Tampereen yliopisto

Joonas Järnvall

OPINTOMENESTYS KAHDEN SUORITUSTAPAVAIHTOEHDON YLIOPISTOMATEMATIIKAN OPINTOJAKSOLLA

Diplomityö
Tekniikan ja luonnontieteiden tiedekunta
Tarkastajat: Yliopisto-opettaja Jani Hirvonen & Yliopistonlehtori Simo Ali-Löytty
Helmikuu 2024

TIIVISTELMÄ

Joonas Järnvall: Opintomenestys kahden suoritustapavaihtoehdon yliopistomatematiikan opintojaksolla
Diplomityö
Tampereen yliopisto
Teknis-luonnontieteellinen koulutusohjelma
Helmikuu 2024

Tutkimuksessa tarkastellaan Tampereen yliopiston opiskelijoiden opintomenestystä opintojaksolla Johdatus todennäköisyyslaskentaan ja tilastolliseen päättelyyn. Tutkimuksen tarkoituksena on tutkia eroja opintomenestyksessä eri opiskelijaryhmien välillä.

Tutkittavan opintojakson suorittamiseen oli annettu opiskelijoille kaksi vaihtoehtoa. Nämä vaihtoehdot olivat tenttipainotteinen videoluento-opetus ja käänteisen opetuksen ja oppimisen menetelmiä käyttävä osallistava opetus. Näiden kahden opiskelijaryhmän opintomenestyksen vertailu on tutkimuksessa keskiössä. Lisäksi tutkimuksessa vertaillaan opintomenestystä tarkasteltavalta opintojaksolta eri sukupuolten välillä sekä eri vuosina opiskelut alottaneiden opiskelijoiden välillä.

Tutkimusaineisto on kerätty vuonna 2022 järjestetyltä Johdatus todennäköisyyslaskentaan ja tilastolliseen päättelyyn -opintojaksolta. Tutkimusluvan antaneilta opiskelijoilta kerättiin vastauksia kysymyksiin opinnoista, sukupuolesta ja opintojen aloitusvuodesta. Lisäksi näiltä opiskelijoilta kerättiin tarkasteltavalta kurssilta arvosanat opintomenestyksen vertailua varten.

Tutkimuksen tuloksissa saatiin selville, että opiskelijat jotka suorittivat kurssin käänteisen oppimisen ja opetuksen menetelmin saivat kurssista parempia arvosanoja kuin tenttipainotteisesti kurssin suorittaneet opiskelijat. Myös naisten ja miesten välillä opintomenestyksessä oli selkeä ero naisten eduksi. Pidempään yliopistossa opiskelleet opiskelijat saivat heikompia tuloksia kuin ensimmäisen vuoden opiskelijat.

Käänteisen opetuksen menetelmillä saatua parempaa opintomenestystä voidaan pitää oletettuna tuloksena. Erilaisten opetusmenetelmien käyttäminen ja säännöllisten viikottaisten harjoitusten käyttö osana opetusta aktivoi opiskelijoita opiskelemaan. Tällaisen aktiivisen oppimisen on nähty aiemminkin antavan parempia oppimistuloksia.

Avainsanat: käänteinen oppiminen, käänteinen opetus, aktiiviinen oppiminen, opintomenestys

ABSTRACT

Joonas Järnvall: Academic success on a university mathematics course with two completion method options
Master of Science Thesis
Tampere University
Master's Degree Programme in Science and Engineering
February 2024

The study examines the academic success of students at the University of Tampere on the course Introduction to probability and statistical inference. The purpose of the study is to investigate the differences in academic success between groups of students.

The students were given two options for passing the course. These alternatives were exam-oriented video lecture teaching and participatory teaching using flipped teaching and flipped learning methods. The comparison of the academic success between these two groups of students is the focus of the study. In addition, the study compares the academic success on the course between different genders and between students who started their studies in different years.

The research material has been collected from the Introduction to Probability and Statistical Inference course which took place in 2022. Answers to questions about studies, gender and the year their studies started were collected from the students who gave consent to participate in the research. In addition, grades from the course under review were collected from these students for the purpose of comparing their success in this course.

