas analüüsis ei käsitleta.

Algandmestikus sisalduvad järgmised tunnused:

- aasta – võistluse toimumise aasta,
- vanuserühm – õpilase vanusele vastav võistlusklass (nt pre-ekolier, ekolier, benjamin, kadett, juunior, student),
- koht – õpilase saavutatud koht vastavas vanuserühmas,
- nimi – osaleja ees- ja perekonnanimi,
- klass – õpilase klass võistluse toimumise ajal,
- kool – õpilase kooli nimi,
- punkte – võistlusel saadud punktisumma.

Andmestik ei sisalda otseselt mitmeid analüüsiks vajalikke tunnuseid, näiteks osaleja sugu ega infot selle kohta, kas sama isik on osalenud erinevatel aastatel. Samuti puuduvad unikaalsed identifikaatorid, mis võimaldaksid osalejaid üheselt tuvastada.

1.2 Andmete puhastamine ja tunnuste loomine

Kuna algandmestikus puudusid mitmed analüüsiks vajalikud tunnused, viidi läbi andmete täiendav puhastamine ning uute tunnuste konstrueerimine.

Esmalt eraldati igast kirjest ees- ja perekonnanimi. Seejärel määrati õpilastele sugu, kuna algandmestikus vastav info puudus.

Soo määramiseks leiti kõik unikaalsed eesnimed, mida oli kokku 2106. Iga nimi kontrolliti käsitsi Eesti Statistikaameti nimeandmebaasis (Statistikaamet, 2026), mille abil hinnati, kas nimi esineb meeste või naiste seas.

Kui nimi esines andmetel ainult ühe soo puhul, määrati sellele vastav sugu. Mitmetähenduslikud nimed (123 unikaalset nime), mis esinesid nii meeste kui naiste seas, märgiti esialgu määramata kategooriasse.

Nende juhtumite puhul kasutati täiendavat infot, näiteks teisi eesnimesid või ees- ja perekonnanime kombinatsiooni. Kui ka see ei võimaldanud soo määramist, teostati täiendav käsitsi kontroll avalike allikate põhjal (nt võistlustulemused või muud avalikult kättesaadavad andmed). Kõigi nimede puhul õnnestus sugu lõpuks määrata, mistõttu määramata soo tõttu ühtegi õpilast analüüsist välja ei jäetud.

Andmete kontrollimisel kasutati üksnes avalikult kättesaadavat infot ning seda tehti ainult käesoleva teadustöö eesmärgil. Kogutud teavet ei kasutatud muudel eesmärkidel ega säilitatud viisil, mis võimaldaks isikute edasist tu-

vastamist. Töös ei kasutatud tundlikke isikuandmeid ning analüüs viidi läbi koondatud kujul, tagades uuritavate isikute privaatsuse.

Pärast soo määramist viidi läbi andmete põhjalik puhastamine. Selle käigus kontrolliti kogu andmestik käsitsi, et tuvastada ja parandada võimalikud vead, nagu kirjavead, nimede erinevad kirjakuju variandid ning juhtumid, kus ees- ja perekonnanimi olid vahetusse läinud.

Lisaks vaadati kirjed läbi tähestikulises järjekorras, et tuvastada sama nimega kirjed. Selliste kirjete puhul võrreldi õpilase klassi, osalemisaastaid ja kooli, et hinnata, kas tegemist on sama isiku või erinevate isikutega. Näiteks ei ole võimalik, et sama isik osaleb varasemal aastal kõrgemas klassis ning hilisemal aastal madalamas klassis.

Juhtudel, kus sama nimega kirjed vastasid erinevatele isikutele, eristati need üksteisest, lisades nimele täiendava tunnuse (nt indeks või sümbol). Samuti parandati ilmseid sisestusvigu, mis võisid põhjustada sama isiku esinemise erinevate kirjetena.

Pärast andmete puhastamist ja isikute eristamist loodi õpilastele unikaalne identifikaator, mis põhines korrigeeritud nimekombinatsioonil.

Järgmisena loodi tunnus, mis kirjeldab kooli tüüpi. Koolid jaotati kahte rühma: eliitkoolid ja tavakoolid. Tallinna koolide puhul lähtuti ülelinnalise vastuvõtuga koolide nimekirjast. Eliitkoolide hulka liigitati järgmised koolid: Tallinna Reaalkool, Tallinna Inglise Kolledž, Tallinna Juudi Kool, Tallinna Prantsuse Lütseum, Tallinna Tõnismäe Reaalkool, Vanalinna Hariduskollegium ning Miina Härma Gümnaasium. Kõik ülejäänud koolid käsitleti tavakoolidena.

Lisaks määrati koolide asukoht. Selleks leiti kõik unikaalsed koolide nimetused ning iga kooli asukoht tuvastati eraldi. Koolid jaotati nelja rühma:

Tallinn, Tartu, Narva ja muud asulad.

Andmete puhastamise käigus eemaldati andmestikust ka kirjed, mis ei kuulunud Eesti koolide hulka, nimelt Stockholmi Eesti Kool ja European School of Brussels.

2 Metoodika

Käesolevas töös kasutatakse kvantitatiivseid statistilisi meetodeid, et analüüsida, millised tunnused on seotud algklassides saavutatud TOP 200 tulemuse püsimisega gümnaasiumiastmes. Analüüsi eesmärk on hinnata, millised tunnused on seotud sellega, kas algklassides edukad õpilased jätkavad osalemist matemaatikavõistlustel ning kas nad jõuavad TOP 200 hulka ka gümnaasiumiastmes.

Oluline on märkida, et erinevates analüüsi etappides kasutatakse erinevat analüüsiühikut. Kirjeldavas analüüsis käsitletakse iga andmestiku kirjet eraldi vaatlusena, kus iga kirje vastab ühele õpilase osalemisele konkreetsel aastal. Seega kajastab see osa analüüsist osalemiste sagedust, mitte unikaalsete õpilaste arvu.

