

Matematiikan ja operaatiotutkimuksen maisteriohjelma

Yliopistomatematiikan arvosanat ja opiskelijasta riippumattomat tekijät

Tutkimus Aalto-yliopiston kandidaattitason matematiikan perus- ja jatkokursseista

Kristian Jakobsson

© 2025

Tämä teos on lisensoitu [Creative Commons](#) “Nimeä-EiKaupallinen-JaaSamoin 4.0 Kansainvälinen” -
käyttöluvalla.



Tekijä Kristian Jakobsson

Työn nimi Yliopistomatematiikan arvosanat ja opiskelijasta riippumattomat tekijät
— Tutkimus Aalto-yliopiston kandidaattitason matematiikan perus- ja
jatkokursseista

Koulutusohjelma Teknillinen fysiikka ja matematiikka

Pääaine Matematiikka ja operaatiotutkimus

Työn valvoja Pekka Pere

Työn ohjaajat Pekka Pere, Pekka Pere

Yhteistyötaho Aalto-yliopiston johtamisen tietopalvelu

Päivämäärä 31. heinäkuuta 2025

Sivumäärä 49+5

Kieli suomi

Tiivistelmä

Tämä tutkielma selvittää, miten opiskelijan sukupuoli, syntymäkuukausi, kansalaisuus, korkeakoulu ja arvosanainflaatio vaikuttavat Aalto-yliopiston matematiikan kurssien arvosanoihin vuosina 2014–2024. Analyysissä tarkastellaan sekä perus- että syventäviä kandidaattikursseja hyödyntäen logistista regressiota ja osittain verrannollisten vastasuhteiden malleja. Tulokset osoittavat, että sukupuolella, kansalaisuudella ja arvosanainflaatiolla on yhteys suoriutumiseen, kun taas syntymäkuukauden eli suhteellisen iän vaikutus on rajallinen. Työssä tutkitaan myös arvosanainflaation kehitystä. Tulokset auttavat hahmottamaan rakenteellisia eroja ja opiskelumenestykseen vaikuttavia tekijöitä korkeakoulumatematiikassa.

Avainsanat logistinen regressio, suhteellinen ikävaikutus, yliopistomatematiikka, sukupuolierot matematiikassa, osittain verrannollisten vastasuhteiden malli, Aalto-yliopisto, arvosanainflaatio

Author Kristian Jakobsson

Title University Mathematics Grades and Demographic Effects — A Study of Aalto University’s bachelor’s level basic and advanced mathematics courses

Degree programme Technical Physics and Mathematics

Major Mathematics and Operations Research

Supervisor and advisor Pekka Pere

Collaborative partner Management Information Services of Aalto University

Date 31 July 2025

Number of pages 49+5

Language Finnish

Abstract

This thesis investigates how sex, birth month, nationality, faculty, and grade inflation affect performance in university mathematics courses at Aalto University between 2014 and 2024. The analysis covers both basic and advanced bachelor’s level mathematics courses, using logistic regression and partial proportional odds models. The study finds that sex, nationality, and grade inflation influence performance, while the effects of birth month (i.e., the relative age effect) are limited. Grade inflation trends are also discussed. The findings shed light on patterns of academic performance in higher education mathematics.

Keywords cumulative logistic regression, logistic regression, relative age effect, university mathematics, sex differences in mathematics, partial proportional odds model, Aalto University, grade inflation

Författare Kristian Jakobsson

Titel Faktorer som påverkar vitsorden i universitetsmatematik: Demografiska skillnader, institutionella skillnader och vitsordsinflation — En studie av Aalto-universitetets grund- och fortsättningskurser på kandidatnivå i matematik

Utbildningsprogram Teknisk fysik och matematik

Huvudämne Matematik och systemanalys

Övervakare Pekka Pere

Handledare Pekka Pere

Samarbetspartner Aalto-universitetets informationsservice

Datum 31 July 2025

Sidantal 49+5

Språk finska

Sammandrag

Detta arbete undersöker hur kön, födelsemånad, medborgarskap, högskola och betygsinflation påverkar vitsorden i universitetsmatematik vid Aalto-universitetet under åren 2014–2024. Analysen omfattar både grundkurser och fördjupade kurser på kandidatnivå. Modellerna som används i analysen är logistisk regression och ordinal regression. Resultaten visar att kön, medborgarskap och betygsinflation påverkar studieprestationer, medan effekten av födelsemånad, d.v.s. den relativa ålderseffekten, är minimal. Studien belyser även utvecklingen av betygsinflation över tid. Resultaten ökar förståelsen för prestationsmönster och bakgrundsfaktorer inom universitetsmatematik.

Nyckelord ordinal logistisk regression, logistisk regression, relativ ålderseffekt, universitetsmatematik, könsskillnader i matematik, partiell proportional odds-modell, Aalto-universitetet, vitsordsinflation

Isoin kiitos kuuluu työn ohjaajalleni Pekka Perelle. Pekka on ollut inspiraationa alusta alkaen, tutkimuksen suunnittelemisesta, tilastollisten menetelmien ymmärtämiseen ja lopulta tekstin viimeistelyyn saakka.

Haluan myös kiittää Aalto-yliopiston johtamisen tietopalvelua yhteistyöstä ja luottamuksesta, erityisesti Riitta Anderssonia ja Jiri Lallimoa. Tietopalvelun luovuttama aineisto mahdollisti tämän tutkielman. Erityisesti arvostan sitä, että aineistoa täydennettiin työn aikana esimerkiksi arvosanainflaation tarkastelua varten.

Otaniemi, 31. heinäkuuta 2025

Kristian Jakobsson

Sisällysluettelo

Tiivistelmä	3
Tiivistelmä (englanniksi)	4
Tiivistelmä (ruotsiksi)	5
Sisällysluettelo	7
Symbolit ja lyhenteet	9
1 Johdanto	10
1.1 Tutkimuksen tavoite ja tutkimuskysymykset	10
1.2 Rajaukset	11
1.3 Tutkielman rakenne	11
2 Kirjallisuuskatsaus	12
2.1 Iän yhteys matematiikan arvosanoihin ja osaamiseen	12
2.2 Sukupuoli ja matematiikan arvosanat	12
2.3 Kansalaisuus ja matematiikan arvosanat	13
2.4 Arvosanat Aalto-yliopiston korkeakouluissa	13
2.5 Arvosananinflaation kehitys koulutusasteilla	14
3 Tutkimusaineisto	15
3.1 Aineiston kuvaus	15
3.2 Syntymäkuukausien jakauma ja vertailu väestöön	21
4 Menetelmät	22
4.1 Logistinen regressio	22
4.1.1 Mallin arviointi: Erottelukykykäyrä ja käyräalainen pinta-ala	23
4.2 Järjestyslogistinen regressiomalli	25
4.2.1 Poikkeamat verrannollisten vastasuhteiden oletuksesta . . .	25
4.2.2 Brantin testi: Verrannollisten vastasuhteiden oletukselle . .	25
4.2.3 Mallin selityskyky: Yhteensopivuusindeksi	26
4.3 Yhteinen diagnostiikka	27
4.3.1 Waldin testi	27
4.3.2 Uskottavuusosamäärätesti	27
4.3.3 Devianssi ja Pearsonin χ^2 -testi	27
4.3.4 Pearsonin korrelaatiokerroin ja McFaddenin pseudo- R^2 . . .	28
4.3.5 Akaiken informaatiokriteeri (AIC)	29
4.3.6 Standardoidut ja kumulatiiviset jäännökset	29
4.4 Yhteenvedo menetelmistä	29

5 Tulokset	31
5.1 Hyväksytyt ja hylätyt suoritukset	31
5.1.1 Peruskurssittain	31
5.1.2 Kaikkien kurssien hyväksymiseen vaikuttavat tekijät	34
5.2 Osittain verrannollisten vastasuhteiden malli	36
5.2.1 Peruskurssien arvosanatasoihin vaikuttavat tekijät	38
5.2.2 Peruskurssittain	38
5.2.3 Kaikkien kurssien arvosanatasoihin vaikuttavat tekijät	41
6 Pohdinta	44
6.1 Tulosten tulkinta ja yhteys kirjallisuuteen	44
6.2 Tulosten merkitys	44
6.3 Tutkimuksen rajoitukset ja luotettavuus	45
7 Johtopäätökset	46
7.1 Keskeiset havainnot	46
7.2 Jatkotutkimusaiheet	46
A Liite - Logistinen regressio	50
B Liite - Koodit	50
B.1 Osuustestit syntymäkuukausille (tammi–maaliskuu ja tammi–kesäkuu)	50
B.2 Logistinen regressio	51
B.3 Verrannollisten vastasuhteiden oletuksen testaus (Brantin testi)	52
B.4 Osittain verrannollisten vastasuhteiden mallit	53

Symbolit ja lyhenteet

Symbolit

$\pi(\mathbf{x})$	hyväksynnän todennäköisyys selittäjillä \mathbf{x}
β_0	logistisen/järjestyslogistisen mallin vakiotermi
β_j	j :nnen selittäjän regressiokerroin
C	Yhteensopivuusindeksi
$\overset{a}{\sim}$	Testisuure noudattaa jakaumaa asympotoottisesti
z	Waldin testisuure $z = \frac{\hat{\beta}_j}{\sigma(\hat{\beta}_j)}$
χ^2	Pearsonin χ^2 -testisuure
G^2	devianssi
χ^2_{df}	χ^2 -jakauma vapausasteilla df

Matemaattiset merkinnät

$\log()$	luonnollinen logaritmi
$\text{logit}(\pi(\mathbf{x}))$	logaritmi todennäköisyyden $\pi(\mathbf{x})$ vastasuhteesta
$\exp()$	eksponenttifunktio

Lyhenteet

AIC	Akaiken informaatiokriteeri
AUC	käyräalainen pinta-ala (<i>area under the ROC curve</i>)
C-indeksi	yhteensopivuusindeksi (<i>concordance index</i>)
df	vapausasteet (<i>degrees of freedom</i>)
GDPR	EU:n yleinen tietosuoja-asetus
GPA	arvosanaindeksi
KARVI	Kansallinen koulutuksen arviointikeskus
UOM-testi	Uskottavuusosamäärätesti
PISA	Kansainvälinen oppimistulosten arviointiohjelma
SCI	Aalto-yliopiston Perustieteiden korkeakoulu
ENG	Aalto-yliopiston Insinööritieteiden korkeakoulu
ELEC	Aalto-yliopiston Sähkötekniikan korkeakoulu
CHEM	Aalto-yliopiston Kemian tekniikan korkeakoulu
ARTS	Aalto-yliopiston Taiteiden ja suunnittelun korkeakoulu
BIZ	Aalto-yliopiston Kauppakorkeakoulu

1 Johdanto

1.1 Tutkimuksen tavoite ja tutkimuskysymykset

Miten syntymäkuukausi, sukupuoli ja kansalaisuus vaikuttavat yliopistotasosten matematiikan kurssien arvosanoihin? Näitä kysymyksiä selvitetään pureutumalla Aalto-yliopiston laajaan opiskelija-aineistoon viimeisten 11. vuoden pohjalta. Lisäksi tarkastellaan, esiintyykö eroja yliopiston eri korkeakoulujen välillä ja miten matematiikan arvosanat ovat kehittyneet viimeisen vuosikymmenen aikana mahdollisen arvosanainflaation näkökulmasta.

Tutkimuksessa sovelletaan kahta regressiomallia: logistista regressiota (logistic regression) ja järjestyslogistista regressiomallia (ordinal regression). Logistinen regressiomalli mahdollistaa hyväksytyjen ja hylättyjen suoritusten erottelun. Järjestyslogistinen malli ottaa huomioon myös hyväksytyjen arvosanojen järjestyksen, esimerkiksi erottelun välttävien ja hyvien arvosanojen välillä, sekä hylätyt suoritukset. Tämä mahdollistaa tarkemman analyysin hyväksytyjen arvosanojen jakautumisesta.

Vaikka matematiikan arvosanoja ja oppimistuloksia on tutkittu runsaasti perus- ja toisen asteen koulutuksessa, yliopistotason tutkimus on vähäisempää. Erityisesti suhteellisen iän vaikutuksia (jatkossa iän vaikutusta), eli opiskelijan omaa ikää suhteessa muihin oman ikäluokan sisällä. Korkeakoulutasolla tätä on tarkasteltu vain harvoissa tutkimuksissa, vaikka aiemmat kansainväliset aineistot osoittavat sen merkityksen jo varhaisiässä. Tämä tutkimuksellinen aukko on merkittävä, sillä on epäselvää, vahvistuuko iän vaikutus ajan myötä matematiikan arvosanoissa, kuten esimerkiksi urheilussa (R. H. Barnsley, Thompson ja P. E. Barnsley 1985), vai heikkeneekö se kehityserojen tasoittuessa.

Tämän lisäksi, aiemmassa tutkimuskirjallisuudessa järjestyslogistista mallia ei ole juurikaan hyödynnetty korkeakoulukontekstissa yhdessä hylättyjen suoritusten kanssa. Mallit tarjoavat kuvan siitä, miten eri muuttujat, esimerkiksi suhteellinen ikä, sukupuoli ja kansalaisuus vaikuttavat matematiikan kurssimenestykseen yliopistotasolla, hyväksytyistä arvosanoista aina hylättyihin suorituksiin saakka.

Tutkimus pyrkii vastaamaan seuraaviin tutkimuskysymyksiin:

- Onko syntymä-ajalla yhteyttä opiskelijoiden matematiikan arvosanoihin?
- Eroavatko matematiikan arvosanat sukupuolten välillä?
- Eroavatko suomalaisten opiskelijoiden ja kansainvälisten opiskelijoiden matematiikan arvosanat?
- Esiintyykö arvosanallisia eroja Aalto-yliopiston eri korkeakoulujen välillä?
 - ARTS – Taiteiden ja suunnittelun korkeakoulu
 - BIZ – Kauppakorkeakoulu
 - CHEM – Kemian tekniikan korkeakoulu
 - ELEC – Sähkötekniikan korkeakoulu

- ENG – Insinööritieteiden korkeakoulu
 - SCI – Perustieteiden korkeakoulu
- Miten matematiikan arvosanjakaumat ovat muuttuneet vuosien 2014–2024 aikana? Voidaanko havaita arvosanainflaatiota, eli yleistä nousua arvosanoissa?

1.2 Rajaukset

Tutkimus keskittyy Aalto-yliopiston kandidaattitason matematiikan opintokokonaisuuksiin, jotka edustavat pääasiassa soveltavaa matematiikkaa. Kurssivalikoimaa on rajattu siten, että mukaan on otettu vain kurssit, joilla on riittävästi suorituksia opiskelijoiden yksityisyyden suojaamisen takaamiseksi. Aineisto kattaa vuodet 2014–2024 ja sisältää tietoa opiskelijoiden syntymäkuukaudesta, sukupuolesta, kansalaisuudesta, korkeakoulukohtaisesta taustasta sekä kurssisuorituksista suoritusajankohtaan. Kurssi-suorituksiin sisältyy hyväksytyjen arvosanojen lisäksi myös hylättyjä suorituksia.

Aineisto ei sisällä tietoa opiskelijoiden aiemmasta koulumenestyksestä tai tarkasta syntymäpäivästä, joten analyysi perustuu kuukausitasoiseen syntymäaikaan eikä huomioi esimerkiksi iän vaikutusta vuotuisella tasolla.

1.3 Tutkielman rakenne

Tutkielma etenee seuraavasti: Ensimmäisessä luvussa esitellään tutkimuksen tausta, tavoitteet ja rajaukset. Toisessa luvussa tarkastellaan aiempaa tutkimusta liittyen syntymäkuukauden, sukupuolen ja kansalaisuuden yhteyksiin oppimistuloksiin sekä arvosanainflaation yleistä trendiä viimeisen kymmenen vuoden aikana. Kolmannessa luvussa kuvataan aineisto. Neljännessä luvussa avataan matemaattiset menetelmät, mukaan lukien järjestyslogistisen regressiomallin soveltaminen. Viidennessä luvussa esitetään tutkimustulokset. Kuudes luku sisältää tulosten pohdinnan ja kriittisen arvioinnin. Seitsemäs ja viimeinen luku kokoaa johtopäätökset ja esittää suosituksia jatkotutkimukselle.

2 Kirjallisuuskatsaus

2.1 Iän yhteys matematiikan arvosanoihin ja osaamiseen

Kansainväliset tutkimukset osoittavat iän yhteyden matematiikan suorituksiin eri koulutustasoilla ja -järjestelmissä. Ruotsissa tämä ilmeni jo vuonna 1995, kun ikäero näkyi selvästi neljäsluokkalaisten pistetasoissa. (Cliffordson ja Gustafsson 2010) Vastaavasti Saksassa havaittiin, että suhteellisesti vanhemmat oppilaat saavuttivat hieman korkeampia matematiikan pisteitä jo peruskoulun toisella luokalla. (Thoren, Heinig ja Brunner 2016) Sama havainto on tehty Turkissa, jossa ikäluokan nuorimmat oppilaat, erityisesti kahdeksannella luokalla, suoriutuivat heikommin matematiikan arvioinnissa, ja iän vaikutus selitti arvosanojen vaihtelua hyvin. (Ünal 2019)

Ilmiö on havaittu myös suomalaisessa perusopetuksessa. Tammikuussa syntyneet oppilaat saivat paremmat päättöarvosanat yhdeksännellä luokalla kuin juuri ennen vuodenvaihdetta syntyneet. (Kaila 2017) Karvi¹ raportoi saman ilmiön vuosien 2018–2020 välillä: ikäluokan vanhimmat kolmasluokkalaiset saavuttivat selvästi paremmat osaamispisteet matematiikassa kuin heidän nuoremmat ikätoverinsa. (Ukkola ja Metsämuuronen 2023) Myös neljännesvuosittainen vertailu osoitti, että vuoden ensimmäisellä kvartaalilla syntyneet suoriutuivat keskimäärin paremmin kuin viimeisellä kvartaalilla syntyneet. (Huusela ja Kilpi 2022)

Sekä kansainvälisesti, että Suomessa iällä on iso vaikutus matematiikan oppimistuloksiin peruskoulussa. Ikäluokan vanhemmat oppilaat suoriutuvat nuorempia oppilaita paremmin sekä osaamistesteissä, että saivat parempia arvosanoja.