The results of the study revealed that students who completed the course using flipped learning and teaching methods received better grades than students who completed the course with a focus on exams. There was also a clear difference in academic success between female and male participants in favor of female ones. Students who studied longer at the university got worse results than first-year students.

The better academic success obtained with the methods of flipped teaching can be considered as an assumed result. Using different teaching methods and using regular weekly exercises as part of teaching activates students to study and learn. This kind of active learning has been seen to give better learning results in the past.

Keywords: flipped learning, flipped teaching, active learning, academic success

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

SISÄLLYSLUETTELO

LYHENTEET JA MERKINNÄT

$E(x)$	Odotusarvo satunnaismuuttujalle x
H_0	Nollahypoteesi eli oletus, etti ryhmien välillä ole eroa
H_1	Vaihtoehtoinen hypoteesi, että ryhmien välillä on eroa
$N_n(\mu, \Sigma)$	Satunnaismuuttuja on multinormaalijakautunut
P-arvo	Kuvaa todennäköisyyttä sille, että jokin tutkittava tapahtuma olisi tapahtunut sattumalta
S_n	Otoskeskiarvojen keskivirheet
$V(x)$	Kovarianssimatriisi satunnaisvektorille x
X_k	Selittävä muuttuja lineaarisessa regressioanalyysissä
Y	Selitettävä muuttuja lineaarisessa regressioanalyysissä
$\bar{X}$	Otoskeskiarvo t-testissä
$\bar{Y}$	Otoskeskiarvo kahden otoksen t-testissä
α	Luottamusväliin kuulumattomien realisoituvien arvojen osuus
χ^2	Khiin neliö -jakauma
ϵ	Satunnaistermi lineaarisessa regressiomallissa
γ	Funktio, joka voidaan tulkita kertoman yleistyksenä
$\hat{X}$	Datamatriisi, joka sisältää vektorit ${x_i}^T$
μ_0	Odotusarvo t-testin yhteydessä
σ^2	Varianssi
a	Regressiokertoimien b pienimmän neliösumman estimaatti
b_k	Selitettävää muuttujaa vastaava regressiokerroin lineaarisessa regressioanalyysissä
c_{ii}	Matriisin ${\hat{X}^T X}^{-1}$ lävistäjäalkio
e	Residuaalivektori regressiomallissa
m	Otoskoko toiselle otokselle t-testissä
n	Otoskoko tutkimusaineistolle
s	Keskihajonta

t	Satunnaismuuttuja t-jakauman yhteydessä
$t(n-1)$	t-jakauma $n-1$ vapausastein
v	t-jakauman vapausasteet
v_u	Varianssin komponentti
x_i	Vektori x_i on pystyvektori, joka sisältää selittävät muuttujat X_k
${s_u}^2$	Varianssin harhaton estimaattori
LaTeX	Ladontajärjestelmä tieteelliseen kirjoittamiseen
TUNI	Tampereen korkeakouluyhteisö (engl. Tampere Universities)
URL	verkkosivun osoite (engl. Uniform Resource Locator)

1. JOHDANTO

Tässä diplomityössä tutkitaan erilaisten opiskelutapojen vaikutusta menestykseen matematiikan korkeakouluopinnoissa. Tutkittavana ryhmänä toimii Tampereen yliopiston opiskoilijoita, jotka suorittivat ensimmäiselle opintovuodelle suunnattuja matematiikan opintoja. Tarkemmin tutkittava ryhmä koostuu opiskelijoista, jotka saivat suorituksen Tampereen yliopiston matematiikan opintojaksolta Johdatus todennäköisyyslaskentaan ja tilastolliseen päättelyyn. Tutkimuksen tarkoituksena on selvittää, miten opintojakson tulokset eroavat kahden opintojakson toteutuksella olleen osallistumistavan välillä. Perinteisemmällä tenttipainotteisella osallistumistavalla kurssin suorittaneiden opiskelijoiden tuloksia verrataan opiskelijoihin, jotka suorittavat kurssin käänteisen oppimisen ja käänteisen opetuksen keinoja sisältävällä osallistavammalla osallistumistavalla.