Edukuse püsimise analüüsis ja logistilises regressioonis käsitletakse aga õpilast ühe korra ehk unikaalse isikuna. Selleks koondatakse sama õpilase korduvad osalemised ning vaadeldakse, kas ta kuulus algklassides vähemalt ühel korral TOP 200 ning kas ta jõudis TOP 200 hulka ka gümnaasiumiastmes vähemalt ühel korral. Selline lähenemine võimaldab keskenduda varajase edu püsimisele, mitte osalemiste arvule.

2.1 Uuritav tunnus ja selgitavad muutujad

Peamiseks uuritavaks tunnuseks on binaarne tunnus, mis kirjeldab varajase edu püsimist. Täpsemalt vaadeldakse, kas õpilane, kes kuulus algklassides TOP 200 hulka, jõuab TOP 200 hulka ka gümnaasiumiastmes. Tunnus defineeritakse järgmiselt:

- $Y = 1$, kui õpilane jõuab TOP 200 hulka ka gümnaasiumiastmes;

- $Y = 0$, kui õpilane ei jõua TOP 200 hulka gümnaasiumiastmes.

Selline tunnuse defineerimine võimaldab hinnata varajase edu püsimist käesolevas töös kasutatud TOP 200 tulemuste põhjal.

Selgitavateks muutujateks on:

- sugu (poiss / tüdruk),
- koolitüüp (eliitkool / tavakool),
- kooli asukoht (Tallinn, Tartu, Narva, muud asulad).

Kategoorilised tunnused kodeeriti binaarsete indikaatormuutujatena, kusjuures võrdlusgrupiks valiti poisid, tavakoolid ja muud asulad. Täpsem kodeerimine on kirjeldatud alapeatükis [2.3](#).

2.2 Statistilise tõenäosuse hindamine

Varajase edu püsimise tõenäosust hinnatakse statistilise tõenäosusena ehk suhtelise sagedusena. Statistiline tõenäosus on statistikas laialdaselt kasutatav lähenemine, kus sündmuse tõenäosust hinnatakse vaatluste suhtelise sagedusena (Ross, [2010](#)).

Olgu N algklassides TOP 200 hulka jõudnud õpilaste arv ning $N(A)$ nende õpilaste arv, kes jõuavad TOP 200 hulka ka gümnaasiumiastmes. Siis hinnatakse tõenäosust kujul

$$P(A) = \frac{N(A)}{N}. \quad (1)$$

Statistilist tõenäosust kasutatakse erinevate tunnuste lõikes, et kirjeldada, kuidas edukuse püsimise tõenäosus varieerub näiteks soo, koolitüübi ja asukoha järgi.

2.3 Logistiline regressioon

Erinevate tegurite samaaegse seose hindamiseks kasutatakse logistilist regressiooni. Logistiline regressioon on statistiline meetod, mida kasutatakse binaarse tunnuse modelleerimiseks (Hosmer, Lemeshow ja Sturdivant, 2013).

Mudelis vaadeldakse tõenäosust, et õpilane jõuab TOP 200 hulka gümnaasiumiastmes. Tähistades selle sündmuse tõenäosust $p_i = P(Y_i = 1)$, esitatakse logistilise regressiooni mudel üldkujul

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki}, \quad (2)$$

kus $X_{1i}, X_{2i}, \dots, X_{ki}$ tähistavad mudelis kasutatud selgitavaid tunnuseid või nende põhjal moodustatud indikaatortunnuseid.

Kuna töös kasutatud selgitavad tunnused on kategoorilised, kodeeritakse need regressioonanalüüsis indikaatortunnustena. Täpsemalt kirjeldatakse tunnuste kodeerimist logistilise regressioonanalüüsi tulemuste esitamisel.

Mudeli parameetrid hinnatakse suurima tõepära meetodil. Koefitsientide eksponentsiaalsed väärtused $\exp(\beta_i)$ vastavad šansisuhetele, mis näitavad, kuidas muutub sündmuse toimumise šanss võrreldes vastava võrdluskategooriaga. Lisaks p-väärtustele arvestatakse tulemuste tõlgendamisel ka šansisuhete suurust, et hinnata seoste praktilist olulisust.

2.4 Statistiline olulisus

Mudeli parameetrite hindamisel kasutatakse olulisuse nivood $\alpha = 0,05$, mis on statistikas tavapäraselt kasutatav lävi hüpoteeside kontrollimiseks (Agresti, 2018). Tulemusi loetakse statistiliselt oluliseks, kui p-väärtus on väiksem kui 0,05.

Tulemuste tõlgendamisel eristatakse statistilist ja praktilist olulisust ning välditakse liigseid järeldusi olukordades, kus seos ei ole statistiliselt kinnitatud.

2.5 Analüüsi piirangud

Analüüsi tulemuste tõlgendamisel tuleb arvestada mitmete piirangutega.

Andmestikus puuduvad unikaalsed identifikaatorid, mistõttu sama nimega isikute eristamine ei pruugi olla täielikult täpne. Isikute tuvastamine põhines nimede, klassi, kooli ja osalemisaastate kombinatsioonil, kuid selline lähenemine võib teatud juhtudel viia selleni, et sama isik on käsitletud mitme erineva isikuna või erinevad isikud on ekslikult ühendatud üheks. See võib mõjutada eelkõige varajase edu püsimise analüüsi, kuna see sõltub otseselt sama isiku jälgimisest ajas.

Samuti määrati sugu nime põhjal, mis võib teatud juhtudel põhjustada vigu, eriti mitmetähenduslike nimede korral. Kuigi kasutati täiendavat käsitsi kontrolli, ei ole võimalik täielikult välistada võimalikke ebatäpsusi.

Lisaks keskendub analüüs ainult nendele õpilastele, kes kuulusid algklassides TOP 200 hulka. Seetõttu ei kajasta tulemused kogu osalejate populatsiooni ning järeldusi tuleb käsitleda pigem kirjeldavate seostena, mitte põhjuslike järeldustena.

Oluline piirang on ka see, et andmestik ei võimalda eristada õpilasi, kes gümnaasiumiastmes võistlusel ei osalenud, nendest, kes osalesid, kuid ei jõudnud TOP 200 hulka. Seetõttu tõlgendatakse tulemusi varajase edu püsimisena TOP 200 tulemuste põhjal, mitte osalemise jätkamise otsese mõõdikuna.