2.2 Sukupuoli ja matematiikan arvosanat

Sukupuolten väliset erot matematiikan osaamisessa on kiistelty ja poliittisesti herkkä tutkimuskohde. (Pinker 2008, Luku 18) Kansainväliset meta-analyysit osoittavat, että sukupuolten välisten erot matematiikan osaamisessa ovat vaihtelevia. Yhdessä laajassa metaregressioanalyysissä havaittiin, että tytöt saavat keskimäärin hieman parempia arvosanoja kuin pojat. (D. Voyer ja S. D. Voyer 2014)

Toisen meta-analyysin perusteella pojat näyttivät suoriutuvan keskimäärin hieman paremmin, erityisesti tehtävissä, jotka perustuivat valintavaihtoehtoihin ja vaativat päättelyä, kun taas tytöt menestyivät paremmin avoimia tai kirjallisia vastauksia edellyttävissä tehtävissä. (Lindberg et al. 2010) Yhdysvaltalaisen koetulosten analyysi ei puolestaan osoittanut tilastollisesti merkitseviä sukupuolieroja keskiarvoissa. (Hyde et al. 2008)

PISA-aineistojen vertailussa pojat suoriutuivat keskimäärin hieman paremmin, mutta erot olivat pienempiä tai jopa kääntyivät tyttöjen eduksi maissa, joissa sukupuolten välinen tasa-arvo on korkeammalla tasolla. (Else-Quest, Hyde ja Linn 2010) PISA 2003-aineiston pohjalta todettiin, että huipputason (99. prosenttiin ylittävät) matematiikkasuoriutumisessa sukupuolten välinen ero on pienempi tasa-arvoisemmissa maissa. (Guiso et al. 2008)

¹Perusopetustasolla Kansallisen koulutuksen arviointikeskus

Toisaalta, Yhdysvaltojen kansallisissa aineistoissa havaittiin, että poikien tuloksissa oli enemmän vaihtelua, mikä johti heidän yliedustukseensa sekä huipputuloksissa että heikommassa päässä. (Hedges ja Nowell 1995) Sama ilmiö havaittiin kansainvälisessä vertailussa, jossa poikien suurempi hajonta toistui useimmissa maissa. (Machin ja Pekkarinen 2008)

Suomessa, PISA-aineistossa vuosilta 2005–2022 perustuen, ei 9. luokan matematiikan osaamisen tasossa havaittu merkittäviä sukupuolieroja, mutta tytöt saivat koulussa parempia arvosanoja ja pojat olivat yliedustettuina huippusuoriutujien joukossa. (Metsämuuronen 2013) Sama ilmiö toistui uusimmissa PISA-tuloksissa. Tytöt suoriutuivat keskimäärin poikia hieman paremmin, mutta pojat saavuttivat korkeampia pistemääriä huippusuoriutujien joukossa. (OECD 2023, s. 124) Aalto-yliopistossa havaittiin, että naiset suoriutuivat hieman paremmin matematiikan peruskursseilla verrattuna miehiin Aalto-yliopistossa. (Miihkinen 2013, s. 60)

Kirjallisuuden perusteella ei voida tehdä yksiselitteisiä johtopäätöksiä, koska tulokset vaihtelevat. Kuitenkin maissa, joissa sukupuolten tasa-arvo on korkeampi, matematiikan osaamistasot ovat tyypillisesti lähempänä toisiaan. Suomessa tytöt saavat keskimäärin hieman parempia arvosanoja, mutta pojat ovat yliedustettuina huippusuoriutujien joukossa.

2.3 Kansalaisuus ja matematiikan arvosanat

Valtiontalouden tarkastusviraston mukaan ensimmäisen polven maahanmuuttajataustaiset oppilaat suoriutuivat matematiikan PISA-kokeessa selvästi heikommin kuin kantasuomalaiset oppilaat. (Tarkastusvirasto 2015, s. 36)

Myöhempi tutkimus tarkensi havaintoa analysoimalla Suomen PISA-aineistoja vuosilta 2003–2018. Tulosten mukaan kaikki maahanmuuttajasukupolvet jäivät keskimäärin kantaväestöstä, mutta erot vaihtelivat selvästi sukupolvien välillä. Heikoimmin menestyivät 1.25-sukupolven oppilaat (yli 13-vuotiaina Suomeen muuttaneet), kun taas 2.5-sukupolven oppilaiden (joista toinen vanhemmista on syntynyt Suomessa) erot kantaväestöön olivat pienimmät. 1.75-sukupolven oppilaat (muuttaneet Suomeen alle 7-vuotiaina) sijoituivat tasaisesti keskivaiheille. (Planting 2022, s. 3–38)

Uusimmat PISA 2022-tulokset osoittavat, että erityisesti ensimmäisen polven maahanmuuttajaoppilaiden matematiikan osaamisessa on yhä merkittävä ero kantaväestöön verrattuna, vaikka erot ovat jonkin verran kaventuneet aiemmista tutkimuskerroista. (OECD 2023, s. 215–217)

2.4 Arvosanat Aalto-yliopiston korkeakouluissa

Aalto-yliopiston sisäisiä arvosanaeroja korkeakoulujen välillä on tutkittu toistaiseksi melko vähän. Ainoassa aiemmassa tutkimuksessa on havaittu merkittäviä eroja Aalto-yliopiston korkeakoulujen välillä myöhemmän opintomenestyksen ennustamisessa matematiikan kurssiarvosanojen perusteella (Miihkinen 2013, s. 74–76).

2.5 Arvosanainflaation kehitys koulutusasteilla

Arvosanainflaatiolla tarkoitetaan ilmiötä, jossa arvosanojen keskiarvot nousevat ilman vastaavaa osaamisen paranemista. Ilmiö on havaittavissa sekä perusopetuksessa että korkeakoulutuksessa sekä kansainvälisesti että Suomessa.

Yhdysvalloissa on raportoitu, että korkeakoulujen myöntämien arvosanojen keskiarvo (GPA) on noussut kaikilla tieteenaloilla keskimäärin noin 0,1² GPA-pistettä per vuosikymmen 1960-luvulta 2010-luvulle. Samalla parhaimpien A-arvosanojen osuus on kolminkertaistunut: vielä 1960-luvulla arvosana A oli harvinainen, mutta 2000-luvun loppuun mennessä siitä oli tullut yleisin arvosana, kattaen keskimäärin 43 % kaikista arvosanoista (Rojstaczer ja Healy 2012). Vastaavasti Yhdysvalloissa, vuosien 1990 ja 2020 välillä, matematiikan ja insinööritieteiden opiskelijoiden keskimääräiset arvosanat nousivat merkittävästi (Education Statistics 1990; Education Statistics 2020).

Suomessa on havaittu samanlainen kehitys: vuosina 2005–2012 yhdeksännen luokan matematiikan arvosanojen keskiarvo nousi 0,4 pistettä (7,4 → 7,8), ilman osoitettavaa osaamisen paranemista (Metsämuuronen 2013). Uusimmat kriteeriuudistukset pyrkivät hillitsemään inflaatiota standardoimalla arviointia, mutta niiden vaikutukset ovat vielä epäselviä (Karvi 2023).

Pidemmällä aikavälillä 9. luokkalaisten arvosanan 10 osuus on kolminkertaistunut ja arvosanan 5 osuus laskenut kolmasosaan vuosien 2000 ja 2021 välillä, vaikka oppilaiden osaamistaso on samanaikaisesti heikentynyt (Metsämuuronen, Nousiainen ja Peltoniemi 2023). Samaan aikaan 9. luokkalaisten PISA-tulokset matematiikassa ovat olleet jyrkässä laskussa Suomessa (OECD 2023, s. 415).

Taulukko 1: Arvosanainflaation keskeisiä havaintoja.

koulutusjärjestelmä	aikaväli	havaittu muutos	lähde
yliopistot (USA)	1960–2010	GPA noussut 0,1 pistettä / vuosikymmen; A-arvosanat kolminkertaistuneet (→ 43 %)	(Rojstaczer ja Healy 2012)
yliopistot (USA)	1990–2020	Matematiikan ja insinööritieteiden keskiarvot nousussa	(Education Statistics 1990; Education Statistics 2020)
perusopetus (Suomi)	2005–2012	Matematiikan keskiarvo nousut 7,4 → 7,8 ilman osaamistason paranemista	(Metsämuuronen 2013)
perusopetus (Suomi)	2000–2021	Arvosanan 10 osuus kolminkertaistunut, arvosanan 5 osuus laskenut kolmasosaan	(Metsämuuronen, Nousiainen ja Peltoniemi 2023)
perusopetus (Suomi)	2000–2021	PISA-tulokset heikentyneet Suomessa	(OECD 2023; Vihriälä 2023)

²Alkuperäinen asteikko: 1–4. Tässä tutkimuksessa on käytössä asteikko 1–5.

3 Tutkimusaineisto

3.1 Aineiston kuvaus

Tutkimuksen aineisto koostuu Aalto-yliopiston ja Aalto-yliopiston avoimen yliopiston opiskelijoiden matematiikan kurssisuorituksista vuosilta 2014–2024. Mukaan on otettu suoritukset, jotka on rekisteröity viimeistään 11.2.2025. Tarkasteltavat kurssit on jaettu kahteen ryhmään:

Peruskurssit

- Differentiaali- ja integraalilaskenta 1 (MS-A010X)
- Differentiaali- ja integraalilaskenta 2 (MS-A020X)
- Matriisilaskenta (MS-A000X)
- Todennäköisyyslaskennan ja tilastotieteen peruskurssi (MS-A050X)

Jatkokurssit

- Statistical Inference (MS-C1620)
- Fourier-analyysi (MS-C1420)
- Lineaarialgebra (MS-C1342)

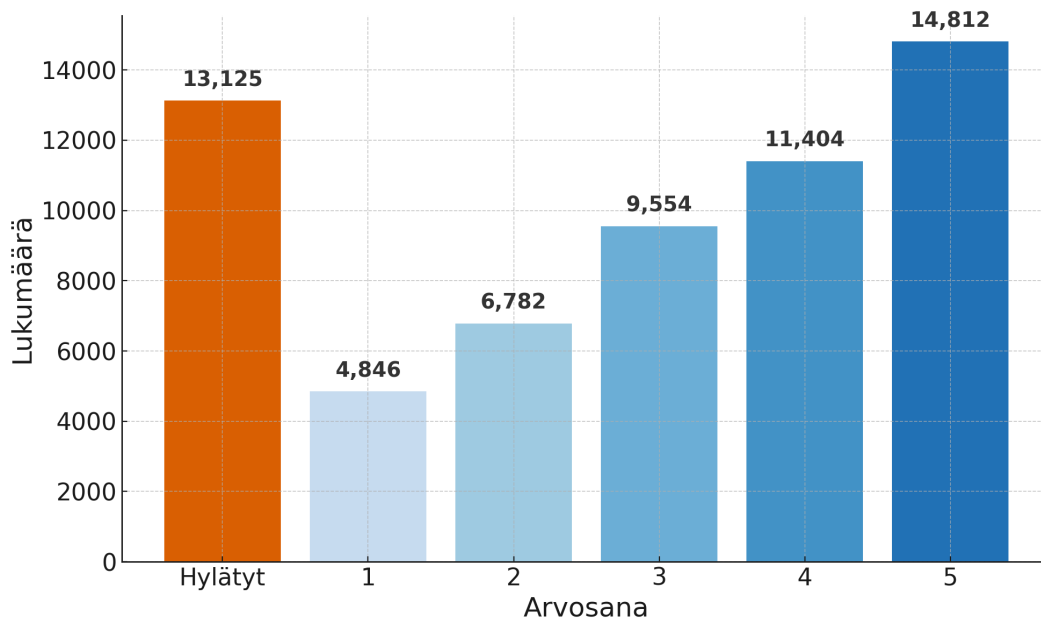
Aineistossa opiskelijatunnisteet on pseudonymisoitu EU:n yleisten tietosuojasetusten (GDPR) mukaisesti siten, että yksittäisten opiskelijoiden kurssisuorituksia voidaan yhdistää, mutta ei tunnistaa. Jokaisesta opiskelijasta on mukana tiedot sukupuolesta, kansalaisuudesta, syntymäkuukaudesta, korkeakoulusta, koulutustyyppistä (esimerkiksi perustutkinto- tai jatko-opiskelija), suoritusvuodesta ja arvosanasta. Aineistoa varten haettiin tutkimuslupa, ja sitä on käsitelty luottamuksellisesti koko analyysin ajan.

Kokonaisaineisto sisältää 62 423 kurssisuoritusta, joista 49 298 on hyväksyttyjä ja 13 125 hylättyjä. Kaikkiin analyysihin on otettu mukaan vain opiskelijoiden ensimmäiset suoritukset, myöhempiä arvosanan korotuksia ei ole huomioitu. Näin pyritään säilyttämään suoritusten välinen riippumattomuusoletus.

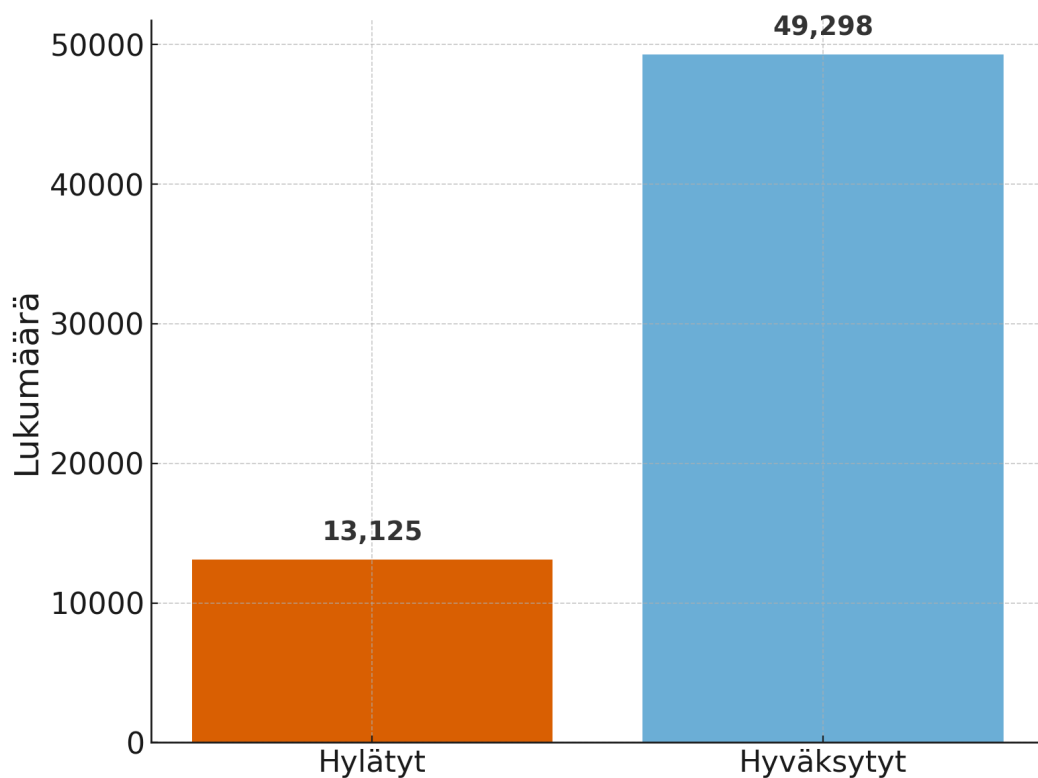
Suoritusten jakautuma on esitetty kuvassa [1b](#), Arvosanojen jakauma on esitetty kuvassa [1a](#) sekä hyväksytyjen että hylättyjen suoritusten jakauma kuvassa [1b](#).

Kurssikohtaiset suoritusmäärät ja hylkäysprosentit vaihtelevat merkittävästi. Kuvasta [2a](#) käy ilmi, että erityisesti peruskursseilla on huomattavasti enemmän suorituksia verrattuna jatkokursseihin. Tämän vuoksi tulevassa analyysissä keskitytään ensin peruskursseihin, joita lähes kaikki opiskelijat ovat suorittaneet. Lisäksi tarkastellaan koko aineistoa, jossa mukana ovat myös jatkokurssit. Kuten kuvasta [2b](#) nähdään, kurssien välillä on selviä eroja myös hylkäysprosentteissa.

Kurssisuoritusten määrä on kasvanut tasaisesti vuosittain vuodesta 2014 alkaen. Tämä kehitys on nähtävissä kuvassa [3](#).

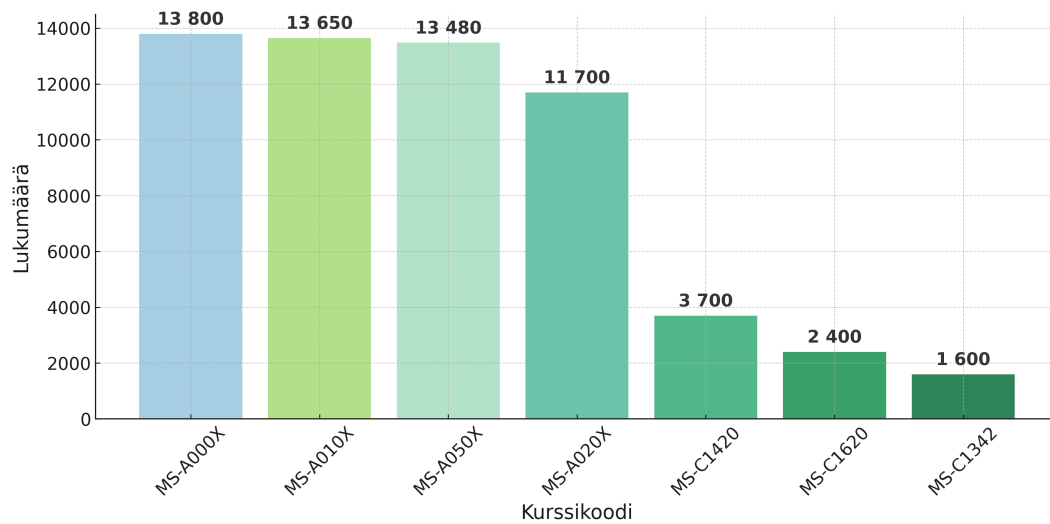


(a) Arvosanojen jakauma (1–5 sekä hylätyt).