Työssä vertaillaan kurssin opiskelijoiden opintomenetystä myös muissa ryhmissa, sekä sukupuolten että eri vuosina opinnot aloittaneiden opiskelijoiden välillä. Opintomenestyksen vertailua tehdään erilaisilla tilastomatematiikan keinoilla, kuten käyttämällä regressioanalyysia analysoitavien tutkimuskyselyn arvosanatietojen yhteydessä. Tämän työn toisessa luvussa esitellään tutkimuskysymykset, joissa kerrotaan tarkemmin tutkimuksen tavoitteista. Työn toisessa luvussa esitellään myös tutkittavan kurssin opetusmenetelmiä ja suoritustapoja.

Kolmas luku käsittelee tutkimuksen taustatietoja. Siinä esitellään erilaisia oppimis- sekä arviointimenetelmiä, joita kurssilla oli käytössä. Neljännessä luvussa tutustutaan matemaattiseen teoriaan työn taustalla. Luvussa esitellään työssä käytettäviä tilastotieteen menetelmiä, kuten regressioanalyysia. Viidennessä luvussa käsitellään tutkimuksen toteutusta. Tutkimusainoistona työssä toimii opintojakson yhteydessä tutkimusluvan antaneiden opiskelijoiden täyttämä kysely, jossa kysyttiin muun muassa heidän sukupuoltaan ja yliopisto-opintojen aloitusvuottaan. Tutkittavien vastaukset on pseudonymisoitu ennen tutkimuksen tekemistä. Kuudennessa luvussa esitellään työstä saatuja tuloksia. Luku pyrkii vastaamaan tutkimuskysymyksiin työssä tehdyn tilastollisen testauksen ja tilastollisten tunnuslukujen avulla.

Yhteenvedossa pohditaan saatujen tulosten merkitystä ja tutkimuskysymysten yhteydessä asetettujen oletusten toteutumista. Luvussa pohditaan myös tutkimuksen luotettavuutta ja siihen vaikuttavia tekijöitä.

2. TUTKIMUKSEN LÄHTÖKOHDAT

Tässä luvussa esitellään tutkimuksen taustat ja kerrotaan miksi tutkimusta tehdään. Ensimmäisessä alaluvussa esitellään tarkasteltavan kurssin toteutus ja suoritustavat. Luvun toisessa alaluvussa esitellään tutkimuskysymykset ja kerrotaan lyhyesti, miten niihin aiotaan tässä tutkimuksessa saada vastauksia.

2.1 Johdatus todennäköisyyslaskentaan ja tilastolliseen päättelyyn

Tässä diplomityössä käsitellään Tampereen yliopiston matematiikan opintojaksoa Johdatus todennäköisyyslaskentaan ja tilastolliseen päättelyyn. Erityisesti tutkitaan opintojakson toteutuksen eri osallistumistapojen yhteyttä opintomenestykseen.

Tutkimuksessa tarkasteltava kurssi toteutettiin käänteisellä opetusmenetelmällä ja pääasiassa verkkototeutuksena. Tältä opintojaksolta jokainen opiskelija pystyi saamaan suorituksen kahdella eri osallistumistavalla. Eri osallistumistapoja nimitetään jatkossa osallistuvaksi suoritustavaksi ja tenttipainotteiseksi suoritustavaksi. Osallistuva suoritustapa painotti kurssin aikana tehtyjen verkko- ja harjoitustehtävien osuutta kurssin arvosanassa. Tässä osallistumistavassa tenttien vaikutus arvosanaan oli reilu kolmasosa, ja kurssin aikasen työskentelyn osuus vajaa kaksi kolmasosaa. Tenttipainotteiseen suoritustapaan sisältyi myös harjoitustehtäviä, mutta arvosana saatiin suoraan lopputentin pisteistä. Opiskelijan ei tarvinnut valita suoraan näiden kahden osallistumistavan välillä, sillä molemmista laskettiin hänelle arvosana, joista parempi tuli hänen lopulliseksi arvosanakseen. Kurssin aiheet oli jaettu viikottaisiin kokonaisuuksiin ja koostuvat todennäköisyyslaskennan käsitteistä kuten satunnaismuuttuja, todennäköisyysjakaumat ja odotusarvo. [12]