3 Andmeanalüüs

3.1 Kirjeldav analüüs

Vaatlusalusel perioodil sisaldab andmestik kokku 15 166 kirjet, kus iga kirje vastab ühele õpilase osalemisele matemaatikavõistlusel ja tema kuulumisele TOP 200 hulka konkreetsel aastal. See tähendab, et andmestik ei kajasta unikaalsete õpilaste arvu, vaid edukate osalemiste sagedust, mistõttu võib sama õpilane esineda andmetes mitmel korral.

Kogu valimi lõikes moodustavad poistest osalemised 63,57% ja tüdrukutest 36,43%, mis viitab sellele, et poisid on TOP 200 hulgas selgelt ülesindatud. Kui seda jaotust vaadata aastate lõikes, siis selgub, et tegemist ei ole üksikute aastate eripäraga, vaid püsiva mustriga.



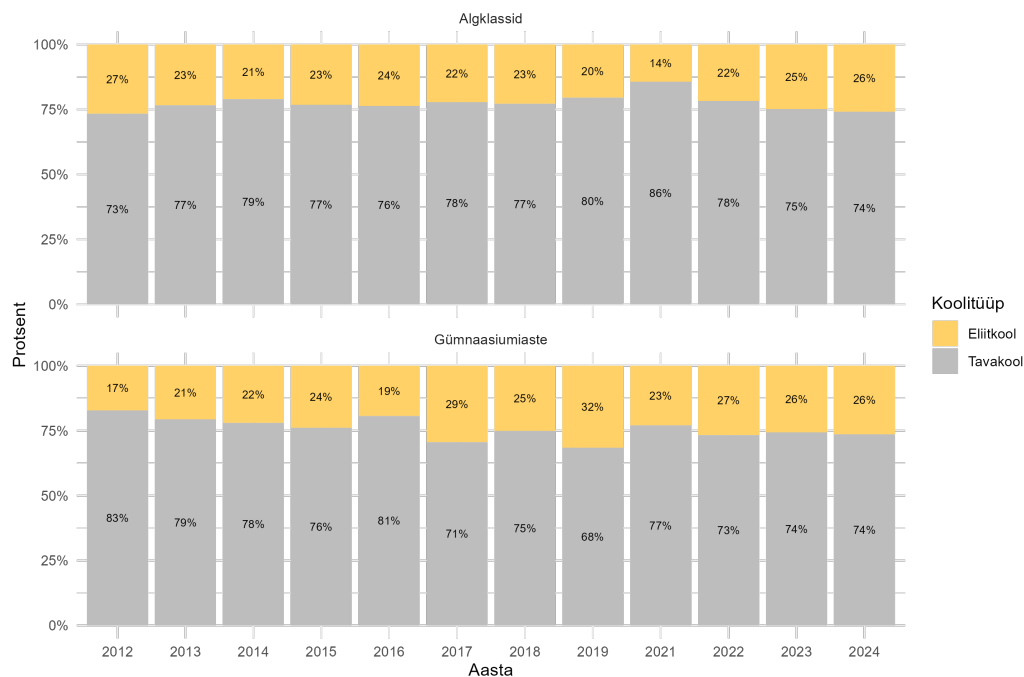
Joonis 1: TOP 200 hulka kuulunud osalemiste sooline jaotus algklassides ja gümnaasiumiastmes aastate lõikes.

Jooniselt 1 selgub, et nii algklassides kui ka gümnaasiumiastmes on poisid TOP 200 hulgas ülesindatud. Algklassides jääb poiste osakaal vaadeldavatel aastatel ligikaudu vahemikku 51% kuni 64%, samas kui tüdrukute osakaal jääb vahemikku 36% kuni 49%. Seega on algklassides sooline jaotus küll poiste kasuks kaldu, kuid erinevus ei ole väga suur ning mõnel aastal moodustavad tüdrukud peaaegu poole TOP 200 hulka kuulunud osalemistest.

Gümnaasiumiastmes on erinevus selgem. Poiste osakaal jääb enamasti ligikaudu 70% ümbrusesse ning ulatub mõnel aastal ka üle 75%, samas kui tüdrukute osakaal jääb enamasti 25% – 30% vahele. Võrreldes algklassidega on seega gümnaasiumiastmes poiste ülekaal TOP 200 hulgas märgatavalt suurem.

Aastate lõikes ei ilmne kummaski kooliastmes selget ühesuunalist kasvavat või kahanevat trendi. Pigem võib öelda, et sooline erinevus on püsiv muster: poisid moodustavad TOP 200 hulgas suurema osa nii algklassides kui ka gümnaasiumiastmes. Samas näitab kooliastmete võrdlus, et see erinevus on gümnaasiumiastmes tugevam. See viitab sellele, et tüdrukute osakaal parimate hulgas väheneb kõrgemas kooliastmes, mis võib olla seotud nii osalemise jätkamise, motivatsiooni kui ka võistlustel osalemise toetamisega.

Järgnevalt vaadeldakse TOP 200 hulka kuulunud osalemiste jaotust koolitüübi lõikes.