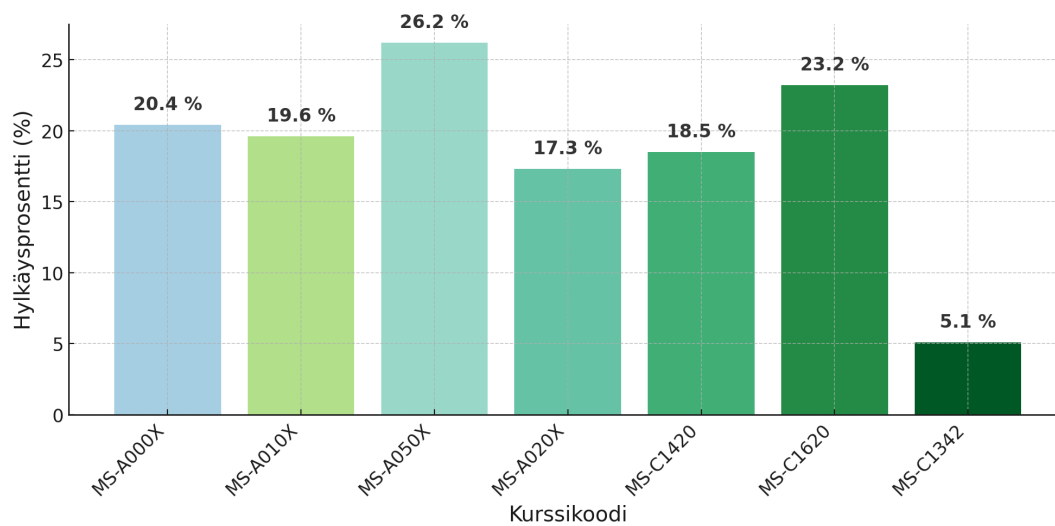


(b) Hyväksytyjen ja hylättyjen kurssisuoritusten jakauma.

Kuva 1: Suoritusten arvosanajakauma ja hyväksytyjen/hylättyjen osuudet. Pylväiden yläpuolella näkyvät suoritusten määrät.

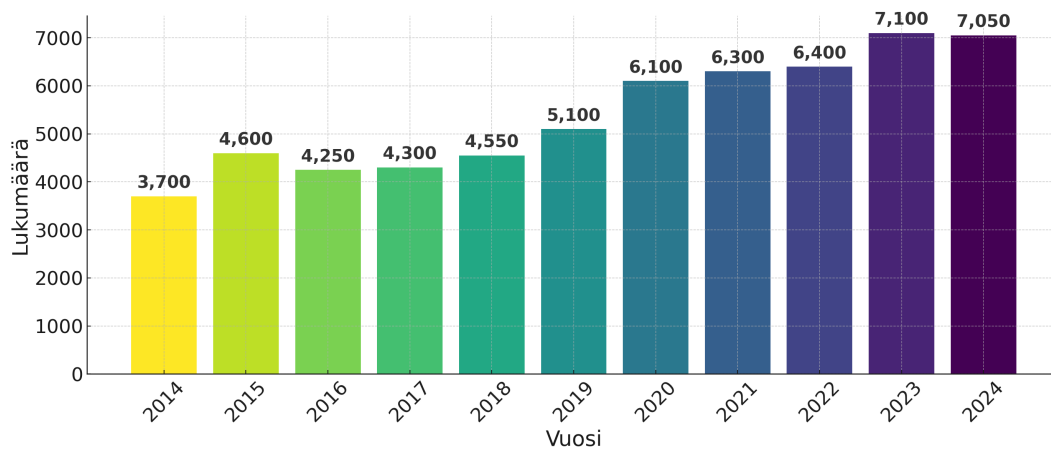


(a) Kurssisuoritusten lukumäärät kurseittain.



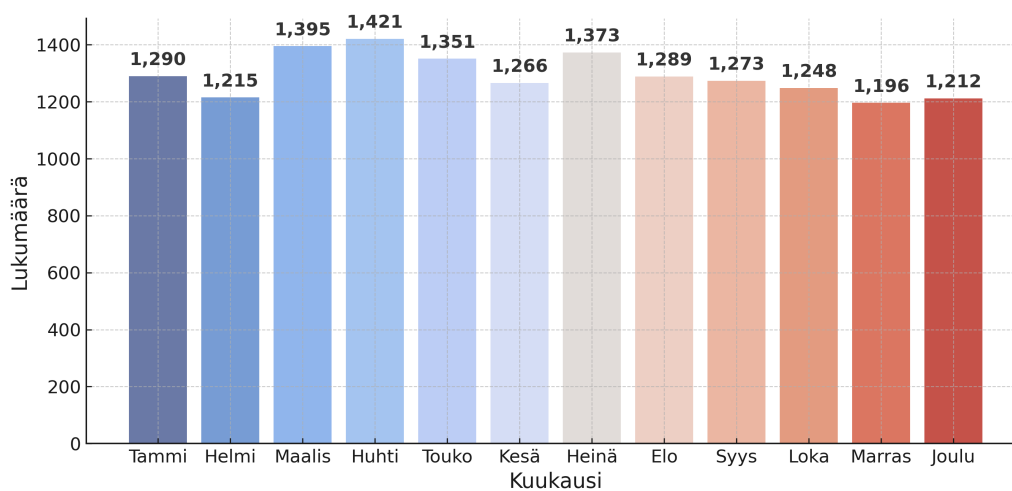
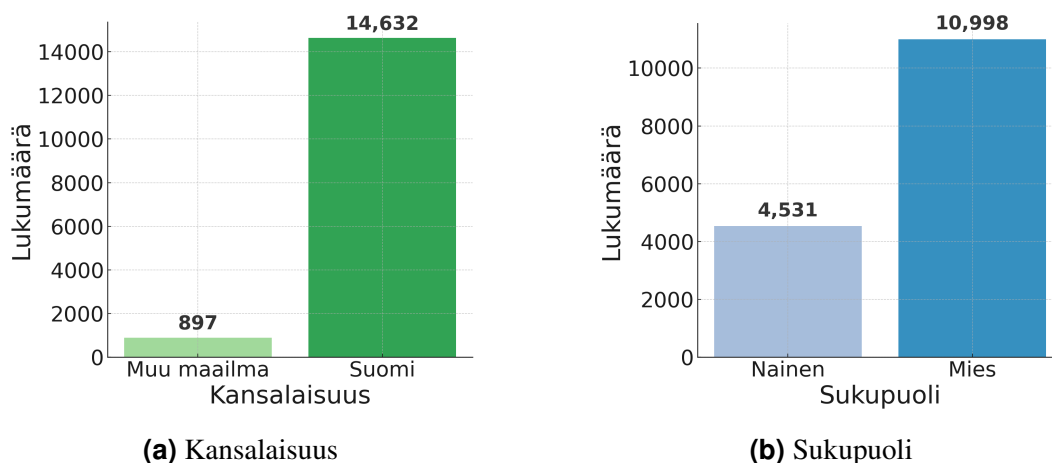
(b) Hylättyjen osuudet kurseittain.

Kuva 2: Kurssikohtaiset suoritusmäärät ja hylkäysprosentit.



Kuva 3: Kurssisuoritukset vuosittain (2014–2024).

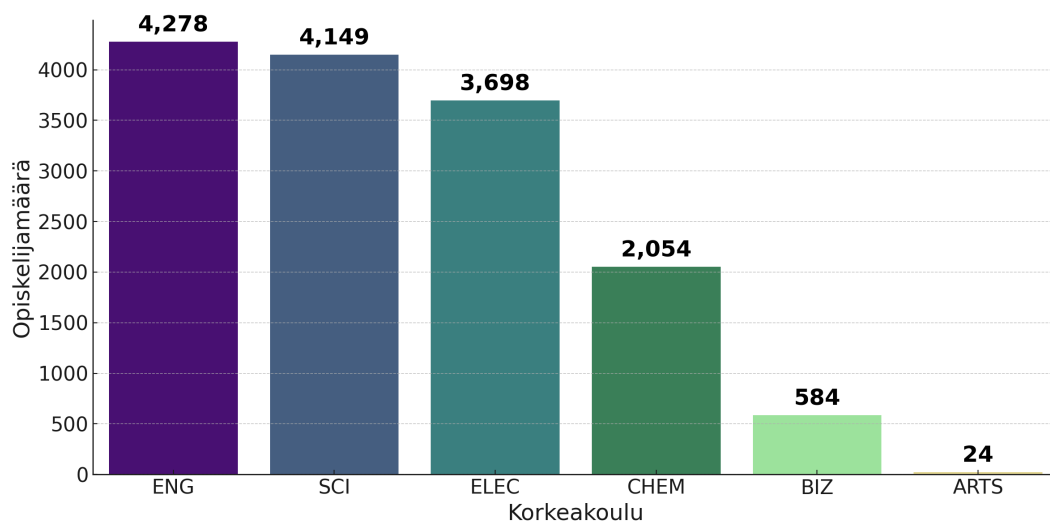
Opiskelijaprofiilit kuvaavat tarkemmin, keitä analysoidut opiskelijat ovat. Kuvassa 4 on esitetty opiskelijoiden kansalaisuus, sukupuolijakauma ja syntymäkuukaudet. Suurin osa opiskelijoista on suomalaisia miehiä, ja syntymäkuukausien jakauma on suhteellisen tasainen. On tärkeää huomioida, että opiskelijoiden määrä poikkeaa kurssisuoritusten kokonaismäärästä, sillä suurin osa opiskelijoista on suorittanut useita matematiikan kursseja. Aineistossa on yhteensä 62 423 kurssisuoritusta, mutta nämä on tehty 15 530 eri opiskelijan toimesta.



Kuva 4: Opiskelijoiden perustiedot: kansalaisuus, sukupuoli ja syntymäkuukausi. Mukana opiskelijat, jotka ovat suorittaneet vähintään yhden tarkastelluista kurseista.

Kuvassa 5 tarkastellaan opiskelijoiden jakaumaa eri Aalto-yliopiston korkeakouluissa. Eniten opiskelijoita on insinöritieteiden korkeakoulusta (ENG) ja perustieteiden korkeakoulusta (SCI).

Aineisto on jaettu kolmeen osakokonaisuuteen: hyväksytyt suoritukset (ensimmäinen hyväksytty suoritus), hylätyt suoritukset sekä kurssien arvosanojen korotukset.



Kuva 5: Korkeakoulujakauma: matematiikan kursseja suorittaneet opiskelijat

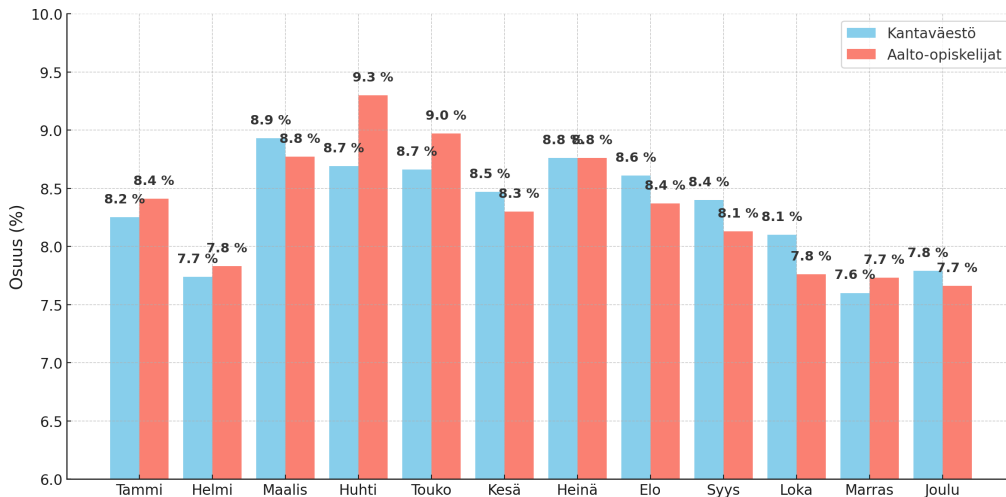
Varsinaisessa analyysissä keskitytään ensisijaisiin suorituksiin, hyväksytyihin ja hylättyihin, jolloin korotetut suoritukset on jätetty tarkastelun ulkopuolelle.

3.2 Syntymäkuukausien jakauma ja vertailu väestöön

Ensin tarkastellaan poikkeako Aalto-yliopiston suomalaiset matematiikan kursien suorittaneiden opiskelijoiden syntymäkuukausijakauma Suomen kantaväestön jakaumasta³ Vertailuun käytettiin Tilastokeskuksen syntymäkuukausidataa vuosina 1984–2007 syntyneistä, koska aineiston opintosuoritukset ajoittuvat vuosille 2014–2024 ja kohdejoukon ikähaarukka on siten arviolta 17–30 vuotta (Tilastokeskus 2025).

Kuvassa 6 on esitetty vertailu kantaväestön ja opiskelijoiden syntymäkuukausien välillä. Kaksi pylvästä jokaiselta kuukaudelta kuvaavat kantaväestön (sininen sävy) ja Aalto-yliopiston opiskelijoiden (punainen sävy) osuutta kyseisen kuukauden syntyneistä.

Kuvasta nähdään, että syntymäkuukausijakaumat poikkeavat toisistaan. Osuustesti vahvasti havainnon. Tammi–maaliskuussa (ensimmäinen kvartaali) syntyneiden osuus ei eronnut tilastollisesti merkitsevästi suomalaisten Aalto-yliopiston opiskelijoiden (25.0 %) ja kantaväestön (24.9 %) välillä ($\chi^2 = 0.28$, $df = 1$, $p = 0.596$). Tammi–kesäkuussa (ensimmäinen puolivuosi) syntyneiden osuus oli opiskelijoilla 51.6 % ja kantaväestöllä 50.7 %, ja ero oli tilastollisesti merkitsevä ($\chi^2 = 16.14$, $df = 1$, $p < 0.001$).



Kuva 6: Suomalaisen Aalto-opiskelijoiden ja Suomen väestön syntymäkuukausijakaumat (1984–2007 syntyneet).

³Huomiona, että ryhmät eivät ole keskenään täysin riippumattomia. Kaikki suomalaiset Aalto-opiskelijat sisältyvät Suomen väestökisteriin, josta kantaväestön data on kerätty. Opiskelijoiden osa kantaväestöstä on kuitenkin merkityksetön.

4 Menetelmät

4.1 Logistinen regressio

Logistista regressiota käytetään, kun vastemuuttuja saa vain kaksi arvoa. Tässä tutkimuksessa ne ovat *hyväksytty* (1) ja *hylätty* (0). Mallin tarkoituksena on estimoida todennäköisyys hyväksynnälle annetun selittäjäjoukon \mathbf{x} , eli opiskelijan taustatietojen perusteella, esim. sukupuoli, kansalaisuus, syntymäkuukausi jne. Mallin ytimessä on logit-muunnos:

$$\text{logit}(\pi(\mathbf{x})) = \log\left(\frac{\pi(\mathbf{x})}{1 - \pi(\mathbf{x})}\right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k = \beta_0 + \mathbf{x}^\top \boldsymbol{\beta}, \quad (1)$$

jossa opiskelijan tiedoilla \mathbf{x} laskettu hyväksymistodennäköisyys on $\pi(\mathbf{x}) = P(Y = 1 \mid \mathbf{x})$. (Agresti 2007, s. 99–100) Yhtälössä (1) β_1, \dots, β_k ovat regressiokertoimia, jotka kuvaavat kunkin taustamuuttujan vaikutusta logit-muunnokseen. Vastaavasti x_1, \dots, x_k ovat selittäviä muuttujia eli taustamuuttujia. Tässä tutkimuksessa suurin osa x_j :stä on kaksiarvoisia indikaattorimuuttujia (esim. sukupuoli, syntynyt alkuvuonna, kansalaisuus)- *Arvosanatrendi* on jatkuva muuttuja. $\mathbf{x}^\top \boldsymbol{\beta}$ on sisätulo vektoreiden \mathbf{x} ja $\boldsymbol{\beta}$ välillä, eli summaa $\sum_{j=1}^k \beta_j x_j$.

Vakiotermi β_0 vastaa logit-muunnoksessa tilannetta, jossa kaikki selittävät muuttajat saavat arvon nolla. Vastaavasti regressiokertoimet β_j kuvaavat kunkin selittävän muuttujan vaikutusta logit-muunnokseen muiden muuttujien vaikutukset vakioituna.

Logit-muunnos hyödyntää tapahtuman todennäköisyyden $\pi(\mathbf{x})$ ja sen vastatapahtuman todennäköisyyden vastasuhdetta (odds). (Agresti 2007, s. 28–29) Vastasuhte tarkoittaa todennäköisyyden π ja sen vastatapahtuman todennäköisyyden $1 - \pi$ välistä suhdetta:

$$\text{Vastasuhte} = \frac{\pi(\mathbf{x})}{1 - \pi(\mathbf{x})} \quad (2)$$

Vastasuhte kertoo, kuinka monta kertaa todennäköisyys $\pi(\mathbf{x})$ on suurempi kuin sen vastatodennäköisyys $1 - \pi(\mathbf{x})$. Logit-muunnos muuntaa vastasuhteen logaritmiksi, jolloin vastasuhdetta on helpompi mallinta.

Todennäköisyys voidaan ratkaista kaavasta (1):

$$\pi(\mathbf{x}) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)}} \quad (3)$$

$$= \frac{e^{\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k}}{1 + e^{\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k}} \quad (4)$$

Ristisuhte $\exp(\beta_j)$ ilmaisee, kuinka moninkertaiseksi vastasuhte muuttuu, kun kyseisessä selittäjässä tapahtuu yhden yksikön muutos, muiden muuttujien pysyessä vakioina. (Agresti 2007, s. 104–105) Luokkamuuttujien kohdalla tämä tarkoittaa vertailua vertailuluokkaan (eli tilanteeseen, jossa luokkamuuttujan arvo on 0). Jatkuvien

muuttujien, aikatrendin, kohdalla $\exp(\beta_j)$ kertoo vaikutuksen vastasuhteeseen, kun muuttujan arvo kasvaa yhdellä yksiköllä.

Laajentamalla tarkastelu koko selittäjäjoukkoon \mathbf{x} , saadaan vastasuhde:

$$\frac{\pi(\mathbf{x})}{1 - \pi(\mathbf{x})} = \exp(\beta_0 + \mathbf{x}^\top \boldsymbol{\beta})$$

Tätä voi kehittää kaavan (5) mukaisesti:

$$\text{Uusi vastasuhde} = \exp(\beta_0) \cdot \exp(\mathbf{x}^\top \boldsymbol{\beta}) = \frac{\pi(\mathbf{x}_0)}{1 - \pi(\mathbf{x}_0)} \cdot \exp(\mathbf{x}^\top \boldsymbol{\beta}) \quad (5)$$

Tässä \mathbf{x}_0 viittaa vertailuprofiiliin. Ristisuhteen $\exp(\mathbf{x}^\top \boldsymbol{\beta})$ tulkinta on:

- $\exp(\mathbf{x}^\top \boldsymbol{\beta}) > 1$: opiskelijalla on suurempi todennäköisyys hyväksyntään kuin vertailuprofiililla,
- $\exp(\mathbf{x}^\top \boldsymbol{\beta}) < 1$: opiskelijalla on pienempi todennäköisyys hyväksyntään kuin vertailuprofiililla,
- $\exp(\mathbf{x}^\top \boldsymbol{\beta}) = 1$: selittäjäjoukon \mathbf{x} ei vaikutusta hyväksyntään.