Kummallakin osallistumistavalla opiskelija keräsi pisteitä, joista kurssin arvosana muodostui. Osallistuvassa suoritustavassa opiskelijat osallistuivat viikoittaisiin harjoitusryhmiin, joissa käsiteltiin viikon aiheita ja tehtiin niihin liittyviä tehtäviä. Harjoitusryhmissä työskenneltiin pienryhmissä ja jokaisella pienryhmällä oli mahdollisuus tehdä harjoitustyö, josta pystyi ansaitsemaan lisäpisteitä. Verkkotehtävät olivat yksi osa kurssin viikoittaisia suorituksia, kuten myös viikoittain palautettava tehtäväpaketti, joka vertais- ja itsearvioitiin palautusta seuraavalla viikolla. Näiden tehtävien lisäksi kurssilla oli elektronisen kurssimateriaalin yhteydessä itsenäisen opiskelun käsitteenmuodostusta edistäviä tehtäviä, joita pystyi tekemään yliopiston Plus-järjestelmässä. Näistä tehtävistä oli mahdollista halutessaan ansaita ylimääräisiä pisteitä. Näistä osasuorituksista saatavien pisteiden kertymistapa osallistuvalla suoritustavalla on esiteltynä taulukossa 2.1.[12]

Suorite	Pisteet
Palautettavat tehtävät ja arvioinnit	6x45 = 270
Pienryhmätyöskentely ja harjoitustyö	7x10 + 80 = 150
Verkkotehtävät	6x35 = 210
Plussamonisteen tehtävät	40
Välitentti	120
Lopputentti	250
Yhteensä	1040

Taulukko 2.1. *Kurssilta kerättävät pisteet osallistuvaan suoritustapaan*

Tenttipainotteisella suoritustavalla kurssin pisteet tulivat suoraan lopputentin 250 mahdollisesta pisteestä kerrottuna neljällä.

Vähimmäisvaatimukset kurssin läpäisylle kummallakin osallistumistavalla oli 500 pistettä eri suoritteista. Lisäksi opiskelijan tuli saada yhteensä vähintään 100 pistettä kurssin kahdesta Exam-järjestelmässä suoritettavasta tentistä ja vähintään 100 pistettä muista kurssin suorituksista. Kurssin arvosanat saatiin suoraan kurssilla kerätyistä pisteistä, kunhan vähimmäisvaatimukset kurssin suorituksista täyttyivät. Kurssin arvosanarajat on esitelty taulukossa 2.2. [12]

Pisteet	500	600	700	800	900
Arvosana	1	2	3	4	5

Taulukko 2.2. *Opintojakson arvosanarajat*

Lopputentin vaikutus tenttipainotteisella tavalla on moninkertainen osallistuvaan suoritustapaan verrattuna. Tenttipainotteisella suoritustavalla lopputentti määrää lähes yksin koko kurssin arvosanan kun taas osallistuvalla suoritustavalla tentin vaikutus on arvosanaan on vain noin neljännes kaikista opintojakson suorituksista. [12]

2.2 Tutkimuskysymykset

Tutkimuksessa käsitellään Johdatus todennäköisyyslaskentaan ja tilastolliseen päättelyyn -opintojakson (JTT-kurssin) opintomenestystä eri osallistumistapojen välillä. Pohditaan siis sopiiko yliopiston matematiikan käänteisen opetustavan kursseille paremmin pienryhmätyöskentely vai itsenäinen suorittaminen, kun kysymystä pohditaan opiskelijoiden opintomenestyksen kannalta. Tutkimuskysymyksinä työssä ovat seuraavat:

1. Muodostuuko JTT-kurssin kahden eri osallistumistavan välillä eroa kurssin opintomenestyksessä?
2. Vaikuttaako opiskelijan sukupuoli tai ikä opintomenestykseen JTT-kurssilla?
3. Saavatko pidempään yliopistossa opiskelleet erilaisia tuloksia JTT-kurssilta kuin ensimmäisen vuoden opiskelijat?