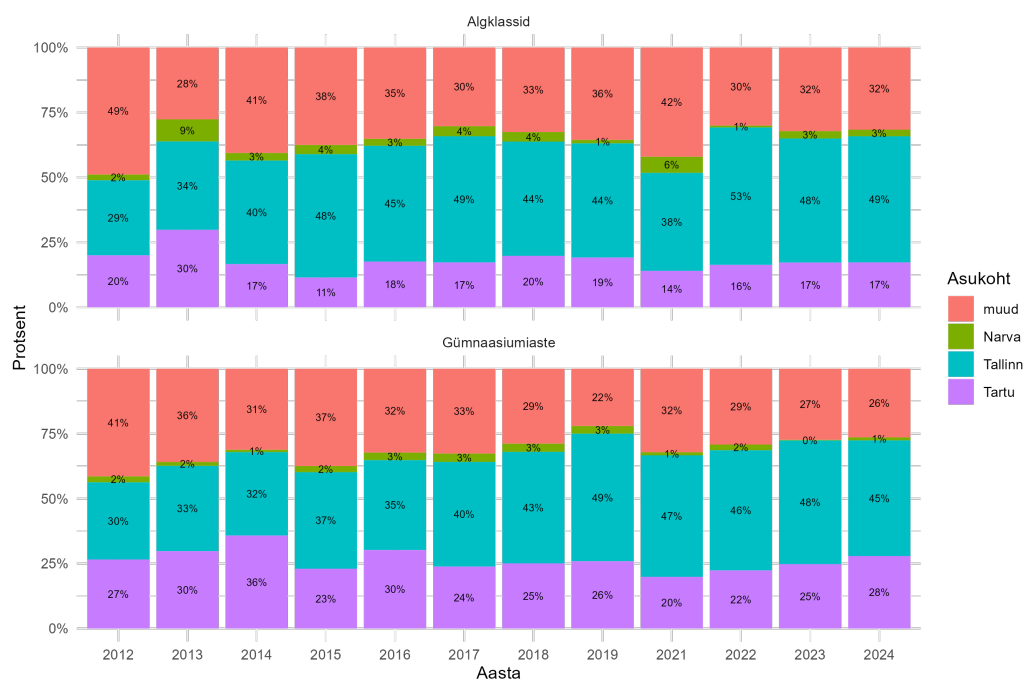
Joonis 2: TOP 200 hulka kuulunud osalemiste jaotus koolitüübi järgi algklassides ja gümnaasiumiastmes aastate lõikes.

Joonise 2 põhjal moodustavad tavakoolide õpilased mõlemas kooliastmes igal aastal enamuse TOP 200 hulka kuulunud osalemistest. Algklassides jääb tavakoolide osakaal enamasti ligikaudu 75% – 80% vahele ning eliitkoolide osakaal ligikaudu 20% – 25% vahele. Üks erandlikum aasta on 2021, kus eliitkoolide osakaal on algklassides väiksem, ligikaudu 14%.

Gümnaasiumiastmes on eliitkoolide osakaal üldiselt veidi suurem ja kõigub ligikaudu 17% kuni 32% vahel. Mitmel aastal moodustavad eliitkoolide õpilased ligikaudu veerandi või isegi peaaegu kolmandiku TOP 200 hulka kuulunud osalemistest. See näitab, et kuigi tavakoolide õpilased moodustavad arvuliselt enamuse, on eliitkoolide osakaal TOP 200 hulgas nende koolide väikest arvu arvestades märkimisväärne.

Kooliastmete võrdlus näitab, et eliitkoolide osakaal on gümnaasiumiastmes sageli suurem kui algklassides. See võib viidata sellele, et tugevama akadeemilise või valikulise taustaga koolide roll muutub kõrgemas kooliastmes nähtavamaks. Samas tuleb seda tõlgendada ettevaatlikult, sest koolitüübi tunnus ei kirjelda üksnes koolikeskkonda, vaid võib peegeldada ka õpilaste varasemat valikut ja koolide vastuvõtutingimusi.

Kui liikuda edasi koolikeskkonnalt laiemasse konteksti, siis kerkib esile ka geograafiline tegur, millistes linnades need õpilased õpivad ning kuidas see seostub nende esindatusega TOP 200 hulgas.



Joonis 3: TOP 200 hulka kuulunud osalemiste jaotus kooli asukoha järgi algklassides ja gümnaasiumiastmes aastate lõikes.

Joonise 3 põhjal on näha, et TOP 200 hulka kuulunud osalemiste jaotus kooli asukoha järgi erineb kooliastmeti. Algklassides moodustavad vaadeldava pe-

rioodi esimesel poolel suure osa osalemistest muude asulate koolide õpilased. Nende osakaal ulatub mõnel aastal üle 40%, kuid perioodi lõpu poole jääb see pigem ligikaudu 30% lähedale. Samal ajal on Tallinna koolide osakaal algklassides üldiselt suurenenud ning ulatub viimastel aastatel ligikaudu poole TOP 200 hulka kuulunud osalemistest. See viitab sellele, et Tallinna koolide esindatus parimate seas on ajas kasvanud ka algklasside tasemel.

Tartu osakaal jääb algklassides enamasti ligikaudu 15% – 20% vahele ning on seetõttu võrreldes Tallinnaga väiksem, kuid siiski püsiv. Narva osakaal on algklassides väike ja kõigub mõne protsendi ümber. Seega on algklassides näha ühelt poolt Tallinna osakaalu kasvu ning teiselt poolt muude asulate osakaalu vähenemist. See võib viidata sellele, et edukate osalemiste koondumine suurematesse keskustesse algab juba varasemates kooliastmetes, kuigi muude asulate roll on algklassides endiselt märgatav.

Gümnaasiumiastmes on Tallinna koolide osakaal veelgi suurem. Mitmel aastal moodustavad Tallinna koolide õpilased ligikaudu 40% – 48% TOP 200 hulka kuulunud osalemistest. Tartu osakaal on gümnaasiumiastmes samuti märgatav ning jääb enamasti ligikaudu veerandi ümbrusesse. Võrreldes algklassidega on muude asulate osakaal gümnaasiumiastmes väiksem ning jääb viimastel aastatel ligikaudu 27% – 32% vahele. Narva osakaal on mõlemas kooliastmes väike, kuid gümnaasiumiastmes on see eriti tagasihoidlik.

Selline jaotus viitab sellele, et kõrgemas kooliastmes koondub TOP 200 hulka jõudmine rohkem suurematesse keskustesse, eelkõige Tallinna ja Tartusse. Algklassides on muude asulate koolide osakaal veel üsna suur, kuid gümnaasiumiastmes väheneb see osakaal märgatavalt. Ühe võimaliku seletusena võib välja tuua koolivõrgu ja õppimisvõimaluste erinevused: suuremates linnades on rohkem koole, suurem õpilaste arv ning sageli ka rohkem võimalusi võistlusteks valmistumiseks. Samuti võib rolli mängida see, et osa tugevamaid

õpilasi liigub hilisemates kooliastmetes suurematesse gümnaasiumidesse.