4.1.1 Mallin arviointi: Erottelukykykäyrä ja käyräalainen pinta-ala

Logistinen regressiomalli mallintaa muuttujien vaikutusta vasteen vastasuhteeseen, jonka avulla hyväksymistodennäköisyys $\pi(\mathbf{x})$ lasketaan. Mallin luokittelu tehdään vertaamalla tätä todennäköisyyttä ennalta asetettuun kynnyksarvoon: jos todennäköisyys ylittää kynnyksen, havainto luokitellaan hyväksytyksi; muussa tapauksessa hylätyksi.

Esimerkiksi kohdan 5.1.2 mallissa⁴, jossa $x_1 = 1$ (nainen) ja $x_2 = 1$ (syntynyt joulukuussa).

$$\text{logit}(\pi(x)) = 1.5 + 0.37x_1 + 0.10x_2,$$

Asettamalla $x_1 = 1$ ja $x_2 = 1$, logit-arvoksi saadaan $1.5 + 0.37 + 0.1 = 1.97$. Tämä vastaa kaavan (4) hyväksyntätodennäköisyyttä

$$\pi(x) = \frac{1}{1 + e^{-1.97}} \approx 0.88$$

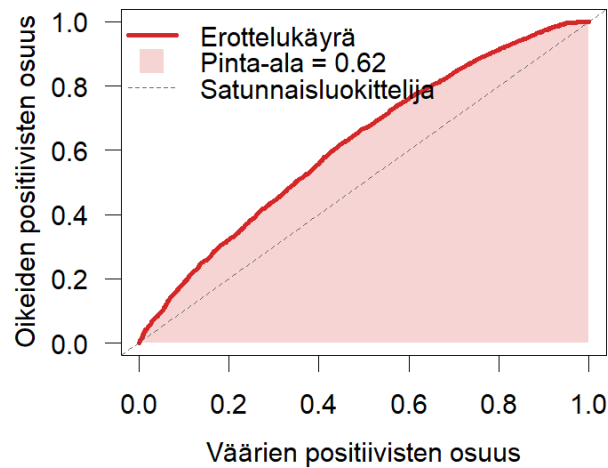
Jos kynnyksarvoksi asetetaan 0.5, havainto luokitellaan hyväksytyksi. Mikäli raja olisi korkeampi, esimerkiksi 0.89, sama havainto tulkittaisiin hylätyksi. Kynnyksarvon muuttaminen vaikuttaa siihen, kuinka monta havaintoa luokitellaan oikein tai väärin, mikä heijastuu mallin oikeiden luokittelujen osuuksissa (true positive rate) ja väärin luokittelujen osuuksissa (false positive rate).

Mallin erotuskykyä arvioidaan erottelukykykäyrällä (roc-curve), jossa vaakakselilla esitetään väärin positiivisten osuus ja pystyakselilla oikeiden positiivisten osuus. Käyrä muodostetaan vaihtelemalla kynnyksarvoa välillä 0–1 ja laskemalla näihin liittyvät arvot kunkin kynnyksen kohdalla (Hosmer, Lemeshow ja Sturdivant 2013, s. 174–177).

⁴Muut kertoimet asetettu nollassi.

Käyräalain pinta-ala (area under the curve) mittaa mallin keskimääräistä kykyä erottaa toisistaan oikeat ja väärät luokitellut tapaukset. Mitä lähempänä käyräalain pinta-ala on yhtä, sitä parempi on mallin erotuskyky. Kuvassa 7 on esimerkki ROC-käyrästä ja tämän pinta-alasta. Arvo voidaan tulkita seuraavasti (Hosmer, Lemeshow ja Sturdivant 2013, s. 177):

- $AUC = 0.5$: Ei parempi kuin satunnainen arvonta
- $AUC > 0.7$: Kohtalainen erotuskyky
- $AUC > 0.8$: Hyvä erotuskyky
- $AUC > 0.9$: Erinomainen erotuskyky



Kuva 7: Esimerkkimalli kohdasta 5.1.2, käyräalain pinta-ala = 0.62.

4.2 Järjestyslogistinen regressiomalli

Kun vastemuuttuja saa useampia järjestysasteikollisia arvoja, logistinen regressiomallia voidaan yleistää järjestyslogistiseksi regressiomalliksi. Tällöin ei tarkastella pelkästään hyväksyty/hylätty asetelmaa, vaan hyväksytyjen suoritusten sisäinen erotteleminen tulee mahdolliseksi. Tässä tutkielmassa vastemuuttujassa on kolme luokkaa paremmuusjärjestyksessä.

Järjestysasteikollisen vastemuuttujan analysointiin soveltuu verrannollisten vastasuhteiden malli (proportional odds model). Mallissa estimoidaan ristisuhteita eli vastasuhteita sille, että havainto sijoittuu tiettyyn luokkaan tai alempiin suhteessa ylempiin luokkiin. Näitä kuvataan kertymätodennäköisyyksien avulla seuraavasti:

$$\begin{aligned}\text{logit}(P(Y \leq j | \mathbf{x})) &= \log \left(\frac{P(Y \leq j | \mathbf{x})}{P(Y > j | \mathbf{x})} \right) \\ &= \alpha_j + \mathbf{x}'\boldsymbol{\beta}, \quad j = 1, \dots, J - 1\end{aligned}\quad (6)$$

Mallin vakiotermit α_j ovat kynnsarvoja, jotka määrittävät rajat kertymätodennäköisyyksille $P(Y \leq j | \mathbf{x})$. Kullekin luokkarajalle $j = 1, \dots, J - 1$, estimoidaan oma kynnsarvo. Jokaista järjestysluokkaa j mallinnetaan samalla regressiokertoimien vektorilla $\boldsymbol{\beta}$. Viimeiselle luokalle J ei tarvita erillistä kynnsarvoa, koska kertymätodennäköisyys sille on aina 1. Mallissa on $J - 1$ kynnsarvoa. Ristisuhdekerroin $\exp(\beta_j)$ kuvaa yksittäisen selittävän muuttujan vaikutusta vastasuhteeseen ja sitä kautta todennäköisyyteen sijoittua luokkaan j tai sitä alempiin. (Agresti 2007, s. 180)

Tutkimuksessa käytetään seuraavaa kolmiportaista luokittelua vasteelle:

- *Luokka 0*: Hylätty
- *Luokka 1*: tyydyttävä (arvosanat 1–3)
- *Luokka 2*: hyvä (arvosanat 4–5)

4.2.1 Poikkeamat verrannollisten vastasuhteiden oletuksesta

Verrannollisten vastasuhteiden malli perustuu oletukseen, että regressiokertoimien β_j yhtälöstä (6) vaikutus on sama kaikilla kynnsarvoilla. (Agresti 2007, s. 180–182) Kun analysoidaan isoja aineistoja tämä oletus ei välttämättä päde: jotkin selittävät muuttujat voivat vaikuttaa eri tavoin eri kynnsarvoilla.

Tällaisissa tapauksissa voidaan soveltaa *osittain verrannollisten vastasuhteiden mallia* (partial proportional odds model), jossa osa muuttujista saa eri kertoimet eri kynnsarvoille, kun taas muilla muuttujilla voidaan olettaa edelleen vakiovaikutus.

4.2.2 Brantin testi: Verrannollisten vastasuhteiden oletukselle

Brantin testin idea perustuu siihen, että sovitetaan $k - 1$ erillistä logistista regressiomallia vastemuuttujan jokaiselle luokitteluportaille ja verrataan näiden mallien

regressiokertoimia toisiinsa. Jos nämä eroavat toisistaan tilastollisesti merkitsevästi, oletus verrannollisista vastasuhteista kaavassa (6) rikkoutuu.

Brantin testisuure vertaa estimoituja regressiokertoimia $\tilde{\mathbf{b}}$ ja niiden kovarianssimatriisia \widehat{V} kontrastimatriisiin D avulla:

$$X^2 = (D\widehat{\mathbf{b}})^\top \left[D\widehat{V}(\tilde{\mathbf{b}})D^\top \right]^{-1} (D\tilde{\mathbf{b}}) \stackrel{a}{\sim} \chi_{(k-2)p}^2,$$

jossa p on selittävien muuttujien määrä. Kontrastimatriisilla D vertaillaan eri kynnsfunktioiden regressiokertoimia toisiinsa.

$$\mathbf{D} = \begin{bmatrix} I & -I & 0 & \cdots & 0 \\ I & 0 & -I & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ I & 0 & 0 & \cdots & -I \end{bmatrix} \quad \tilde{\mathbf{b}} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{k-1} \end{bmatrix} \in \mathbb{R}^{(k-1)p \times 1} \quad \beta_j = \begin{bmatrix} \beta_{j1} \\ \beta_{j2} \\ \vdots \\ \beta_{jp} \end{bmatrix}$$

Testi toteutetaan tutkielmassa sekä koko mallille että muuttujakohtaisesti (Brant 1990, s. 1173).

Nollahypoteesi on, että verrannollisten vastasuhteiden oletus pätee, eli selittävien muuttujien vaikutus on sama kaikilla vasteen kynnyksillä. Kun testin p -arvo on pieni, nollahypoteesi hylätään, ja päätellään, että verrannollisten vastasuhteiden oletus ei päde ja vaikutus ei ole sama kaikilla kynnsarvoilla.

4.2.3 Mallin selityskyky: Yhteensopivuusindeksi

Yhteensopivuusindeksi (concordance index) mittaa osuutta kaikista havaintopareista, joissa malli järjestää havainnot oikein vasteen perusteella. Indeksii voidaan tulkita myös todennäköisyytenä, että satunnaisesti valittu pari havaintoja, joilla on eri vastearvot, on mallin sovituksen mukaan järjestyksessä oikein.

$$C = \Pr\{\hat{y}_i > \hat{y}_k \mid y_i > y_k\} = \frac{M}{N},$$

M on parien määrä, joissa mallin sovitteet ovat samassa järjestyksessä kuin arvosanat, ja N on kaikkien eri arvosanan saaneiden parien määrä (Harrell 2015, s. 257).

Yhteensopivuusindeksi yleistää erottelukäyränalaista pinta-alaa (ks. kohta 4.1.1) tilanteisiin, joissa vastemuuttujalla on järjestysasteikollinen. Jos vaste on kaksiarvoinen, indeksi saa saman arvon kuin käyränalainen pinta-ala. Indeksien arvot vaihtelevat välillä:

- 0.5 merkitsee satunnaista järjestystä.
- 1 viittaa täydelliseen järjestykseen mallin sovitteissa.

Harrell pitää arvoa $C \geq 0,8$ hyvästä erottelukyvystä. (Harrell 2015, s. 257) Huomiona, että yhteensopivuusindeksi mittaa mallin kykyä järjestää havaintoja vasteen mukaan, mutta ei arvioi sitä, kuinka hyvin mallin sovittamat todennäköisyydet vastaavat havaittuja jakaumia. (ks. esim. Agresti 2010, s. 65).

4.3 Yhteinen diagnostiikka

Seuraavat testit ja tunnusluvut hyödynnetään sekä kaksiarvoiselle logistiselle regressiolle että järjestyslogistiselle regressiolle.

4.3.1 Waldin testi

Waldin testi arvioi yksittäisten selittävien muuttujien tilastollista merkitsevyyttä. Testisuure perustuu kertoimen ja sen keskihajonnan suhteeseen:

$$z = \frac{\hat{\beta}_j}{\sigma(\hat{\beta}_j)}, \quad z^2 \stackrel{a}{\sim} \chi_1^2 \quad (7)$$

Kaavan (7) avulla lasketaan z -arvo, josta voidaan johtaa p -arvo χ^2 -jakaumasta. (Agregti 2010, s. 60–61) Nollahypoteesi on, että tarkasteltavan selittävän muuttujan regressiokerroin on nolla eli muuttujalla ei ole vaikutusta vasteeseen. Kun testin on pieni, nollahypoteesi hylätään ja muuttuja tulkitaan tilastollisesti merkitseväksi.

4.3.2 Uskottavuusosamäärätesti

Uskottavuusosamäärätesti (UOM-testi, likelihood ratio test) vertailee kahden mallin istuvuutta: täyttä mallia ilman rajoituksia ja sen rajoitettua versiota, jossa osalle parametreille on asetettu rajoituksia. Testisuure lasketaan näiden mallien logaritmisten uskottavuuksien erotuksen avulla:

$$G^2 = 2 [\log L_{\text{täysi malli}} - \log L_{\text{rajoitettu malli}}] \stackrel{a}{\sim} \chi_g^2$$

jossa g on rajoitettujen parametrien määrä. Testin nollahypoteesi on, että rajoitetut parametrit eivät paranna mallin selitysvoimaa, joten täysi ja rajoitettu malli ovat tilastollisesti yhtä hyviä. Nollahypoteesi kumotaan, mikäli testisuureen p -arvo on riittävän pieni. Tällöin lisättyjen muuttujien katsotaan parantavan mallin istuvuutta merkitsevästi. (Harrell 2015, s. 205) Uskottavuusosamäärätesti on monikäyttöinen ja sitä hyödynnetään esimerkiksi devianssin määrittelyssä kohdassa 4.3.3 sekä Akaiken informaatiokriteerin laskennassa kohdassa 4.3.5.

4.3.3 Devianssi ja Pearsonin χ^2 -testi

Sekä devianssi että Pearsonin χ^2 -testi arvioivat sovitetun mallin sopivuutta vertaamalla sitä täydellisesti sopivaan eli saturoituneeseen malliin, mutta eri tavoin.

Devianssin testisuure perustuu sovitetun mallin ja saturoituneen mallin logaritmisten uskottavuuksien erotukseen:

$$D = -2 [\log L_{\text{malli}} - \log L_{\text{sat}}] \stackrel{a}{\sim} \chi^2,$$

jossa L_{sat} on täydellisesti sopivan mallin uskottavuus. Pienempi D viittaa parempaan istuvuuteen (Agregti 2010, s. 67).

Pearsonin χ^2 -testi vertaa mallin tuottamia sovitteita \hat{y}_i havaittuihin arvoihin y_i :

$$\chi^2 = \sum_i \frac{(y_i - \hat{y}_i)^2}{\hat{y}_i}.$$

jossa summaus ulottuu kaikkien luokkien i yli. Testisuure X^2 noudattaa χ^2 -jakaumaa, jonka vapausasteet määräytyvät kaavalla $(r - 1)(c - 1)$, jossa r ja c ovat havaintotaulukon rivien ja sarakkeiden lukumäärät (Agresti 2010, s. 68). Testiä käytetään myös kohdassa 3.2, jossa verrataan Aallon suomalaisten opiskelijoiden syntymäkuukausijakaumaa Suomen kantaväestöön.

Devianssi ja Pearsonin χ^2 -testi täydentävät logistisen mallin erottelukyvyn arviointia (ks. kohta 4.1.1) sekä järjestyslogistisen mallin yhteensopivuuksindeksiä (ks. kohta 4.2.3). Erottelukäyräanalainen pinta-ala ja yhteensopivuuksindeksi mittaavat mallin kykyä järjestää havainnot oikeaan järjestykseen, devianssi ja Pearsonin χ^2 -testi arvioivat, kuinka hyvin mallin sovittamat todennäköisyydet vastaavat havaittua jakaumaa.

4.3.4 Pearsonin korrelaatiokerroin ja McFaddenin pseudo- R^2

Pearsonin korrelaatiokerroin r mittaa lineaarista yhteyttä havaittujen vasteiden y_i ja mallin sovitteiden \hat{y}_i välillä:

$$r = \frac{\sum_i (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_i (y_i - \bar{y})^2} \sqrt{\sum_i (\hat{y}_i - \bar{\hat{y}})^2}},$$

jossa summataan kaikkien havaintoparien yli (y_i, \hat{y}_i) , eli jokaiselle havainnolle ja sitä vastaavalle mallin sovitteelle. Kaavassa $r = 1$ tarkoittaa täydellistä positiivista lineaarista yhteyttä, $r = 0$ ei yhteyttä, ja $r = -1$ täydellistä negatiivista yhteyttä. Pearsonin korrelaatiokerroin olettaa, että vasteet heijastavat luokkien välisiä etäisyyksiä. Tämä ei välttämättä pidä paikkaansa järjestysasteikollisessa aineistossa, jossa tiedetään vain luokkien järjestys, ei etäisyyksiä niiden välillä (Agresti 2007, s. 41).

McFaddenin pseudo- R^2 mittaa, kuinka paljon mallin logaritminen uskottavuus suurenee verrattuna malliin, jossa ei ole selittäviä muuttujia ollenkaan:

$$R_{\text{McFadden}}^2 = 1 - \frac{\log L_{\text{sovitettu malli}}}{\log L_{\text{nolla malli}}},$$

jossa $\log L_{\text{malli}}$ on sovitetun mallin logaritminen uskottavuus ja $\log L_{\text{nolla}}$ nollamallin logaritminen uskottavuus. Arvoja välillä 0.2–0.4 pidetään hyvinä sosiaalitieteellisissä tutkimuksissa, mutta myös tätä pienemmät arvot antavat tilastollisesti merkitseviä yhteyksiä (McFadden 1974, s. 122).

Toisin kuin devianssi ja Pearsonin χ^2 -testi, jotka arvioivat mallin *istuvuutta* vertaamalla mallin sovittamia todennäköisyyksiä ja havaittuja arvoja frekvenssien tai uskottavuuksien tasolla (ks. kohta 4.3.3), Pearsonin korrelaatiokerroin ja McFaddenin pseudo- R^2 tarkastelevat mallin *selityskykyä*. Ne mittaavat, kuinka hyvin taustamuuttujat, kuten sukupuoli ja syntymäkuukausi, pystyvät selittämään vasteen vaihtelua. Näin ollen ne eivät kuvaa suoraan mallin ja havaintojen yhteensopivuutta, vaan mallin tuottamien sovitteiden ja todellisten vasteiden välistä yhteyttä.