Kurssin arvosanojen avulla saadaan vastuksia ensimmäiseen tutkimuskysymykseen. Muihin tutkimuskysymyksiin etsitään vastuksia analysoimalla kurssilta kerätyn kyselyn tuloksia, kun ne yhdistetään arvosanatietoihin. Kyselyn yhteydessä kurssin opiskelijoilta kysyttiin muun muassa heidän sukupuoltaan ja ikäänsä. Tutkimuksessa vertaillaan ainoastaan opiskelijoita, joilla on kurssista vähintään yksi tenttisuoritus ja jotka ovat antaneet tutkimusluvan.

Ennakko-oletus ensimmäiseen tutkimuskysymykseen on, että kurssin pelkällä tentillä suorittavat opiskelijat eivät saa yhtä korkeita arvosanoja kurssista, kuin ne jotka osallistuvat säännöllisesti harjoituksiin. Oletus perustuu siihen, että jatkuvalla oppimisella, jossa opiskelu jaetaan tasaisesti pitkin kurssia harjoitustehtävien ja ryhmätöiden avulla, on aikaisemmissa tutkimuksissa havaittu antavan parempia oppimistuloksia.

Toisen tutkimuskysymyksen kohdalla ennakko-oletuksena on, että naiset saavat miehiä parempi arvosanoja. Perusteluna tähän oletukseen voidaan pitää esimerkiksi tyttöjen parempia Pisa-tuloksia [21]. Kolmannen tutkimuskysymyksen kohdalla voitaisiin olettaa, että pidempään yliopistossa opiskelleet saavat parempia tuloksia kurssista, sillä heillä on enemmän kokemusta matematiikan opiskelusta yliopistossa. On kuitenkin todennäköisempää, että pidempään opiskelleet opiskelijat saavat huonompia arvosanoja, sillä he mahdollisesti uusivat kurssia ja käyvät sitä siten toista kertaa. Kurssia uusivilta opiskelijoilta ei voida olettaa huippuarvosanoja toisellakaan suorituskerralla.

3. TUTKIMUKSEN TAUSTAT

Tässä luvussa perehdytään Johdatus todennäköisyyslaskentaan ja tilastolliseen päättelyyn -kurssin taustalla olevaan teoriaan opetuksesta. Tämä luku esittelee kurssilla käytettyjä opetus- ja arviointimenetelmiä. Opetusmenetelmissä tutustutaan käänteiseen oppimisen ja käänteiseen opetukseen, josta käytetään myös nimeä flippaus. Flippauksen tutkiminen ja vaikutus opintomenestykseen yliopistossa, on yksi syistä miksi tutkimusta alettiin tekemään. Aktiivisen oppimisen malli ja sen opetusmenetelmät ovat myös kurssilla olennaisessa osassa erityisesti kurssin osallistavassa suoritustavassa. Lisäksi luvussa tutustutaan opintojakson arviointimenetelmien kuten vertais- ja itsearvioinnin taustalla olevaan teoriaan.

3.1 Käänteinen oppiminen ja käänteinen opetus

Käänteinen opetus eli flippaus on viime vuosina paljon suosiota saanut opetusmuoto. Käänteinen opetus (Flipped classroom) eroaa käänteistä oppimisesta (Flipped learning) siten, että käänteinen oppiminen on enemmänkin ideologia, jossa omaehtoinen oppiminen voi tapahtua. Käänteisellä opetuksella tarkoitetaan opetusmallia, jossa opetuksen painotus siirtyy pois opettajakeskeisestä opetusmallista. Opiskelijoita kannustetaan tällöin itsenäiseen oppimiseen, mutta samalla annetaan myös keinoja toteuttaa sitä. Käänteiselle opetukselle tyypillisiä opetusmenetelmia ovat opetusvideot, erilaiset pohdintatehtävät ja ryhmässä opitun asian soveltaminen. [20] Tutkittavalla kurssilla käänteinen opetus näkyi usealla tavalla. Kurssilla uuden aiheen opiskeluun käytettiin opetusvideoita, joihin opiskelijat tutustuvat ennen yhteisiä tapaamisia. Yhteisissä pienryhmätapaamisessa tehtiin tehtäviä opiskelluista aiheista ja keskusteltiin niiden sisällöstä.[12] Yleisesti flippauksella nähdään olevan paljon positiivisia vaikutuksia opiskelijoiden oppimiseen aktivoimisessa ja opiskelun monipuolistamisessa. Siten flippaukseen liityvät tutkimukset ovatkin ajankohtaisia ja kiinnostavia tutkimusaiheita.