Asukohtadevahelised erinevused ei pruugi siiski tuleneda ainult linna suuruselt või koolide arvust. Oluline võib olla ka kooli ja õpetajate poolne huvi ning aktiivsus õpilaste võistlustel osalemise toetamisel. Kui mõnes koolis pööratakse matemaatikavõistlustele rohkem tähelepanu, suunatakse õpilasi aktiivsemalt osalema või pakutakse neile paremat ettevalmistust, võib see kajastuda ka TOP 200 hulka jõudnud osalemiste jaotuses. Seetõttu tuleb asukoha järgi ilmnevaid erinevusi tõlgendada kui koolide, õppimisvõimaluste ja osalemisaktiivsuse omavaheliste seoste võimalikku kirjeldavat mustrit, mitte otsest piirkondlikku seletust.

Samas tuleb rõhutada, et tegemist on kirjeldava analüüsiga ning nende erinevuste põhjal ei saa teha põhjuslikke järeldusi. Joonis näitab, kuidas TOP 200 hulka kuulunud osalemised jagunevad kooli asukoha järgi, kuid ei võimalda eraldi hinnata, kuivõrd erinevused tulenevad koolide arvust, õpilaste liikumisest, õpetajate aktiivsusest, koolitüübist või muudest teguritest.

3.2 Edukuse püsimise tõenäosus

Igal aastal osaleb Känguru matemaatikavõistlusel suur hulk õpilasi, kuid osalemine ja edukus ei pruugi olla kogu koolitee jooksul järjepidevad. Õpilane võib saavutada häid tulemusi juba algklassides, kuid see ei tähenda tingimata, et ta jõuab sarnaste tulemusteni ka hilisemates kooliastmetes. Seetõttu on oluline vaadata, kui suur osa varakult silma paistnud õpilastest jõuab TOP 200 hulka ka gümnaasiumiastmes.

Selles analüüsis käsitletakse õpilast ühe korra ehk unikaalse isikuna. See tähendab, et kui õpilane jõudis algklassides TOP 200 hulka mitmel aastal, arvestatakse teda selles alapeatükis siiski ühe õpilasena. Seega ei analüüsita

siin TOP 200 hulka jõudmise kordade arvu, vaid seda, kas algklassides edukas olnud õpilane jõudis hiljem TOP 200 hulka ka gümnaasiumiastmes.

Olgu N algklassides TOP 200 hulka jõudnud õpilaste arv ning $N(A)$ nende õpilaste arv, kes jõudsid TOP 200 hulka ka gümnaasiumiastmes. Sarnase tulemuse saavutamise tõenäosus hinnatakse statistilise tõenäosusena ehk suhtelise sagedusena (vt alapeatükk 2.2).

Andmete põhjal kuulus vaadeldaval perioodil algklassides TOP 200 hulka kokku 2601 õpilast. Neist 202 jõudis TOP 200 hulka ka gümnaasiumiastmes. Seega on tõenäosus, et algklassides TOP 200 hulka jõudnud õpilane saavutab sarnase tulemuse ka gümnaasiumis,

$$P(A) = \frac{202}{2601} \approx 0,078 \quad (\text{vt valem (1)}).$$

See tähendab, et ligikaudu 7,8% algklassides TOP 200 hulka jõudnud õpilastest jõuab TOP 200 hulka ka gümnaasiumiastmes. Teisisõnu, enamik varakult häid tulemusi saavutanud õpilastest ei saavuta hiljem sama taset või ei osale enam gümnaasiumiastmes võistlusel.

Üldine tõenäosus näitab, et varajane edu kandub edasi vaid väikese osa õpilaste puhul. Samas ei ole kõik õpilasrühmad tingimata samas olukorras. Juba kirjeldavas analüüsis ilmnes, et TOP 200 hulgas on poisse rohkem kui tüdrukuid, mistõttu on loomulik vaadata, kas erinevus säilib ka siis, kui uurida varajase edu püsimist gümnaasiumiastmeni.

Tabel 1: Algklasside TOP 200 õpilaste tõenäosus jõuda TOP 200 hulka gümnaasiumis soo lõikes

| Sugu | Õpilaste arv | Jõudis TOP 200 hulka | Tõenäosus |
|----------|--------------|----------------------|-----------|
| Poisid | 1462 | 142 | 0,097 |
| Tüdrukud | 1139 | 60 | 0,053 |

Tabelist 1 selgub, et poiste puhul on tõenäosus jõuda TOP 200 hulka ka gümnaasiumiastmes 9,7%, samas kui tüdrukute puhul on see 5,3%. Seega on poistel ligikaudu 1,8 korda suurem tõenäosus säilitada sarnane edukus gümnaasiumini. See tulemus on kooskõlas varasema tähelepanekuga, et poiste osakaal TOP 200 hulgas on üldiselt suurem. Siin lisandub sellele aga ajamõõde: erinevus ei ilmne ainult ühel kooliastmel, vaid püsib ka siis, kui vaadelda varajase edu jätkumist.

Soo kõrval võib varajase edu püsimist mõjutada ka koolikeskkond. Kui õpilane õpib koolis, kus matemaatikale pööratakse rohkem tähelepanu, kus konkurents on tugevam või kus võistlusteks valmistumist toetatakse süsteemsemalt, võib see suurendada tõenäosust, et algklassides saavutatud edu jätkub ka hiljem.

Tabel 2: Algklasside TOP 200 õpilaste tõenäosus jõuda TOP 200 hulka gümnaasiumis koolitüübi lõikes

| Koolitüüp | Õpilaste arv | Jõudis TOP 200 hulka | Tõenäosus |
|-----------|--------------|----------------------|-----------|
| Tavakool | 2071 | 138 | 0,067 |
| Eliitkool | 530 | 64 | 0,121 |

Eliitkoolide õpilaste puhul on tõenäosus jõuda TOP 200 hulka ka gümnaa-

siumiastmes 12,1%, tavakoolide õpilaste puhul aga 6,7% (vt tabel 2). See tähendab, et eliitkoolide õpilaste tõenäosus säilitada sarnane edukus on ligikaudu 1,8 korda suurem. Tulemus sobitub eelmises alapeatükis kirjeldatud mustriaga, mille järgi eliitkoolide õpilased on TOP 200 hulgas oma koolide väikest arvu arvestades ülesindatud. Siin on näha, et erinevus ei piirdu ainult TOP 200 üldise jaotusega, vaid avaldub ka edukuse püsimises ajas.