4.3.5 Akaiken informaatiokriteeri (AIC)

Akaiken informaatiokriteeri (AIC) tasapainottaa mallin selitysvoinan ja yksinkertaisuuden: se suosii malleja, jotka selittävät aineistoa hyvin vähäparametrisesti:

$$\text{AIC} = -2 \log L + 2k,$$

jossa L on mallin maksimointi uskottavuus ja k on mallin vapausasteiden eli parametrien lukumäärä. Pienempi AIC-arvo viittaa parempaan malliin (Harrell 2015, s. 204).

4.3.6 Standardoidut ja kumulatiiviset jäännökset

Devianssi ja Pearsonin χ^2 -testi kohdassa 4.3.3 antavat yleiskuvan mallin sopivuudesta, mutta eivät kerro, jossa yksittäisissä soluissa malli mahdollisesti epäonnistuu. Tätä varten tarkastellaan *standardoituja jäännöksiä*, jotka mittaavat yksittäisen solun havaintojen n_{ij} ja mallin sovituksen $\hat{\mu}_{ij}$ erotusta:

$$r_{ij} = \frac{n_{ij} - \hat{\mu}_{ij}}{\sigma(n_{ij} - \hat{\mu}_{ij})},$$

jossa $\sigma(n_{ij} - \hat{\mu}_{ij})$ on jäännöksen keskihajonnan estimaatti. Tässä tutkimuksessa käytetään $|r_{ij}| > 3$ poikkeavien solujen tunnistamiseen. (Agresti 2010, s. 73–74)

Järjestyslogistisessa mallinnuksessa voidaan hyödyntää myös kumulatiivisiin summiin perustuvia standardoituja jäännöksiä. Ne mittaavat poikkeamia tilanteissa, joissa kiinnostuksen kohteena on vasteen luokkarajojen ylittäminen. Tietyille havaintoryhmälle i ja vasteen luokalle j kumulatiivinen jäännös määritellään kaavalla

$$r_{ij}^{\text{cum}} = \frac{\sum_{k=1}^j n_{ik} - \sum_{k=1}^j \hat{\mu}_{ik}}{\sqrt{\sum_{k=1}^j \hat{\mu}_{ik} [1 - P(Y \leq j | x_i)]}},$$

osoittaja kuvaa havaittujen ja odotettujen kumulatiivisten frekvenssien erotusta ja nimittäjä mallin mukaista keskihajontaa kyseiselle summalle. Vaikka nämä jäännökset eivät noudata täsmällisesti normaalijakaumaa, arvojen suuruutta voidaan tulkita vastaavasti kuin standardoitujen jäännösten kohdalla: suuret arvot ($|r_{ij}^{\text{cum}}| > 3$) viittaavat siihen, että malli ei sovi hyvin kyseiseen havaintoryhmään tai luokkarajaan (Agresti 2010, s. 73–74).

4.4 Yhteenveto menetelmistä

Taulukko 2 kokoaa keskeiset analyysikeinot ja niiden roolin mallien arvioinnissa.

Taulukko 2: Menetelmien merkitys tutkimuksessa.

menetelmä	kohta tekstissä	tarpeellisuus tutkimuksessa
Logistinen regressio	Kohta 4.1	Mallintaa kaksiarvoista vastemuuttujaa (hyväksytyt/hylätyt) ja mahdollistaa taustamuuttujien vaikutusten tulkinnan vastasuhteiden kautta.
Erottelukyky ja käyräanalainen pinta-ala	Kohta 4.1.1	Arvioi logistisen mallin kykyä erottaa luokat toisistaan; mittaa luokittelun tarkkuutta.
Verrannollisten vastasuhteiden malli	Kohta 4.2	Yleistää logistisen regression useampi-luokkaiseen vasteeseen käyttäen kertymätodennäköisyyksiä ja näiden vastasuhteita.
Brantin testi	Kohta 4.2.2	Testaa, päteekö järjestyslogistisessa mallissa oletus verrannollisista vastasuhteista.
Yhteensopivuusindeksi	Kohta 4.2.3	Mittaa järjestyslogistisen mallin kykyä järjestää havainnot oikein; yleistää käyräanalaisen pinta-alan moniluokkaiselle vasteelle.
Waldin testi	Kohta 4.3.1	Testaa yksittäisten mallikertoimien tilastollista merkitsevyyttä.
Uskottavuusosamäärätesti	Kohta 4.3.2	Yleinen testimenetelmä, joka vertaa kahden mallin sopivuutta logaritmisesti uskottavuudella.
Devianssi ja Pearsonin χ^2 -testi	Kohta 4.3.3	Arvioivat, kuinka hyvin mallin sovittamat todennäköisyydet vastaavat havaittua jakaumaa.
Pearsonin korrelaatiokerroin ja McFaddenin pseudo- R^2	Kohta 4.3.4	Mittaavat kuinka hyvin selittävät muuttujat selittävät vasteen vaihtelua.
Akaiken informaatiokriteeri (AIC)	Kohta 4.3.5	Tasapainottaa mallin selitysvoiman ja selityskertoimien määrän rangaisten monimutkaisista malleista.
Standardoidut ja kumulatiiviset jäännökset	Kohta 4.3.6	Kertoo, yksittäisissä soluissa tai havaintoryhmissä malli ei sovi hyvin.

5 Tulokset

Sekä logistiset regressiomallit kohdissa 5.1.1 ja 5.1.2, että verrannollisten vastasuh- teiden mallit kohdissa 5.2.1 ja 5.2.3 rakennetaan seuraavien muuttujien pohjalta:

Taulukko 3: Logistisen regressiomallin selittävät muuttujat, niiden vertailuluokat ja selitykset.

muuttuja	vertailuluokka	kuvaus
Sukupuoli	Mies	Naisopiskelijoiden suoritumista verrataan mies- opiskelijoihin.
Kansalaisuus	Suomi	Kansainvälisten opiskelijoiden tuloksia verrataan suomalaisiin opiskelijoiden tuloksiin.
Korkeakoulu	SCI	Muiden korkeakoulujen suorituksia verrataan SCI-opiskelijoihin suorituksiin.
Syntynyt Alkuvuonna	Huhti–Marraskuu	Tammi–maaliskuussa syntyneitä verrataan huh- ti–marraskuuhun syntyneisiin.
Syntynyt Joulukuussa	Huhti–Marraskuu	Joulukuussa syntyneitä verrataan huh- ti–marraskuuhun syntyneisiin.
Vuodet 2020–2021	2014–2019, 2023–2024	Suorituksia vuosilta 2020–2021 verrataan pan- demiaa edeltävään ja jälkeiseen aikaan.
Vuosi 2022	2014–2019, 2023–2024	Vuoden 2022 vaikutuksia verrataan pandemiaa edeltävään ja jälkeiseen aikaan.
Arvosanatrendi	–	Jatkuva muuttuja arvosanojen lineaarisen kehi- tyksen arvioimiseksi ajan suhteen.

Selittävät muuttujat on koottu taulukkoon 3. Valtaosa muuttujista on kaksiluok- kaisia, mikä sopii hyvin logistisen ja järjestyslogistisen regressiomallin perusperi- aatteeseen: mallin tavoitteena on arvioida, miten yksittäinen ominaisuus vaikuttaa hyväksytyksi tulemisen todennäköisyyteen tai arvosanan todennäköisyysjakaumaan, suhteessa valittuun vertailuluokkaan.

Muuttuja Arvosanatrendi on jatkuva ja mittaa arvosanojen kehitystä vuosien välillä. Lisäksi Vuodet 2020–2021 ja Vuosi 2022 on eroteltu toisistaan, koska etäopetuksen toteutuksessa oli olennaisia eroja: vuosina 2020–2021 opetus oli lähes täysin etämuotoista, kun taas vuonna 2022 toteutus vaihteli jaksoittain.

5.1 Hyväksytyt ja hylätyt suoritukset

5.1.1 Peruskursseittain

Analyysissä käytettiin Aalto-yliopiston vuosien 2014–2024 kurssisuoritusaineistoa, joka kattaa seuraavat matematiikan peruskurssit:

- Differentiaali- ja integraalilaskenta 1
- Differentiaali- ja integraalilaskenta 2
- Matriisilaskenta
- Todennäköisyyslaskennan ja tilastotieteen peruskurssi

Ensin sovittiin logistinen regressiomalli, jossa hyväksytyksi tulemista selitettiin opiskelijan sukupuolella, kansalaisuudella, korkeakoululla, syntymäkuukaudella, koronavuosilla sekä arvosanojen pitkittäistrendillä. Muuttuja *syntynyt alkuvuonna* ei ollut tilastollisesti merkitsevä ($p = 0,97$) ja poistettiin mallista. Malliksi sovittiin:

```
glm(hyväksytty ~ Sukupuoli + Kansalaisuus + Korkeakoulu + 'Syntynyt Alkuvuonna'
+ 'Syntynyt Joulukuussa' +
'Korona 2020-2021' + 'Korona 2022' + Arvosanatrendi,
data = aineisto, family = binomial)
```

Taulukko 4: Perusmallin selittävät muuttujat, UOM-testin arvot ja vaikutukset hyväksymistodennäköisyyteen.

Vertailuryhmä: suomalainen mies, syntynyt huhtikuun ja marraskuun välillä.

muuttuja	estimaatti	keskivirhe	UOM-testi	p -arvo	$\hat{\pi}$ ⁵
(Vakiotermi)	1.43	0.025	—	—	0.81
Nainen	0.40	0.026	242.24	$< 2 \cdot 10^{-16}$	0.86
Kansainv. opiskelija	-0.14	0.057	6.28	0.012	0.78
Korkeakoulu	—	—	1424.57	$< 2 \cdot 10^{-16}$	—
BIZ	-0.89	0.067			0.63
CHEM	-0.43	0.037			0.73
ELEC	-0.14	0.032			0.79
ENG	-0.32	0.030			0.75
ARTS	-1.46	0.303			0.49
Avoin yliopisto	-2.00	0.056			0.36
Syntynyt joulukuussa	0.13	0.043	9.39	0.0022	0.83
Vuodet 2020–2021	0.53	0.031	317.05	$< 2 \cdot 10^{-16}$	0.88
Vuosi 2022	-0.11	0.036	9.53	0.0020	0.79
Arvosanatrendi	-0.04	0.004	117.69	$< 2 \cdot 10^{-16}$	0.80

Huomioita taulukosta 4.

- Uskottavuusosamäärätesti (*UOM-testi*, ja tämän p -arvo) on laskettu vertaamalla kutakin termiä sisältävää mallia sitä ilman kyseistä termiä ("drop-term"-vertailu). Raportoidut arvot ovat uskottavuusosamäärä-testisuureita, ja tulkin-taa täydennettiin tarkastelemalla AIC:n muutosta: jos muuttujan lisääminen pienentää AIC-arvoa merkittävästi, pääteltiin sen parantavan mallin sopivuutta.

⁵Mallin sovittama hyväksymistodennäköisyys jokaiselle muuttujalle.

- Hyväksymistodennäköisyys on saatu logistisen funktion kaavan (4) avulla

$$\hat{\pi}_j = \frac{\text{vastasuhde}_j}{1 + \text{vastasuhde}_j}.$$

Perusmallin jäännösanalyysi paljasti useita poikkeavia havaintoja, joissa standardoitu Pearsonin jäännös ylitti arvon 2. Poikkeamat painottuivat erityisesti naisopiskelijoihin, jotka opiskelivat *SCI*- tai *ELEC*-korkeakoulussa, olivat suomalaisia ja suorittivat kurssin koronavuosina 2020–2021. Tämä viittasi mahdollisiin interaktiovaikutuksiin.

Sovitettiin laajennettu interaktiomalli, joka sisälsi korkeakoulun ja sukupuolen välisen interaktion sekä kolmoisinteraktion korkeakoulun, sukupuolen ja vuoden 2020–2021 välillä. Interaktiomalli paransi useita mittareita: AIC pieneni 103 yksikköä, McFaddenin R^2 kasvoi arvoon 0.0390 ja käyräalainen pinta-ala nousi hieman (0.613 → 0.616). Laajennettu malli on muotoa:

```
glm(hyväksytty ~ Kansalaisuuskattegoria +
  'Op.oik. organisaatio' * Sukupuoli +
  'Syntynyt Joulukuussa' + 'Korona 2020_2021' + 'Korona 2022' +
  arvosanatrendi_keskiarvona +
  'Op.oik. organisaatio':'Korona 2020_2021' +
  'Op.oik. organisaatio':Sukupuoli:'Korona 2020_2021',
  data = aineisto, family = binomial)
```

Taulukko 5: Perusmallin ja interaktiomallin yhteenveto.

malli	n	AIC	McFadden- R^2	Pearson r	käyräalainen pinta-ala
Perusmalli	52 547	52 080	0.0364	0.208	0.613
Interaktiomalli	52 528	51 977	0.0390	0.215	0.616

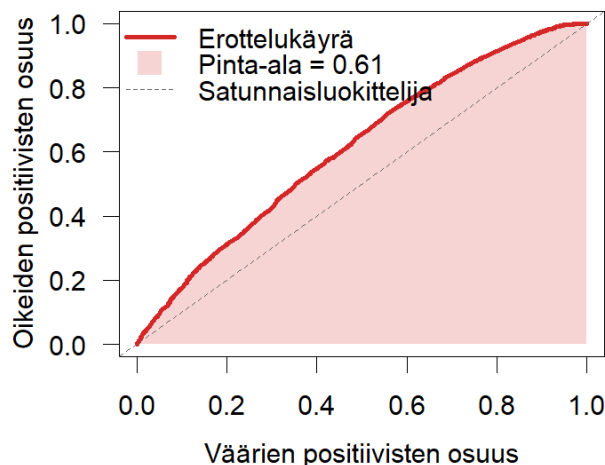
Taulukko 6: Pearson- ja devianssijäännösten jakauma molemmissa malleissa.

	Perusmalli	Interaktiomalli
Pearsonin jäännökset		
min	-3.44	-3.84
1. kvart. / mediaani / 3. kvart.	0.366 / 0.463 / 0.533	0.360 / 0.457 / 0.526
max	1.41	1.60
$ r > 3$ (lukumäärä)	102	136
Devianssijäännökset		
min	-2.26	-2.35
1. kvart. / mediaani / 3. kvart.	0.502 / 0.623 / 0.707	0.494 / 0.616 / 0.700
max	1.48	1.59
$ r > 3$ (lukumäärä)	0	1

Vaikka interaktiomalli tuotti hieman paremman selityskyvyn ja yhteensopivuuden kuin perusmalli (taulukko 5), jäännöstarkastelu osoitti, että suuria jäännöksiä ($|r| > 3$)

esiintyi edelleen molemmissa malleissa (taulukko 6). Interaktiomallissa noin 0.26% havainnoista.

Suurimmat jäännökset esiintyivät koronavuosina 2020–2021 naisopiskelijoilla, erityisesti ELEC- ja BIZ-kouluista sekä kansainvälisten opiskelijoiden ryhmissä. Esimerkiksi matriisilaskennan kurssilla malli ennusti useissa tapauksissa hyväksyttyä suoritusta, vaikka arvosana oli todellisuudessa hylätty.



Kuva 8: ROC-käyrä peruskurssien logistiselle mallille. Harmaa katkoviiva vastaa satunnaisluokittelijaa. AUC = 0.613.

5.1.2 Kaikkien kurssien hyväksymiseen vaikuttavat tekijät

Laajennettiin aineisto koskemaan kaikkia Aalto-yliopistossa vuosina 2014–2024 suoritettuja matematiikan kursseja ($n = 60244$). Samoin kuin kohdassa 5.1.1 sovitettiin ensin perusmalli ilman interaktiovaikutuksia. Myös tässä mallissa muuttuja *syntynyt alkuvuonna* osoittautui tilastollisesti ei-merkittäväksi ($p = 0.84$) ja poistettiin mallista.

```
glm(hyväksytty ~ Sukupuoli + Kansalaisuus + Korkeakoulu + 'Syntynyt Alkuvuonna'
+ 'Syntynyt Joulukuussa' +
'Korona 2020-2021' + 'Korona 2022' + Arvosanatrendi,
data = aineisto, family = binomial)
```

Taulukko 7: Perusmallin ja interaktiomallin yhteenvedo (kaikki matematiikan kurssit).

malli	n	AIC	McFadden- R^2	Pearson r	käyräalainainen pinta-ala
Perusmalli	60 244	58,839	0.0376	0.214	0.619
Interaktiomalli	60 219	58,731	0.0404	0.220	0.622

Taulukko 8: Pearson- ja devianssijäännösten jakauma molemmissa malleissa (kaikki kurssit).

	Perusmalli	Interaktiomalli
Pearsonin jäännökset		
min	-3.56	-3.91
1. kvart. / mediaani / 3. kvart.	0.369 / 0.468 / 0.538	0.363 / 0.461 / 0.530
max	1.45	1.63
$ r > 3$ (lukumäärä)	166	150
Devianssijäännökset		
min	-2.34	-2.46
1. kvart. / mediaani / 3. kvart.	0.505 / 0.628 / 0.712	0.499 / 0.622 / 0.705
max	1.53	1.61
$ r > 3$ (lukumäärä)	0	1

Taulukko 9: Perus- ja interaktiomallin välinen vertailu (kaikki kurssit).

vertailu	ΔAIC	ΔR^2	UOM-arvo (df)	p -arvo
Perus \rightarrow Interaktio	-108	+0.0028	149.6 ($df = 35$)	$< 10^{-16}$

Perusmallin tulokset heijastavat pitkälti peruskurssien mallia: naisilla oli korkeampi todennäköisyys tulla hyväksytyksi, kansainvälisillä opiskelijoilla todennäköisyys oli matalampi, ja koronapandemian aikaisina vuosina 2020–2021 hyväksytyjen osuus oli poikkeuksellisen korkea. Lisäksi kurssien arvosanatrendi hyväksymistä laski ajan myötä.

Myös tässä aineistossa standardoidut jäännökset paljastivat toistuvia poikkeamia erityisesti suomalaisilla naisopiskelijoilla ELEC- ja SCI-korkeakouluissa koronavuosina 2020–2021. Tämä motivoi interaktiotermin lisäämisen.

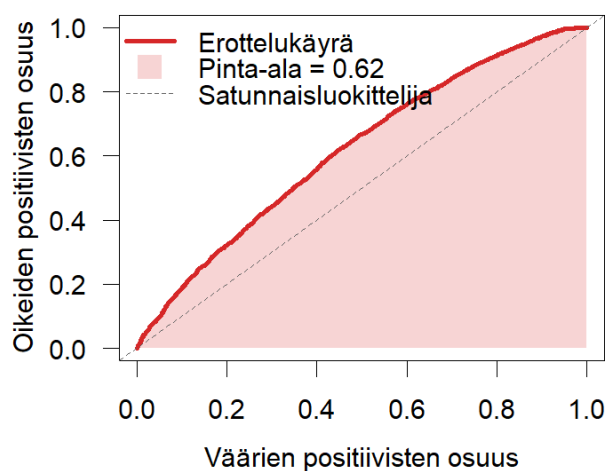
Sovitettiin laajennettu malli, jossa mukana olivat korkeakoulun ja sukupuolen välinen interaktio sekä kolmoisinteraktio vuosien 2020–2021 kanssa. Interaktiomalli paransi istuvuutta selkeästi (AIC väheni 108 yksikköä, McFaddenin R^2 kasvoi arvoon 0.0404 ja AUC nousi arvoon 0.622). Uskottavuusosamäärätesti osoitti mallien välisen eron olevan tilastollisesti erittäin merkitsevä ($p < 10^{-16}$, taulukko 9).

Interaktiomallin parantunut istuvuus osoittaa, että hyväksymistodennäköisyyteen vaikuttavat paitsi yksittäiset taustatekijät myös niiden yhdistelmät. Erityisesti korkeakoulun, sukupuolen ja pandemiavuoden yhteisvaikutus. Silti yksittäisiä suuria jäännöksiä (noin 0.25%, jossa $|r| > 3$) jäi jäljelle, mikä viittaa siihen, ettei mallissa ole täydellinen istuvuus.

Taulukko 10: Mallin muuttujat, UOM-arvot ja vaikutukset hyväksymistodennäköisyyteen.

Vertailuryhmä: suomalainen mies, syntynyt huhtikuun ja marraskuun välillä.

muuttuja	estimaatti	keskivirhe	UOM-testi	<i>p</i> -arvo	$\hat{\pi}^6$
(Vakiotermi)	1.52	—	—	—	0.82
Nainen	0.37	—	233.92	$< 2 \cdot 10^{-16}$	0.87
Kansainv. opiskelija	-0.07	—	2.04	0.153	0.81
Korkeakoulu	—	—	1615.33	$< 2 \cdot 10^{-16}$	—
BIZ	-0.90	—			0.65
CHEM	-0.53	—			0.73
ELEC	-0.23	—			0.78
ENG	-0.42	—			0.75
ARTS	-1.53	—			0.50
Avoin yliopisto	-2.10	—			0.36
Syntynyt joulukuussa	0.10	—	6.23	0.013	0.83
Vuodet 2020–2021	0.60	—	455.91	$< 2 \cdot 10^{-16}$	0.89
Vuosi 2022	-0.18	—	28.84	$< 10^{-7}$	0.79
Arvosanatrendi	-0.04	—	124.29	$< 2 \cdot 10^{-16}$	0.81



Kuva 9: ROC-käyrä kaikkien kurssien logistiselle mallille. Harmaa katkoviiva vastaa satunnaisluokittelijaa.

5.2 Osittain verrannollisten vastasuhteiden malli

Analyysi aloitettiin tutkimalla päteekö kohdan 4.2.1 verrannollisten vastasuhteiden oletus malleille. Tätä testattiin Brantin testillä (ks. kohta 4.2.2). Tulokset on esitetty seuraavissa taulukoissa. Omnibus-rivi kuvaa verrannollisten vastasuhteiden oletuksen

⁶Mallin sovittama hyväksymistodennäköisyys jokaiselle muuttujalle.

kokonaistestiä, kun taas muuttujakohtaiset rivit kertovat missä muuttujissa oletus on rikkoutunut.

Taulukko 11: Brantin testin tulokset verrannollisten vastasuhteiden oletukselle (peruskurssit).

muuttuja	χ^2	<i>df</i>	<i>p-arvo</i>
Omnibus-testi	1337.18	13	0
Nainen	7.02	1	0.01
Kansainvälinen opiskelija	59.41	1	0
Avoin yliopisto	79.44	1	0
ELEC (vs SCI)	137.05	1	0
BIZ (vs SCI)	51.18	1	0
CHEM (vs SCI)	303.02	1	0
ENG (vs SCI)	393.73	1	0
ARTS (vs SCI)	2.53	1	0.11
Tammikuu–maaliskuu	1.91	1	0.17
Joulukuu	0.61	1	0.43
Korona 2020/21	38.46	1	0
Korona 2022	84.71	1	0
Arvosanatrendi	9.08	1	0

Taulukko 12: Brantin testin tulokset verrannollisten vastasuhteiden oletukselle (kaikki kurssit).

muuttuja	χ^2	<i>df</i>	<i>p-arvo</i>
Omnibus-testi	1381.46	13	0
Nainen	7.42	1	0.01
Kansainvälinen opiskelija	68.76	1	0
Avoin yliopisto	103.61	1	0
ELEC (vs SCI)	231.28	1	0
BIZ (vs SCI)	75.18	1	0
CHEM (vs SCI)	260.12	1	0
ENG (vs SCI)	335.52	1	0
ARTS (vs SCI)	3.88	1	0.05
Tammikuu–maaliskuu	1.01	1	0.31
Joulukuu	0.34	1	0.56
Korona 2020/21	17.61	1	0
Korona 2022	73.37	1	0
Arvosanatrendi	12.03	1	0

Brantin testin perusteella molemmissa aineistoissa päädyttiin soveltamaan osittaisen verrannollista vastasuhteiden mallia, jossa verrannollisten vastasuhteiden oletus säilytettiin syntymäkuukausimuuttujien osalta, mutta muiden muuttujien kohdalla ei.

5.2.1 Peruskurssien arvosanatasoihin vaikuttavat tekijät

5.2.2 Peruskursseittain

Aineistona käytettiin samaa Aalto-yliopiston vuosien 2014–2024 peruskurssiaineistoa kuin logistisessa regressiossa (ks. kohta 31). Nyt ei analysoitu pelkästään hyväksytyjen ja hylättyjen arvosanojen välistä eroa, vaan hyväksytyt arvosanat jaettiin kolmeen järjestysluokkaan:

- Luokka 0: Hylätty
- Luokka 1: tyydyttävä (arvosanat 1–3)
- Luokka 2: hyvä (arvosanat 4–5)

Muuttuja Korkeakoulu ei ole mukana päävaikutuksena (numeerisen epävakauden vuoksi), mutta sitä käytettiin interaktiossa. Ensimmäisessä vaiheessa sovitettiin perusmalli ilman interaktiotermiä, oletuksella että kaikki selittävät muuttujat täyttävät verrannollisuusoletuksen. Malli sovitettiin R:n VGAM paketin `vglm()`-funktioilla.

Verrannollisuusoletuksen testaus osoitti, että tämä rikkoutui tilastollisesti merkitsevästi monien muuttujien kohdalla (ks. taulukko 11). Tämän vuoksi siirryttiin osittaisen verrannollisten vastasuhteiden rakenteeseen, jossa oletus säilyi muuttujilla Syntynyt_alkuvuonna ja Syntynyt_joulukuussa. Muut muuttajat saivat omat kertoimet jokaisella kynnyksellä.

```
vglm(Arvosana_luokittelu ~ Sukupuoli + Kansalaisuus +  
Syntynyt_alkuvuonna + Syntynyt_joulukuussa +  
korona_2020_2021 + korona_2022 + arvosanatrendi_keskiarvona,  
data = aineisto,  
family = cumulative(link = "logitlink",  
parallel = FALSE ~ Sukupuoli + Kansalaisuus +  
korona_2020_2021 + korona_2022 +  
arvosanatrendi_keskiarvona))
```

Jäännösanalyysi viittasi lisäksi mahdollisiin interaktioihin korkeakoulun ja sukupuolen sekä koronavuosien ja sukupuolen välillä. Tämän vuoksi mallia laajennettiin vaiheittain:

```
vglm(Arvosana_luokittelu ~ Sukupuoli + Kansalaisuus +  
Syntynyt_alkuvuonna + Syntynyt_joulukuussa +  
korona_2020_2021 + korona_2022 + arvosanatrendi_keskiarvona +  
Sukupuoli:korona_2020_2021 + 'Op.oik. organisaatio':Sukupuoli,  
data = aineisto,  
family = cumulative(link = "logitlink",  
parallel = FALSE ~ Sukupuoli + Kansalaisuus +  
korona_2020_2021 + korona_2022 +  
arvosanatrendi_keskiarvona +  
Sukupuoli:korona_2020_2021 +  
'Op.oik. organisaatio':Sukupuoli))
```

Lopullinen malli sisälsi molemmat interaktiot. Mallien vertailu esitetään alla:

Taulukko 13: Mallien vertailu peruskurssien osalta.

malli	AIC	McFadden- R^2	Pearson r	yhteensopivuusindeksi
Perusmalli	105 522	0.0121	0.133	0.572
Korona \times Sukupuoli	105 504	0.0123	0.133	0.573
Korkeakoulu \times Sukupuoli	103 280	0.0334	0.208	0.640
<i>Molemmat interaktiot</i>	<i>103 260</i>	<i>0.0336</i>	<i>0.209</i>	<i>0.640</i>

Mallin kokonaissovittuvuus oli kohtalainen. Pearsonin χ^2 -testi ei viitannut poikkeamiin todennäköisyyksissä havaintojen suhteen ($p = 0.388$), kun taas devianssitesti oli merkitsevä ($p = 0.00047$), mikä viittaa poikkeamaan saturaatiomallista. Devianssin pieni p -arvo on odotettavissa suurilla otoksilla, jolloin pienetkin poikkeamat johtavat tilastollisesti merkitsevään p -arvoon.

Alla on esitetty lopullisen mallin standardoitujen Pearsonin jäännösten jakauma:

Taulukko 14: Standardoitujen Pearsonin jäännösten jakauma.

kynnysfunktio	minimi	1. kvart.	mediaani	keskiarvo	3. kvart.	maksimi
$\text{logit}(P(Y \leq 2))$	-1.78	-0.93	0.33	0.001	0.86	3.19

Kaksi jäännöstä ylitti arvon kolme: molemmat olivat BIZ-korkeakoulusta, mies-opiskelijoita vuodelta 2022. Laajempaa poikkeamaa ei havaittu. Kumulatiivisissa jäännöksissä ei ollut poikkeamia.

Taulukko 15: Mallin selittävät muuttajat, uskottavuusosamäärätestin arvot ja vaikutukset kertymätodennäköisyyksiin.

Vertailuryhmä: suomalainen mies, syntynyt huhtikuun ja marraskuun välillä.

muuttuja	kynnys	estimaatti	keskivirhe	UOM-testi ⁷	<i>p</i> -arvo	ristisuhde
Sukupuoli (Nainen)	Kynnys 1	-0.298	0.025	14.50	$1.4 \cdot 10^{-4}$	0.742
	Kynnys 2	-0.192	0.020	0.06	0.812	0.825
Kansainv. opiskelija	Kynnys 1	0.249	0.055	6.02	0.0141	1.283
	Kynnys 2	-0.255	0.048	30.72	$2.98 \cdot 10^{-8}$	0.775
Syntynyt alkuvuonna	Kynnys 1	0.012	0.025	1.07	0.301	1.012
	Kynnys 2	0.012	0.025	1.07	0.301	1.012
Syntynyt joulukuussa	Kynnys 1	-0.132	0.043	8.68	0.00321	0.877
	Kynnys 2	-0.132	0.043	8.68	0.00321	0.877
Vuodet 2020–2021	Kynnys 1	-0.425	0.029	136.90	$< 2 \cdot 10^{-16}$	0.654
	Kynnys 2	-0.668	0.023	476.95	$< 2 \cdot 10^{-16}$	0.513
Vuosi 2022	Kynnys 1	0.096	0.035	5.94	0.0148	1.101
	Kynnys 2	-0.206	0.031	48.58	$3.17 \cdot 10^{-12}$	0.814
Arvosanatrendi	Kynnys 1	0.0426	0.0035	67.67	$< 2 \cdot 10^{-16}$	1.044
	Kynnys 2	0.0264	0.0030	70.28	$< 2 \cdot 10^{-16}$	1.027

Taulukko 16: Mallin sovittamat todennäköisyydet arvosanaluokkiin (peruskurssit).

profiili	hylätty	1–3	4–5
Vertailuluokka (Mies, Suomi, SCI)	0.1849	0.2852	0.5299
Nainen	0.1548	0.3128	0.5324
Kansainvälinen opiskelija	0.2078	0.1933	0.5990
Syntynyt alkuvuonna	0.1880	0.2872	0.5248
Syntynyt joulukuussa	0.1709	0.2755	0.5536
Vuodet 2020–2021	0.1313	0.1941	0.6745
Vuosi 2022	0.1988	0.2162	0.5850
Arvosanatrendi +1	0.1895	0.2871	0.5234

Taulukko 16 osoittaa, että sukupuolella oli vaikutus: naisten hylkäytodennäköisyys oli matalampi ja heidän todennäköisyytensä saada hyvä arvosana oli melkein sama kuin

⁷Laskettu R:n VGAM-paketin `lrt.stat()`-funktiolla.

miehillä (0.5324 ja 0.5299). Kansainväliset opiskelijat erottuivat selvimmin: heillä oli korkeampi riski hylätyksi tulemiseen, mutta myös selvästi suurempi todennäköisyys saada hyvä arvosana. Syntymäkuukauden vaikutukset jäivät pieneksi, joskin joulukuussa syntyneet menestyivät paremmin kuin alkuvuonna syntyneet. Koronavuosien vaikutus oli merkittävä: vuosina 2020–2021 arvosanat olivat selvästi korkeampia ja hylkäykset harvinaisempia. Vuonna 2022 tilanne heikkeni hieman, mutta vielä selvästi parempi kuin vertailutilanteessa. Arvosanatrendi oli hieman laskussa.

5.2.3 Kaikkien kurssien arvosanatasoihin vaikuttavat tekijät

Laajennettiin analyysi koskemaan kaikkia kursseja kuten logistisessa regressiossa. (ks. kohta 34). Arvosanat jaettiin jälleen järjestysasteikolliseen muotoon seuraaviin kolmeen luokkaan:

- Luokka 0: hylätty
- Luokka 1: tyydyttävä (arvosanat 1–3)
- Luokka 2: hyvä (arvosanat 4–5)

Mallina käytettiin osittain verrannollisten vastasuhteiden mallia. Verrannollisuusoletus säilytettiin vain syntymäkuukauteen liittyvillä muuttujilla (Syntynyt alkuvuonna, Syntynyt joulukuussa). Muille muuttujille oletus ei pätenyt (ks. taulukko 12).

```
vglm(
  Arvosana_luokittelu ~ Sukupuoli + Kansalaisuus +
  Syntynyt_alkuvuonna + Syntynyt_joulukuussa +
  korona_2020_2021 + korona_2022 + arvosanatrendi_keskiarvona +
  Sukupuoli:korona_2020_2021 + 'Op.oik. organisaatio':Sukupuoli,
  data = aineisto,
  family = cumulative(link = "logitlink",
    parallel = FALSE ~ Sukupuoli + Kansalaisuus +
    korona_2020_2021 + korona_2022 +
    arvosanatrendi_keskiarvona +
    Sukupuoli:korona_2020_2021 +
    'Op.oik. organisaatio':Sukupuoli)
)
```

Lopullinen malli sisälsi sekä koronavuosien että korkeakoulun ja sukupuolen väliset yhteisvaikutukset. Mallien vertailu esitetään taulukossa 17.

Taulukko 17: Mallien vertailu kaikkien kurssien osalta.

malli	AIC	mcFadden R^2	Pearson r	yhteensopivuusindeksi
Perusmalli	121101.5	0.0128	0.140	0.572
Korona \times Sukupuoli	121071.5	0.0131	0.141	0.579
Korkeakoulu \times Sukupuoli	118509.1	0.0342	0.218	0.640
<i>Molemmat interaktiot</i>	<i>118478.5</i>	<i>0.0345</i>	<i>0.219</i>	<i>0.642</i>

Lopullisen mallin istuvuus oli kohtalainen: Pearsonin χ^2 -testi ei viitannut poikkeamiin havaintojen ja mallin sovitteiden välillä ($p = 0.40$), mutta devianssitestin tulos ($p = 0.0026$) viittasi poikkeamiin täydellisestä mallista.

Alla on esitetty kummankin kynnysfunktion standardoitujen Pearsonin jäännösten jakaumat:

Taulukko 18: Standardoitujen Pearsonin jäännösten jakauma (viimeinen kynnysfunktio).

	minimi	1. kvart.	mediaani	keskiarvo	3. kvart.	maksimi
$\text{logit}(P(Y \leq 2))$	-1.766	-0.933	0.339	0.001	0.879	2.897

Jäännöksissä ei havaittu poikkeamia yli kolmen arvoa.

Taulukko 19: Mallin selittävät muuttujat, uskottavuusosamäärätestin arvot ja vaikutukset kertymätodennäköisyyksiin.

Vertailuryhmä: suomalainen mies, syntynyt huhtikuun ja marraskuun välillä.

muuttuja	kynnys	estimaatti	UOM-testi ⁸	p -arvo	ristisuhde
Sukupuoli (Nainen)	Kynnys 1	-3.296	10.86	$9.80 \cdot 10^{-4}$	0.037
	Kynnys 2	-0.102	0.01	0.919	0.903
Kansainv. opiskelija	Kynnys 1	1.548	2.39	0.122	4.700
	Kynnys 2	-7.572	57.34	$3.67 \cdot 10^{-14}$	0.001
Syntynyt alkuvuonna	Yhteinen	1.027	1.06	0.304	2.794
Syntynyt joulukuussa	Yhteinen	-2.484	6.17	0.0130	0.083
Vuodet 2020–2021	Kynnys 1	-14.502	210.30	$< 10^{-47}$	< 0.001
	Kynnys 2	-23.961	574.15	$< 10^{-127}$	< 0.001
Vuosi 2022	Kynnys 1	4.848	23.50	$1.25 \cdot 10^{-6}$	127.45
	Kynnys 2	-3.699	13.68	$2.16 \cdot 10^{-4}$	0.0247
Arvosanatrendi	Kynnys 1	8.576	73.54	$< 10^{-17}$	5 300.5
	Kynnys 2	8.593	73.84	$< 10^{-17}$	5 393.8

Alla olevassa taulukossa on esitetty sovitetut todennäköisyydet eri opiskelijaprofileille:

⁸Laskettu R:n VGAM-paketin `lrt.stat()`-funktioilla.