Soo ja koolitüübi seoseid eraldi vaadates ilmnesid mõlema tunnuse puhul selged erinevused. Samas ei ole need tunnused tegelikkuses üksteisest täielikult eraldiseisvad, sest koolikeskkond ja õpilaste valik võivad poiste ja tüdrukute puhul erineda. Seda arvestades on huvitav vaadata, kuidas muutub tõenäosus siis, kui vaadelda sugu ja koolitüüpi samaaegselt.

Tabel 3: Alklasside TOP 200 õpilaste tõenäosus jõuda TOP 200 hulka gümnaasiumis soo ja koolitüübi lõikes

| Sugu ja koolitüüp | Õpilaste arv | Jõudis TOP 200 hulka | Tõenäosus (%) |
|---------------------|--------------|----------------------|---------------|
| Poisid, tavakool | 1150 | 96 | 8,3 |
| Poisid, eliitkool | 312 | 46 | 14,7 |
| Tüdrukud, tavakool | 921 | 42 | 4,6 |
| Tüdrukud, eliitkool | 218 | 18 | 8,3 |

Tabelist 3 selgub, et eliitkoolis õppimine on seotud suurema tõenäosusega säilitada varajane edu nii poiste kui ka tüdrukute puhul, kuid erinevuse suurus on erinev.

Poiste puhul suureneb tõenäosus tavakooliga võrreldes 8,3% pealt 14,7%-ni, mis tähendab ligikaudu 6,4 protsendipunkti suurus kasv. Tüdrukute puhul on vastav kasv 4,6%-lt 8,3%-ni ehk 3,7 protsendipunkti. Seega on koolitüübiga seotud erinevus mõlema soo puhul märkimisväärne, kuid poiste puhul

tugevam. Veelgi selgemalt tuleb erinevus esile äärmusi võrreldes. Kui võrrelda poisse eliitkoolides ja tüdrukuid tavakoolides, siis on tõenäosus vastavalt 14,7% ja 4,6%, mis tähendab enam kui kolmekordset erinevust.

Huvitav on ka see, et sooline erinevus ei kao ka eliitkoolides. Kuigi mõlema soo puhul tõenäosus suureneb, on poiste tulemus endiselt kõrgem kui tüdrukutel. Samas tavakoolides on vahe veidi väiksem, mis viitab sellele, et koolikeskkond võib olla seotud soolise erinevuse avaldumisega. Selline muster näitab, et varajase edu püsimine ei erine ainult ühe tunnuse lõikes, vaid seda tuleb vaadelda mitme tunnuse taustal.

Kuna eelmises alapeatükis ilmnesis erinevused ka piirkondade vahel, eriti Tallinna, Tartu ja Narva võrdluses, on oluline vaadata, kas geograafiline erinevus kajastub ka varajase edu jätkumises.

Tabel 4: Alklasside TOP 200 õpilaste tõenäosus jõuda TOP 200 hulka gümnaasiumis asukoha lõikes

| Asukoht | Õpilaste arv | Jõudis TOP 200 hulka | Tõenäosus |
|---------|--------------|----------------------|-----------|
| Tallinn | 1153 | 86 | 0,075 |
| Tartu | 436 | 46 | 0,106 |
| Narva | 86 | 3 | 0,035 |
| Muud | 926 | 67 | 0,072 |

Tabelist 4 selgub, et kõrgeim tõenäosus jõuda TOP 200 hulka ka gümnaasiumiastmes on Tartu koolide õpilastel, kelle puhul on vastav tõenäosus 10,6%. Tallinna koolide õpilaste puhul on tõenäosus 7,5% ning muude asulate puhul 7,2%. Narva puhul on tõenäosus kõige madalam, 3,5%, kuid selle tulemuse tõlgendamisel tuleb arvestada väikese vaatluste arvuga.

See tulemus täiendab eelmises alapeatükis tehtud tähelepanekut Tartu suh-

teliselt tugeva esindatuse kohta. Tartu ei paista silma ainult TOP 200 üldises jaotuses, vaid ka selles, et algklassides edukad õpilased jõuavad seal sagedamini TOP 200 hulka ka gümnaasiumiastmes. Tallinna puhul on pilt teistsugune: kuigi Tallinna osakaal TOP 200 hulgas on aastate jooksul suurenenud, ei ole varajase edu püsimise tõenäosus seal kõige kõrgem. Narva madal tulemus viitab omakorda sellele, et lisaks väiksemale üldisele esindatusele on seal ka varajase edu jätkumine harvem.

Kirjeldav analüüs näitab seega, et varajase edu püsimine erineb nii soo, koolitüübi kui ka kooli asukoha lõikes. Samas ei võimalda selline võrdlus veel öelda, kas need tunnused on tulemusega seotud iseseisvalt või peegeldavad osaliselt üksteisega seotud nähtusi. Näiteks võib koolitüübi seos olla osaliselt seotud kooli asukohaga ning asukoha seos võib omakorda peegeldada erinevusi koolikeskkonnas.

Seetõttu kasutatakse järgmises alapeatükis logistilist regressioonimudelit, mis võimaldab hinnata soo, koolitüübi ja kooli asukoha seost tõenäosusega jõuda TOP 200 hulka ka gümnaasiumiastmes, arvestades neid tegureid samaaegselt.

3.3 Logistiline regressioonanalüüs

Eelmises alapeatükis ilmnis, et varajase edu püsimine erineb nii soo, koolitüübi kui ka kooli asukoha lõikes ning lisaks võivad need tunnused olla omavahel seotud, tekitades märkimisväärseid erinevusi õpilaste edasises edukuses. Samas ei võimalda kirjeldav analüüs eristada, kas tegemist on iseseisvate seostega või peegeldavad need tunnused üksteisega seotud nähtusi, mistõttu on vajalik hinnata nende seost tulemusega samaaegselt.