Taulukko 20: Mallin sovitetut todennäköisyydet arvosanaluokkiin (kaikki kurssit).

profili	hylätty	1–3	4–5
Vertailuluokka (Mies, Suomi, SCI)	0.1725	0.2878	0.5397
Nainen	0.1490	0.3102	0.5407
Kansainvälinen opiskelija	0.1846	0.1926	0.6228
Syntynyt alkuvuonna	0.1752	0.2898	0.5350
Syntynyt joulukuussa	0.1620	0.2797	0.5583
Koronavuodet 2020–2021	0.1145	0.2001	0.6854
Vuosi 2022	0.1967	0.2370	0.5663
Arvosanatrendi +1	0.1767	0.2897	0.5335

Kuten peruskursseillakin, sukupuolen vaikutus jäi kaksiosaiseksi. Naisilla oli pienempi tulla hylätyksi kurssissa, mutta heidän todennäköisyytensä saada hyvä arvosana oli lähes identtinen miesten kanssa. Kansainväliset opiskelijat erosivat selvästi: heillä oli toisaalta suurempi hylkäysriski, mutta myös selvästi suurempi todennäköisyys saada hyvä arvosana.

Alkuvuoden syntymäkuukausieffekti jäi pieneksi, mutta joulukuussa syntyneet menestyivät hieman paremmin. Koronavuosien vaikutus näkyi tässäkin aineistossa selvänä: vuosina 2020–2021 arvosanat olivat korkeampia ja hylkäysten osuus pienempi, kun taas vuonna 2022 kehitys taittui. Kokonaisarvosanatrendi oli hieman laskeva.

6 Pohdinta

6.1 Tulosten tulkinta ja yhteys kirjallisuuteen

Tulokset osoittivat, että naisilla oli pienempi hylkäysprosentti kuin miehillä, mutta parhaimpien suorittajien joukossa ei löytynyt eroja. Tämä havainto on linjassa aiempiin tuloksiin nähden (Lindberg et al. 2010; Hyde et al. 2008), joissa sukupuolten väliset erot eivät olleet merkittäviä.

Syntymäkuukauden osalta alkuvuonna syntyneillä ei havaittu etua suhteessa loppuvuonna syntyneisiin, mikä poikkeaa selvästi suhteellisen iän vaikutusta käsittelevistä tutkimuksista (esim. Thoren, Heinig ja Brunner 2016; Kaila 2017; Ukkola ja Metsämuuronen 2023), joissa alkuvuonna syntyneet ovat yleensä suoriutuneet merkittävästi paremmin. Lisäksi joulukuussa syntyneiden hieman parempi menestyminen oli yllättävä tulos, joka ei myöskään vastaa aiempaa kirjallisuutta. On mahdollista, että osa joulukuussa syntyneistä on aloittanut koulun vuotta myöhemmin esimerkiksi myöhäisemmän kouluvalmiuden vuoksi, mutta tutkimusaineistosta tätä ei voida varmistaa.

Kansainvälisten opiskelijoiden osalta tulokset olivat kaksijakoisia: heitä esiintyi vähän enemmän hylättyjen joukossa, mutta merkittävästi enemmän parhaiden suoritusten joukossa. Tämä ei täysin vastaa aiempaa PISA-pohjaista tutkimusta (esim. Planting 2022; OECD 2023), jossa maahanmuuttajataustaiset opiskelijat keskimääräinen osaaminen on heikompaa kuin kantaväestöstä. Korkeakoulujen väliset erot noudattelivat Miihkisen aiempia havaintoja: Perustieteiden korkeakoulun opiskelijoilla oli keskimäärin paremmat suoritukset verrattuna muihin korkeakouluihin. (Miihkinen 2013)

Arvosanainflaation osalta tulokset viittasivat siihen, että arvosanat ovat laskeneet hieman pitkällä aikavälillä, mikä on päinvastaista esimerkiksi Yhdysvaltojen ja Suomen aiemmissa trendeissä havaittuun arvosanainflaatioon (esim. Rojstaczer ja Healy 2012; Karvi 2023). Tämä voi johtua siitä, että opiskelijamäärät ovat kasvaneet, jolloin suurempi opiskelijajoukko lisää matemaattisen osaamisen variaatiota ja heikentää läpäisyprosentteja, kuten kuvassa 3 havaittiin.

Koronavuosina 2020–2021 arvosanat kuitenkin kohosivat, mahdollisesti etäopetuksen vuoksi. Vuonna 2022 arvosanat laskivat jyrkästi, mikä näkyy myös laskevassa arvosanatrendissä koronavuosien poikkeustuloksista huolimatta. On epäselvää, johtuuko lasku opiskelijamäärään kasvusta, heikentyneestä osaamisesta vai muista tekijöistä.

6.2 Tulosten merkitys

Syntymäkuukauden vaikutus arvosanoihin oli tilastollisesti merkityksetön, viitaten siihen, että suhteellisen iän ilmiö, joka on voimakas peruskoulutasolla, vaimenee ja käytännössä katoaa yliopistolla. Tämä eroaa esimerkiksi urheilun kontekstista, jossa suhteellinen ikäetu säilyy ja korostuu aikuisikäen asti (R. H. Barnsley, Thompson ja P. E. Barnsley 1985).

Kansainvälisten opiskelijoiden tulokset osoittivat yllättävää kaksijakoisuutta: heitä esiintyi sekä hylättyjen että huippuosaajien joukossa enemmän kuin suomalaistaus-

taisissa opiskelijoissa. Tämä laajentaa aiempaa keskustelua, jossa on usein verrattu vain kantasuomalaisia ja maahanmuuttajataustaisia opiskelijoita (esim. Planting 2022; OECD 2023). Aiemmat tutkimukset eivät tarkastele erityisesti huippuosajia, mutta tämän tutkimuksen tulokset viittaavat siihen, että suomalaisten matematiikan ydinosaajien taso on kansainvälisesti kilpailukykyinen. Kansainvälisessä joukossa osaamisen hajonta on kuitenkin suurempaa suorituksissa.

Myös arvosanainflaation osalta tutkimus tarjoaa uutta tietoa. Kyseisissä matematiikan kursseissa kymmenen vuoden tarkastelujaksolla ei havaittu arvosanainflaatiokehitystä. Poikkeuksena kuitenkin koronavuodet 2020–2021, jolloin etäopetuksen seurauksena arvosanat kohosivat selvästi. Tämä viittaa arviointikäytäntöjen herkkyyteen poikkeusoloissa.

6.3 Tutkimuksen rajoitukset ja luotettavuus

Tutkimuksen aineisto kattoi useita vuosikursseja ja laajan opiskelijajoukon, mikä vahvistaa tulosten yleistettävyyttä Aalto-yliopiston sisällä. Toisaalta tulokset eivät välttämättä siirry sellaisenaan muihin suomalaisiin tai kansainvälisiin korkeakouluihin.

Logististen ja osittaisen verrannollisten vastasuhteiden selitysasteet jäivät pieniksi (Pearson r ja McFadden- R^2 mittarit jäivät aina alle 0.22 ja 0.04), mikä on tyypillistä sosiaalitieteellisille aineistoille, mutta silti osoittaa, että valtaosa arvosanoihin liittyvästä vaihtelusta johtuu muista ei-demograafisista tekijöistä tai arvosanainflaatiosta.

Myös verrannollisten vastasuhteiden oletus kumoutui sekä peruskurssien että kaikkien kurssien järjestyslogistisissa malleissa, mikä viittaa siihen, että muuttujien vaikutukset ovat monimuotoisempia kuin yksinkertainen järjestysmalli olettaa. Lisäksi luokkien sisäinen vaihtelu korostui erityisesti koronavuosina, jolloin malli sovitettiin keskimäärin korkeampia arvosanoja. Aineistossa esiintyi jonkin verran poikkeavia havaintoja (jäännös > 3), mikä viittaa epäyhtenäisiin vaikutuksiin ryhmien sisällä.

Koronavuosien aikana opinnot siirtyivät vaihtelevasti lähi- ja etäopetuksen välillä, eikä aineiston perusteella ole mahdollista päätellä, mitkä suorituksista tehtiin lähiopetuksessa ja mitkä etäopetuksessa.

7 Johtopäätökset

7.1 Keskeiset havainnot

Tulosten perusteella naisten hyväksymisprosentti kursseilla oli korkeampi, mutta parhaimpien arvosanojen osuuksissa ei löytynyt sukupuolten välillä eroja. Syntymäkuukauden merkitys jäi vähäiseksi: joulukuussa syntyneillä havaittiin lievästi parempia arvosanoja, mutta muut kuukaudet olivat käytännössä ennallaan. Suhteellisen iän vaikutus ei tullut esille, mikä tukee käsitystä sen vaimenemisesta iän myötä, toisin kuin esimerkiksi urheilussa, jossa ilmiö säilyy aikuisuuteen saakka (R. H. Barnsley, Thompson ja P. E. Barnsley 1985).

Kansainvälisten opiskelijoiden suorituksissa näkyi suurempi hajonta: heidän joukossaan oli sekä enemmän hylättyjä suorituksia että enemmän huippuosaajia verrattuna kantasuomalaisiin opiskelijoihin. Tämä täydentää aiempaa maahanmuuttajataustaista tutkimusta laajentamalla vertailun kansainväliseen ydinosajajoukkoon.

Arvosanainflaatiota ei havaittu analysoiduilla matematiikan kursseilla 11 vuoden tarkastelussa, mikä on poikkeavaa suhteessa kansainvälisiin trendeihin. Koronavuosina arvosanat kuitenkin nousivat, mikä voi johtua etäopetuksen ja joustavampien arviointikäytäntöjen vaikutuksesta.

7.2 Jatkotutkimusaiheet

Jatkossa tutkimusta olisi hyödyllistä laajentaa koskemaan myös maisteri- ja tohtori-kursseja, jotta saataisiin kattavampi kuva matematiikan suorituksista koko korkeakoulupolun aikana. Lisäksi tarkempi tieto opiskelijoiden iästä suoritusten ajankohtana mahdollistaisi laajemman ekonometristen mallien käytön, kuten syntymäkuukauden ja ikävuosien vaikutusten erottelun (Lee ja Lemieux, Lee ja Lemieux 2010)

Arvosanainflaation kehitystä olisi syytä seurata pidemmällä aikajänteellä ja laajentaa tarkastelu myös muille oppiaineille matematiikan lisäksi, jotta voitaisiin arvioida, onko ilmiö luonteeltaan yleinen vai ainekohtainen ja miten se on kehittynyt monen vuosikymmenen yli.

Myös opiskelijakohtaisten taustatekijöiden, kuten sosioekonomisen aseman, aikaisemman koulumenestyksen ja opiskelutapojen, yhdistäminen rekisteriaineistoon voisi tarjota syvempää ymmärrystä yliopistomenestyksen taustalla vaikuttavista tekijöistä.

Viitteet

- Agresti, A. (2007). *An Introduction to Categorical Data Analysis*. 2. Hoboken, NJ: Wiley. ISBN: 978-0-471-22618-5.
- (2010). *Analysis of Ordinal Categorical Data*. 2. painos. Hoboken, NJ: Wiley. ISBN: 978-0-470-08289-8.
- Barnsley, R. H., A. H. Thompson ja P. E. Barnsley (1985). “The Relative Age Effect in Sport: A Replication and Extension with Ice-Hockey Players”. *Canadian Journal of Applied Sport Sciences* 10.2, s. 87–93.
- Brant, R. (1990). “Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression”. *Biometrics* 46.4, s. 1171–1178. DOI: [10.2307/2532457](https://doi.org/10.2307/2532457).
- Cliffordson, C. ja J. E. Gustafsson (2010). “Effects of Age and Schooling on Performance in Mathematics: A Longitudinal Study of Factors Influencing Achievement in Sweden”. Teoksessa: *4th IEA International Research Conference*. URL: https://www.iea.nl/sites/default/files/2019-04/IRC2010_Cliffordson_Gustafsson.pdf.
- Education Statistics, National Center for (1990). *National Postsecondary Student Aid Study: 1990 Undergraduates*. Viitattu 28.4.2025. URL: <https://nces.ed.gov/datalab/powerstats/14-national-postsecondary-student-aid-study-1990-undergraduates/averages-medians-percents>.
- (2020). *National Postsecondary Student Aid Study: 2020 Undergraduate Students*. Viitattu 28.4.2025. URL: <https://nces.ed.gov/datalab/powerstats/157-national-postsecondary-student-aid-study-2020-undergraduate-students/averages-medians-percents>.
- Else-Quest, N. M., J. S. Hyde ja M. C. Linn (2010). “Cross-National Patterns of Gender Differences in Mathematics: A Meta-Analysis”. *Psychological Bulletin* 136.1, s. 103–127. DOI: [10.1037/a0018053](https://doi.org/10.1037/a0018053).
- Guiso, L. et al. (2008). “Culture, Gender, and Math”. *Science* 320.5880, s. 1164–1165. DOI: [10.1126/science.1154094](https://doi.org/10.1126/science.1154094).
- Harrell, F. E. (2015). *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. 2. painos. Cham: Springer. ISBN: 978-3-319-19424-0. DOI: [10.1007/978-3-319-19425-7](https://doi.org/10.1007/978-3-319-19425-7).
- Hedges, L. V. ja A. Nowell (1995). “Sex Differences in Mental Test Scores, Variability, and Numbers of High-Scoring Individuals”. *Science* 269.5220, s. 41–45. DOI: [10.1126/science.7604277](https://doi.org/10.1126/science.7604277).
- Hosmer, D. W., S. Lemeshow ja R. X. Sturdivant (2013). *Applied Logistic Regression*. 3. painos. Hoboken, NJ: Wiley. ISBN: 978-0-470-58247-3.
- Huusela, M. ja N. Kilpi (2022). “Suhteellisen Iän Yhteys PISA 2018 -Tutkimuksen Lukutaidon ja Matematiikan Arviointialueiden Tuloksiin Suomessa”. Pro gradu -tutkielma. Pro gradu -tutkielma. Tampereen Yliopisto. URL: <https://trepo.tuni.fi/bitstream/handle/10024/139541/HuuselaKilpi.pdf>.
- Hyde, J. S. et al. (2008). “Gender Similarities Characterize Math Performance”. *Science* 321.5888, s. 494–495. DOI: [10.1126/science.1160364](https://doi.org/10.1126/science.1160364).

- Kaila, M. (2017). *The Effects of Relative School Starting Age on Educational Outcomes in Finland*. Tekninen raportti 84. VATT Institute for Economic Research.
- Karvi, Kansallinen Koulutuksen Arviointikeskus (2023). *Korkeakoulujen Arviointikäytännöt ja Oppimistulosten Arviointi*. Tekninen raportti. URL: https://www.karvi.fi/sites/default/files/sites/default/files/documents/KARVI_2223.pdf.
- Lee, D. S. ja T. Lemieux (2010). “Regression Discontinuity Designs in Economics”. *Journal of Economic Literature* 48.2, s. 281–355. DOI: [10.1257/jel.48.2.281](https://doi.org/10.1257/jel.48.2.281).
- Lindberg, S. M. et al. (2010). “New Trends in Gender and Mathematics Performance”. *PMC*.
- Machin, S. ja T. Pekkarinen (2008). “Global Sex Differences in Test Score Variability”. *Science* 322.5906, s. 1331–1332. DOI: [10.1126/science.1162573](https://doi.org/10.1126/science.1162573).
- McFadden, D. (1974). “Conditional Logit Analysis of Qualitative Choice Behavior”. Teoksessa: *Frontiers in Econometrics*. Toim. P. Zarembka. New York: Academic Press. Luku Four, s. 104–142.
- Metsämuuronen, J., toim. (2013). *Perusopetuksen Matematiikan Oppimistulosten Pitkittäisarviointi V. 2005–2012*. Vol. 2013:4. Koulutuksen Seurantaratortit. Helsinki: Opetushallitus / KARVI.
- Metsämuuronen, J., S. Nousiainen ja H. Peltoniemi (joulukuu 2023). *Matematiikkaa COVID-19-Pandemian Varjossa – Matematiikan Oppimistulosten Syventävien Analyysien Julkistamistilaisuus*. Webinaari, Kansallinen Koulutuksen Arviointikeskus (Karvi). URL: <https://www.karvi.fi/sites/default/files/sites/default/files/documents/Webinaari%20141223.pdf>.
- Miihkinen, A. (2013). “Vertaileva Tutkimus Opintomenestystä Selittäivistä Tekijöistä Aalto-Yliopiston Kolmessa Korkeakoulussa”. Pro gradu -tutkielma. Aalto-Yliopisto, Insinööritieteiden Korkeakoulu. URL: <https://urn.fi/URN:NBN:fi:aalto-2020122859932>.
- OECD (2023). “PISA 2022 Results (Volume I): The State of Learning and Equity in Education”, s. 89. URL: https://www.oecd.org/en/publications/pisa-2022-results-volume-i-and-ii-country-notes_ed6fbcc5-en/finland_6991e849-en.html.
- Pinker, S. (2008). *The Blank Slate: The Modern Denial of Human Nature*. Revised. New York: Penguin Books.
- Planting, M. (2022). “An Analysis of the Relationship Between Immigrant Generation and Math Performance”. Pro gradu -tutkielma. Stockholm University.
- Rojstaczer, S. ja C. Healy (2012). “Where A Is Ordinary: The Evolution of American College and University Grading, 1940–2009”. *Teachers College Record* 114.7, s. 1–23.
- Tarkastusvirasto, Valtiontalouden (2015). *Maahanmuuttajaoppilaat ja Perusopetuksen Tuloksellisuus*. Tekninen raportti 3/2015. VTV. URL: <https://www.vtv.fi/app/uploads/2018/06/15103549/maahanmuuttajaoppilaat-perusopetuksen-tuloksellisuus.pdf>.
- Thoren, K., E. Heinig ja M. Brunner (2016). “Relative Age Effects in Mathematics and Reading: Investigating the Generalizability Across Students, Time and Classes”. *Frontiers in Psychology* 7, s. 679. DOI: [10.3389/fpsyg.2016.00679](https://doi.org/10.3389/fpsyg.2016.00679).