Mudelis vaadeldakse tõenäosust, et õpilane, kes kuulus algklassides TOP 200 hulka, jõuab TOP 200 hulka ka gümnaasiumiastmes. Tähistades selle sünd-

muse muutujaga Y , kus $Y = 1$ tähendab TOP 200 hulka jõudmist gümnaasiumis ja $Y = 0$ vastupidist, modelleeritakse tõenäosust $p = P(Y = 1)$ logistilise regressiooni abil:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 \cdot \text{tüdruk} + \beta_2 \cdot \text{eliitkool} \\ + \beta_3 \cdot \text{Narva} + \beta_4 \cdot \text{Tallinn} + \beta_5 \cdot \text{Tartu} \quad (\text{vt valem (2)}).$$

Mudelis on kõik selgitavad tunnused kodeeritud indikaatortunnustena. Tunnus *tüdruk* saab väärtuse 1 juhul, kui õpilane on tüdruk, ja 0 juhul, kui õpilane on poiss. Tunnus *eliitkool* saab väärtuse 1 juhul, kui õpilane õppis eliitkoolis, ja 0 juhul, kui õpilane õppis tavakoolis. Asukoha puhul kasutatakse kolme indikaatortunnust: *Narva*, *Tallinn* ja *Tartu*. Vastav tunnus saab väärtuse 1 juhul, kui kool asub selles linnas, ja 0 vastasel juhul. Kui kõik kolm asukoha indikaatortunnust on väärtusega 0, kuulub õpilane võrdlusgruppi ehk muude asulate koolide hulka.

Seega on mudeli võrdlusgrupiks poisid, tavakoolide õpilased ja muude asulate koolide õpilased. Mudeli koefitsiendid näitavad, kuidas muutub gümnaasiumiastmes TOP 200 hulka jõudmise šanss võrreldes selle võrdlusgrupiga.

Tabel 5: Logistilise regressiooni tulemused gümnaasiumiastmes TOP 200 hulka jõudmise kohta

| Muutuja | Hinnang (β) | Šansisuhe (OR) | p-väärtus |
|---------------|---------------------|----------------|-----------|
| (Intercept) | -2,141 | – | < 0,001 |
| Sugu (tüdruk) | -0,681 | 0,51 | < 0,001 |
| Eliitkool | 0,632 | 1,88 | 0,001 |
| Narva | -0,944 | 0,39 | 0,116 |
| Tallinn | -0,485 | 0,62 | 0,011 |
| Tartu | -0,018 | 0,98 | 0,932 |

Tabelist 5 on näha, et mudeli tulemused kinnitavad varasemas analüüsis ilmnunud mustreid. Šansisuhe (OR) näitab, mitu korda muutub sündmuse toimumise šanss võrreldes võrdlusgrupiga. Soo seos tulemusega on selgelt eristatav: tüdrukute šanss jõuda TOP 200 hulka gümnaasiumiastmes on väiksem kui poistel. Šansisuhe 0,51 tähendab, et tüdrukute šanss on ligikaudu 49% väiksem ning erinevus on statistiliselt väga tugev ($p < 0,001$).

Koolitüübi seos on veelgi selgem: eliitkoolide õpilaste šanss jõuda TOP 200 hulka on ligikaudu 1,9 korda suurem kui tavakoolide õpilastel (OR = 1,88, $p = 0,001$), mis haakub eelmises alapeatükis nähtud erinevustega ja näitab, et koolitüübiga seotud erinevus püsib ka teiste tegurite arvesse võtmisel.

Asukoha seos on mitmetahulisem. Tallinna koolide puhul on šanss väiksem võrreldes muude asulate koolidega (OR = 0,62, $p = 0,011$), mis tähendab ligikaudu 38% väiksemat šanssi. See tulemus on kooskõlas varasema tähelepanekuga, et kuigi Tallinna osakaal TOP 200 hulgas on suur, ei ole varajase edu püsimine seal kõige kõrgem. Narva ja Tartu puhul ei ole seos statistiliselt oluline, mis tähendab, et nende erinevused võivad olla tingitud juhuslikust

kõikumisest või väiksemast vaatluste arvust.

Kokkuvõttes näitab analüüs, et gümnaasiumiastmes TOP 200 hulka jõudmise tõenäosus on seotud nii soo, koolitüübi kui ka osaliselt kooli asukohaga. Samas tuleb arvestada, et tegemist on vaatlusandmetega ning saadud seoseid ei saa tõlgendada põhjuslike seostena.

3.4 Tulemuste arutelu

Analüüsi tulemused näitavad, et TOP 200 hulka jõudmine gümnaasiumiastmes on seotud nii soo kui ka koolitüübiga. Logistilise regressiooni põhjal on need seosed statistiliselt olulised, mis viitab sellele, et tegemist ei ole juhuslike erinevustega.

Sooline erinevus on üks kesksemaid tulemusi. Tüdrukute šanss jõuda TOP 200 hulka on väiksem kui poistel ka siis, kui arvesse võtta koolitüüpi ja asukohta. Kuna erinevus püsib ka regressioonimudelis, ei saa seda seletada ainult koolikeskkonna või asukoha kaudu.

Oluline on rõhutada, et seda tulemust ei saa tõlgendada kui erinevust võimetes. Varasemad uuringud on näidanud, et poiste ja tüdrukute matemaatilised võimed on keskmiselt väga sarnased (Else-Quest, Hyde ja Linn, 2010). Seetõttu tuleb otsida selgitusi pigem teistest teguritest.

Ühe võimaliku seletusena võib välja tuua erinevused motivatsioonis ja enesekindluses. Samuti võib rolli mängida see, kuidas õpilasi suunatakse, näiteks kas neid julgustatakse osalema võistlustel või mitte. Ei saa välistada ka laiemaid hoiakuid, kus matemaatikat peetakse rohkem poiste valdkonnaks. Sellised tegurid võivad olla seotud sellega, kes jätkab osalemist ja kes mitte. Koolitüübi seos on samuti selgelt nähtav. Eliitkoolide õpilaste šanss jõuda TOP 200 hulka on oluliselt suurem. Seda võib osaliselt selgitada valikuga,

sest nendesse koolidesse pääsevad sageli juba tugevamad õpilased. Samas võib rolli mängida ka keskkond, näiteks tugevam konkurents või paremad võimalused valmistuda võistlusteks.