- Tilastokeskus (2025). *Elävänä syntyneet kuukausittain, 1984–2007*. Viitattu 20. heinäkuuta 2025. Saatavissa: https://pxweb2.stat.fi/PxWeb/pxweb/fi/StatFin/StatFin__synt/statfin_synt_pxt_12dl.px/table/tableViewLayout1/. Helsinki.
- Ukkola, A. ja J. Metsämuuronen (2023). *Matematiikan ja Äidinkielen Taidot Alkuopetuksen Aikana – Perusopetuksen Oppimistulosten Pitkittäisarviointi 2018–2020*. Julkaisut 1:2023. Verkojulkaisu, saatavilla ResearchGate-palvelusta. Helsinki: Kansallinen Koulutuksen Arviointikeskus (KARVI). ISBN: 978-952-206-754-8. URL: <https://www.researchgate.net/publication/367361616>.
- Vihriälä, V. (2023). “PISA-Tulosten Heikentyminen Mysteeri, Jolle Ei Voi Mitään?” *Vesavihriala.fi*.
- Voyer, D. ja S. D. Voyer (2014). “Gender Differences in Scholastic Achievement: A Meta-Analysis”. *Psychological Bulletin* 140.4, s. 1174–1204. DOI: [10.1037/a0036620](https://doi.org/10.1037/a0036620).
- Ünal, A. (2019). “The Impact of Relative Age Effect on Mathematics Achievement”. *International Education Studies* 12.6, s. 39. URL: https://www.researchgate.net/publication/333470096_The_Impact_of_Relative_Age_Effect_on_Mathematics_Achievement.

A Liite - Logistinen regressio

Mallin hyväksymistodennäköisyydet $\hat{\pi}$ on johdettu yhtenäisellä tavalla molemmissa malleissa, joiden tulokset esitetään taulukoissa 4 ja 10.

Laskennan pohjana toimii mallin vakiotermistä johdettu *perustason vastasuhde*, joka tarkoittaa tilannetta, jossa kaikki selittävät muuttujat on asetettu nolaksi. Perustason vastasuhde määritetään logit-muunnoksen avulla kaavan (1) mukaisesti. Peruskurssien mallin vakio-termi on $\beta_0 = 1.431$, jolloin:

$$\text{vastasuhde}_{\beta_0} = e^{1.431} \approx 4.183$$

Kunkin muuttujan vastasuhde saadaan kertomalla perustason vastasuhde kyseisen muuttujan kertoimella eksponentoituna. Esimerkiksi naisopiskelijoiden osalta, joiden estimaatti on 0.399, laskenta etenee seuraavasti:

$$\text{vastasuhde}_{\text{nainen}} = 4.183 \cdot e^{0.399} \approx 6.234$$

Tämän jälkeen hyväksymistodennäköisyys $\hat{\pi}$ voidaan laskea ratkaisemalla $\hat{\pi}$ kaavasta (2):

$$\text{vastasuhde} = \frac{\hat{\pi}}{1 - \hat{\pi}} \Leftrightarrow \hat{\pi} = \frac{\text{vastasuhde}}{1 + \text{vastasuhde}} \quad (\text{A1})$$

Naisopiskelijoiden osalta hyväksymistodennäköisyys on näin:

$$\hat{\pi}_{\text{nainen}} = \frac{6.234}{1 + 6.234} \approx 0.862$$

Kaikki mallien $\hat{\pi}$ -arvot on laskettu vastaavalla tavalla käyttäen vakio-termin määrittämää perustasoa, ja niitä verrataan perustason hyväksymistodennäköisyyteen:

$$\hat{\pi}_0 = \frac{4.183}{1 + 4.183} \approx 0.807$$

B Liite - Koodit

B.1 Osuustestit syntymäkuukausille (tammi–maaliskuu ja tammi–kesäkuu)

```
# Kirjastot
library(readxl)
library(dplyr)

# Lue aineisto (vaihda polku tarvittaessa)
data <- read_excel("oma_tiedosto.xlsx")

# Suomalaisten opiskelijoiden syntymäkuukaudet
student_months <- data %>%
  filter(Kansalaisuuskatgoria == "Suomi") %>%
  count('Syntyma-kuukausi') %>%
  mutate('Syntyma-kuukausi' = as.integer('Syntyma-kuukausi')) %>%
  arrange('Syntyma-kuukausi')
```

```

# Varmistetaan, että kaikki kuukaudet mukana
all_months <- tibble('Syntyma-kuukausi' = 1:12)
student_months <- left_join(all_months, student_months, by = "Syntyma-kuukausi") %>%
  mutate(n = ifelse(is.na(n), 0, n))

# Väestön jakauma (syntyneet Suomessa 1984-2007)
population_months <- c(
  120801, 115204, 129912, 126343, 125134, 123742,
  127086, 125872, 122103, 117849, 110654, 112932
)

# Tammi-maaliskuu
student_q1 <- sum(student_months$n[1:3])
population_q1 <- sum(population_months[1:3])
total_students <- sum(student_months$n)
total_population <- sum(population_months)

q1_test <- prop.test(
  x = c(student_q1, population_q1),
  n = c(total_students, total_population),
  correct = FALSE
)
cat("Proportion test - tammi-maaliskuu:\n"); print(q1_test)

# Tammi-kesäkuu
student_half1 <- sum(student_months$n[1:6])
population_half1 <- sum(population_months[1:6])

halfyear_test <- prop.test(
  x = c(student_half1, population_half1),
  n = c(total_students, total_population),
  correct = FALSE
)
cat("\nProportion test - tammi-kesäkuu:\n"); print(halfyear_test)

```

B.2 Logistinen regressio

```

# Kirjastot mallinnukseen ja taulukoihin
library(readxl); library(dplyr); library(lmtest); library(pROC)
library(broom); library(stringr); library(purrr); library(tibble)
library(knitr); library(VGAM)

options(max.print = 1e6, width = 200)

# Ladataan aineisto
data <- read_excel("oma_tiedosto.xlsx")

# Suodatetaan peruskurssit
peruskurssit <- data %>%
  filter(grepl("MS-A000|MS-A010|MS-A020|MS-A050", 'Opintojakson koodi'))

# Muodostetaan aineisto
aineisto <- peruskurssit %>%
  mutate(
    Suoritusvuosi = as.numeric(Suoritusvuosi),
    hyväksytty = ifelse(Arvostelu == "Hylätty", 0, 1),
    syntynyt_alkuvuonna = ('Syntyma-kuukausi' %in% 1:3)*1,
    syntynyt_joulukuussa = ('Syntyma-kuukausi' == 12)*1,
    korona_2020_2021 = (Suoritusvuosi %in% c(2020, 2021))*1,
    korona_2022 = (Suoritusvuosi == 2022)*1,
    arvosanatrendi_keskiarvona = Suoritusvuosi - mean(Suoritusvuosi, na.rm = TRUE),
    'Op.oik. organisaatio' = relevel(factor('Op.oik. organisaatio'), ref = "SCI"),
    Kansalaisuuskattegoria = relevel(factor(Kansalaisuuskattegoria), ref = "Suomi"),
    Sukupuoli = relevel(factor(Sukupuoli), ref = "Mies")
  ) %>%
  filter(!is.na(Arvostelu), Arvostelu %in% c("Hylätty", "1", "2", "3", "4", "5"))

```

```

# Apufunktio metriikoiden tulostamiseen
metriikat <- function(model, data, tag = "") {
  cat("\n===== \n")
  cat(">>> TULOKSET:", tag, "\n")
  cat("===== \n")

  # AIC ja McFaddenin R²
  null_mod <- update(model, . ~ 1)
  pseudo_r2 <- 1 - as.numeric(logLik(model)) / as.numeric(logLik(null_mod))
  cat("AIC:", AIC(model), "\n")
  cat("McFadden R²:", round(pseudo_r2, 4), "\n")

  # Pearsonin R²
  ennusteet <- predict(model, type = "response")
  cat("Pearson R²:", round(cor(data$hyväksytyt, ennusteet)^2, 4), "\n")

  # ROC AUC
  roc_obj <- roc(data$hyväksytyt, ennusteet)
  cat("AUC:", round(auc(roc_obj), 4), "\n")

  # Devianssi- ja Pearson-²
  dev_val <- deviance(model); df_dev <- df.residual(model)
  cat("Devianssi ²(", df_dev, ") = ", round(dev_val, 2),
      ", p = ", signif(1 - pchisq(dev_val, df_dev), 3), "\n", sep = "")
  pear <- sum(residuals(model, type = "pearson")^2)
  cat("Pearson ²(", df_dev, ") = ", round(pear, 2),
      ", p = ", signif(1 - pchisq(pear, df_dev), 3), "\n\n", sep = "")

  # drop1-LRT
  cat("--- drop1() LRT ---\n"); print(drop1(model, test = "LRT"))

  # Standardoidut residuaalit ja poikkeamat
  h <- hatvalues(model)
  rP <- residuals(model, type = "pearson") / sqrt(1 - h)
  rD <- residuals(model, type = "deviance") / sqrt(1 - h)
  cat("\nPearson-residuaaleja > 3: ", sum(abs(rP) > 3),
      "\nDevianssiresiduaaleja > 3: ", sum(abs(rD) > 3), "\n", sep = "")
}

# Perusmalli
malli_perus <- glm(hyväksytyt ~ Kansalaisuuskattegoria + 'Op.oik. organisaatio' + Sukupuoli +
  syntynyt_joulukuussa + korona_2020_2021 + korona_2022 + arvosanatrendi_keskiarvona,
  data = aineisto, family = binomial)
metriikat(malli_perus, aineisto, tag = "PERUSMALLI")

# Interaktiomalli
malli_inter <- glm(hyväksytyt ~ Kansalaisuuskattegoria + 'Op.oik. organisaatio'*Sukupuoli +
  syntynyt_joulukuussa + korona_2020_2021 + korona_2022 + arvosanatrendi_keskiarvona +
  'Op.oik. organisaatio':korona_2020_2021 +
  'Op.oik. organisaatio':Sukupuoli:korona_2020_2021 +
  'Op.oik. organisaatio':Sukupuoli:korona_2020_2021:Kansalaisuuskattegoria,
  data = aineisto, family = binomial)
metriikat(malli_inter, aineisto, tag = "INTERAKTIOMALLI")

```

B.3 Verrannolisten vastasuhteiden oletuksen testaus (Brantin testi)

```

# Kirjastot
library(readxl); library(dplyr); library(VGAM)

# Lue aineisto
data <- read_excel("oma_tiedosto.xlsx")

# Esikäsittely: peruskurssit ja luokat

```

```

df <- data %>%
  filter(grepl("MS-A000|MS-A010|MS-A020|MS-A050", 'Opintojakson koodi')) %>%
  mutate(
    Suoritusvuosi = as.numeric(Suoritusvuosi),
    grade_grouped = case_when(
      Arvostelu == "Hylätty" ~ "0",
      Arvostelu %in% c("1","2","3") ~ "1",
      Arvostelu %in% c("4","5") ~ "2"
    ),
    syntynyt_joulukuussa = factor(iffelse('Syntyma-kuukausi' == 12, 1, 0)),
    korona_2020_2021 = factor(iffelse(Suoritusvuosi %in% c(2020,2021), 1, 0)),
    korona_2022 = factor(iffelse(Suoritusvuosi == 2022, 1, 0)),
    arvosanatrendi_keskiarvona = Suoritusvuosi - mean(Suoritusvuosi, na.rm=TRUE),
    Organisaatio = factor('Op.oik. organisaatio',
      levels=c("SCI", setdiff(unique('Op.oik. organisaatio'), "SCI"))),
    Kansalaisuus = factor(Kansalaisuuskattegoria,
      levels=c("Suomi", setdiff(unique(Kansalaisuuskattegoria), "Suomi"))),
    Sukupuoli = factor(Sukupuoli,
      levels=c("Mies", setdiff(unique(Sukupuoli), "Mies")))
  ) %>%
  filter(!is.na(grade_grouped)) %>%
  mutate(grade_grouped = ordered(grade_grouped, levels=c("0","1","2")))

# PO- ja NPO-mallit
po_model <- vglm(grade_grouped ~ Sukupuoli + Kansalaisuus + Organisaatio +
  syntynyt_joulukuussa + korona_2020_2021 + korona_2022 +
  arvosanatrendi_keskiarvona,
  family = cumulative(parallel=TRUE, link="logit"), data = df)
npo_model <- vglm(grade_grouped ~ Sukupuoli + Kansalaisuus + Organisaatio +
  syntynyt_joulukuussa + korona_2020_2021 + korona_2022 +
  arvosanatrendi_keskiarvona,
  family = cumulative(parallel=FALSE, link="logit"), data = df)

# LRT PO vs NPO
lrt_stat <- 2 * (as.numeric(logLik(npo_model)) - as.numeric(logLik(po_model)))
df_diff <- attr(logLik(npo_model),"df") - attr(logLik(po_model),"df")
p_val <- pchisq(lrt_stat, df=df_diff, lower.tail=FALSE)
cat("Likelihood Ratio Test (PO-oletus)\n")
cat("Chi² = ", round(lrt_stat,3), ", df = ", df_diff, ", p-arvo = ", format.pval(p_val, 4), "\n")

# McFadden R² (valinnainen)
R2_po <- 1 - as.numeric(logLik(po_model)) / as.numeric(logLik(update(po_model, .~1)))
R2_npo <- 1 - as.numeric(logLik(npo_model)) / as.numeric(logLik(update(npo_model, .~1)))
cat("McFadden R² (PO): ", round(R2_po,4), "\n")
cat("McFadden R² (NPO): ", round(R2_npo,4), "\n")

```

B.4 Osittain verrannollisten vastasuhteiden mallit

```

# Kirjastot
library(readxl); library(dplyr); library(VGAM); library(pROC)
library(tibble); library(stringr); library(forcats)

# Lue aineisto
data <- read_excel("oma_tiedosto.xlsx")

# Peruskurssit ja muuttujat
aineisto <- data %>%
  filter(
    str_detect('Opintojakson koodi', "^(MS-A000|MS-A010|MS-A020|MS-A050)"),
    !'Op.oik. organisaatio' %in% c("NULL","ARTS"),
    Arvostelu %in% c("Hylätty","1","2","3","4","5")
  ) %>%
  mutate(
    Suoritusvuosi = as.numeric(Suoritusvuosi),
    Arvosana_luokittelu = ordered(case_when(
      Arvostelu == "Hylätty" ~ "Hylätty",

```

```

    Arvostelu %in% c("1","2","3") ~ "1-3",
    TRUE ~ "4-5"),
    levels = c("Hylätty","1-3","4-5")),
    accepted = (Arvostelu %in% c("4","5"))*1,
    Syntynyt_alkuvuonna = ('Syntyma-kuukausi' %in% 1:3)*1,
    Syntynyt_joulukuussa = ('Syntyma-kuukausi' == 12)*1,
    korona_2020_2021 = (Suoritusvuosi %in% c(2020,2021))*1,
    korona_2022 = (Suoritusvuosi == 2022)*1,
    arvosanatrendi_keskiarvona = Suoritusvuosi - mean(Suoritusvuosi, na.rm=TRUE),
    'Op.oik. organisaatio' = fct_relevel(factor('Op.oik. organisaatio'), "SCI"),
    Kansalaisuus = fct_relevel(factor(Kansalaisuuskattegoria), "Suomi"),
    Sukupuoli = factor(Sukupuoli, levels = c("Mies","Nainen"))
  )

# Mallikaavat
k_perus <- Arvosana_luokittelu ~ Sukupuoli + Kansalaisuus +
  Syntynyt_alkuvuonna + Syntynyt_joulukuussa +
  korona_2020_2021 + korona_2022 + arvosanatrendi_keskiarvona
k_korona <- update(k_perus, . ~ . + Sukupuoli:korona_2020_2021)
k_koulu <- update(k_perus, . ~ . + 'Op.oik. organisaatio':Sukupuoli)
k_full <- update(k_korona, . ~ . + 'Op.oik. organisaatio':Sukupuoli)

# PPO-perheet
fam_perus <- cumulative("logitlink",
  parallel = FALSE ~ Sukupuoli + Kansalaisuus + korona_2020_2021 + korona_2022 + arvosanatrendi_keskiarvona)
fam_korona <- update(fam_perus, parallel = FALSE ~ . + Sukupuoli:korona_2020_2021)
fam_koulu <- update(fam_perus, parallel = FALSE ~ . + 'Op.oik. organisaatio':Sukupuoli)
fam_full <- update(fam_korona, parallel = FALSE ~ . + 'Op.oik. organisaatio':Sukupuoli)

# Sovitus
m_full <- vglm(k_full, family = fam_full, data = aineisto)

# Referenssi-olio
ref <- tibble(
  Sukupuoli = factor("Mies", levels = levels(aineisto$Sukupuoli)),
  Kansalaisuus = factor("Suomi", levels = levels(aineisto$Kansalaisuus)),
  'Op.oik. organisaatio' = factor("SCI", levels = levels(aineisto$'Op.oik. organisaatio')),
  Syntynyt_alkuvuonna = 0, Syntynyt_joulukuussa = 0,
  korona_2020_2021 = 0, korona_2022 = 0, arvosanatrendi_keskiarvona = 0
)

# Ennuste-funktio
predict_scenario <- function(new_row, label, digits = 4) {
  nd <- data.frame(new_row, check.names = FALSE)
  nd$Sukupuoli <- factor(nd$Sukupuoli, levels = levels(aineisto$Sukupuoli))
  nd$Kansalaisuus <- factor(nd$Kansalaisuus, levels = levels(aineisto$Kansalaisuus))
  nd$'Op.oik. organisaatio' <- factor(nd$'Op.oik. organisaatio',
    levels = levels(aineisto$'Op.oik. organisaatio'))
  pr <- predict(m_full, newdata = nd, type = "response")
  colnames(pr) <- c("P(Hylätty)", "P(1-3)", "P(4-5)")
  cat("\n", label, "\n", sep = ""); print(round(pr, digits))
}

# Skenaariot
predict_scenario(ref, "Referenssihenkilö")
r1 <- ref; r1$Sukupuoli <- factor("Nainen", levels = levels(aineisto$Sukupuoli))
predict_scenario(r1, "Nainen")
r2 <- ref; r2$Kansalaisuus <- factor("Muu", levels = levels(aineisto$Kansalaisuus))
predict_scenario(r2, "Ei-Suomi")
r3 <- ref; r3$Syntynyt_joulukuussa <- 1
predict_scenario(r3, "Syntynyt joulukuussa")
r4 <- ref; r4$korona_2020_2021 <- 1
predict_scenario(r4, "Suoritus korona 2020-2021")
r5 <- ref; r5$korona_2022 <- 1
predict_scenario(r5, "Suoritus 2022")
r6 <- ref; r6$'Op.oik. organisaatio' <- factor("ELEC",
  levels = levels(aineisto$'Op.oik. organisaatio'))
predict_scenario(r6, "Koulu ELEC")

```