Asukoha seost tuleb tõlgendada ettevaatlikumalt. Kuigi kirjeldavas analüüsis ja tõenäosustes ilmnemised erinevused piirkondade vahel, ei olnud kõik asukohaga seotud seosed regressioonis statistiliselt olulised. See tähendab, et nende erinevuste kohta ei saa teha kindlaid järeldusi. Samas võib täheldada, et teatud piirkondades on edukuse püsimine suurem kui teistes, mis võib olla seotud erinevustega koolides, õpetamises või osalemisvõimalustes. Osaliselt võivad erinevused asukohtade vahel olla seletatavad ka erinevustega eliit- ja tavakoolide vahel, sest eliitkoolid paiknevad peamiselt Tallinnas ja Tartus.

Lisaks näitas analüüs, et varajane edu kandub edasi vaid vähestel juhtudel. See viitab sellele, et edukus matemaatikavõistlustel ei ole püsiv, vaid sõltub aja jooksul muutuvatest teguritest. Võimalik, et osa õpilasi kaotab huvi või valib teise suuna, samas kui teised jätkavad ja arenevad edasi.

Kokkuvõttes viitavad tulemused sellele, et õpilaste edukus matemaatikavõistlustel ei ole seotud ainult individuaalsete võimetega, vaid sõltub ka keskkonnast ja võimalustest. Eriti soolise erinevuse puhul on tõenäoline, et rolli mängivad sotsiaalsed ja hariduslikud tegurid, mitte võimete erinevus.

Kokkuvõte

Käesoleva bakalaureusetöö eesmärk oli uurida, kas algklassides matemaatikavõistlustel edukad õpilased jõuavad TOP 200 hulka ka gümnaasiumiastmes ning millised tunnused on selle tulemusega seotud.

Töös analüüsiti Känguru matemaatikavõistluse andmeid aastatest 2012–2024. Edukuse kirjeldamiseks ja selgitamiseks kasutati kolme lähenemist: kirjeldavat analüüsi, statistilise tõenäosuse hindamist ning logistilist regressiooni.

Kirjeldav analüüs näitas, et TOP 200 hulgas on poisid ülesindatud ning eliitkoolide õpilaste osakaal on nende koolide väikest arvu arvestades suur. Samuti ilmneseid erinevused kooli asukoha lõikes.

Statistilise tõenäosuse analüüsist selgus, et varajane edu kandub edasi vaid vähestel juhtudel. Ligikaudu 7,8% algklassides TOP 200 hulka jõudnud õpilastest jõuab TOP 200 hulka ka gümnaasiumiastmes. Samuti ilmnese, et poisitel ja eliitkoolide õpilastel on suurem tõenäosus säilitada varajane edukus.

Logistilise regressiooni tulemused kinnitasid, et nii sugu kui ka koolitüüp on statistiliselt olulisel määral seotud TOP 200 hulka jõudmisega gümnaasiumiastmes. Tüdrukute šanss on väiksem võrreldes poistega, samas kui eliitkoolide õpilaste šanss on suurem võrreldes tavakoolide õpilastega. Kooli asukoha seos ei olnud kõigi rühmade puhul statistiliselt oluline.

Töö tulemused näitavad, et matemaatikavõistlustel edukuse püsimine ei ole seotud ainult varasema tulemusega, vaid ka soo, koolitüübi ja osaliselt kooli asukohaga.

Edasistes uuringutes võiks kasutada täpsemaid andmeid ning täiendavaid tunnuseid, mis võimaldaksid paremini selgitada täheldatud erinevusi ning hinnata erinevate tunnuste seoseid täpsemalt.

Kasutatud allikad

- 98th Percentile (2023). *Impact of Kangaroo Math on Students*. URL: <https://www.98thpercentile.com/blog/impact-of-kangaroo-math-usa-on-students/> (vaadatud 06.04.2025).
- Agresti, Alan (2018). *Statistical Methods for the Social Sciences*. 5. väljaanne. Pearson.
- Duncan, Greg J., Chantelle J. Dowsett, Amy Claessens, Katherine Magnuson, Aletha C. Huston, Pamela Klebanov, Linda S. Pagani, Leon Feinstein, Mimi Engel, Jeanne Brooks-Gunn, Holly Sexton, Kathryn Duckworth ja Crista Japel (2007). "School readiness and later achievement". *Developmental Psychology* 43.6, lk. 1428–1446. DOI: [10.1037/0012-1649.43.6.1428](https://doi.org/10.1037/0012-1649.43.6.1428).
- Else-Quest, Nicole M., Janet Shibley Hyde ja Marcia C. Linn (2010). "Cross-national patterns of gender differences in mathematics: A meta-analysis". *Psychological Bulletin* 136.1, lk. 103–127. DOI: [10.1037/a0018053](https://doi.org/10.1037/a0018053).
- Hosmer, David W., Stanley Lemeshow ja Rodney X. Sturdivant (2013). *Applied Logistic Regression*. 3. väljaanne. Hoboken, NJ: John Wiley & Sons.
- OECD (2018). *Equity in Education: Breaking Down Barriers to Social Mobility*. PISA. Paris: OECD Publishing. DOI: [10.1787/9789264073234-en](https://doi.org/10.1787/9789264073234-en).
- Ross, Sheldon M. (2010). *Introduction to Probability Models*. 10. väljaanne. Burlington, MA: Academic Press.
- Statistikaamet (2026). *Nimede statistika*. URL: <https://stat.ee/nimed> (vaadatud 01.06.2026).
- Tartu Ülikooli Teaduskool (2024). *Känguru võistluse arhiiv*. URL: <https://teaduskool.ut.ee/et/ainevoistlused/kanguru-voistluse-arhiiv> (vaadatud 08.04.2025).

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Darja Rohi,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose "Matemaatikavõistlustel osalemist ja edukust mõjutavad tegurid", mille juhendaja on PhD Krista Fischer, reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.
2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.
3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.
4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Darja Rohi

10.06.2026