Aiempaa tutkimusta käänteisen opetuksen vaikutuksesta matematiikan oppimistuloksiin ei ole kovin paljon. Eräässä Uuden-Seelannin yliopistossa tehdyssä tutkimuksessa [9] tutkittiin käänteisen ja perinteisen opetuksen eroavaisuutta samalla matematiikan laskennon kurssilla. Tutkimuksessa tarkasteltiin 690 opiskelijan suuruista otosta, joista osa suoritti kurssia perinteisemmillä menetelmillä ja osalla opetusmuotona oli erilaisia käänteisen opetuksen menetelmiä. Tuloksena tutkimuksessa todettiin käänteiseen opetukseen osallistuvien opiskelijoiden saaneen noin 8% parempia tuloksia. Erityisesti tutkimuksessa todettiin käänteisestä opetuksesta olleen hyötyä opiskelijoille, joilla oli vahvat matemaattiset kyvyt, mutta vähän kokemusta laskemisesta. Heidän kohdallaan ero koetuloksissa oli noin 10% parempi kuin perinteiseen opetukseen osallistuneilla opiskelijoilla.

Toisessa tutkimuksessa flippauksen vaikutusta matematiikan oppimiseen tutkittiin Ruotsin Kuninkaallisessa teknillisessä korkeakoulussa. [3] Tutkimuksessa vertailtiin opiskelijoiden oppimistuloksia käänteisen opetuksen ryhmän ja perinteisen opetuksen ryhmän välillä. Tuloksia kerättiin kurssin alussa ja lopussa pidettävien kokeiden tuloksista. Tutkimukseen osallistui 639 oppilasta, joista 226 osallistui käänteiseen opetukseen ja 413 suoritti kurssin perinteisen luento-opetuksen avulla. Käänteinen opetus sisälsi ennen luentoa katsottavia opetusvideoita, interaktiivisia luentoja ja yhteisiä ongelmanratkaisu tilaisuuksia. Käänteisen ryhmän normalisoituneet tulokset loppukokeesta olivat 13% parempia kuin vertailtavalla ryhmällä. Lisäksi tuloksia vertailtiin edellisen vuoden loppukokeen tuloksiin ja huomattiin, että käänteisellä opetuksella kurssin hylättyjen arvosanojen määrä oli 56% pienempi verrattuna edelliseen vuoteen.

3.2 Aktiivinen oppiminen

Tutkittavalla kurssilla eräs tärkeä opetusmenetelmä oli aktiivinen oppiminen. Aktiivisella oppimisella tarkoitetaan sellaisia opiskelumenetelmiä, joissa opiskelija kertaa tai käyttää apunaan jo oppimaansa tietoa. Sen vastakohtana voidaan pitää passiivista oppimista, missä opiskelija ei pyri soveltamaan oppimaansa tietoa. Passiivisessa oppimisessa opiskelija saattaa esimerkiksi lukea kirjan kappaletta yhä uudelleen läpi. Aktiivinen oppimismenetelmä vastaavaan tilanteeseen voisi olla esimerkiksi tehdä aiheestä tehtäviä tai esitelmä. Kurssilla viikottaiset palautettavat tehtävät ja verkkotehtävät voidaan laskea aktiiviseksi oppimiseksi. Myös pienryhmissä tehtävät harjoitustyöt kuuluvat aktiivisen oppimisen alle, sillä niiden tekemisessä käytettiin kurssilla jo opittua tietoa hyväksi. [12]

Aktiivisen oppimisen opiskelutekniikat voidaan jakaa valmistautumiseen, osallistumiseen, käyttöön ja arviointiin. Aktiivinen valmistautuminen tarkoittaa tilannetta, jossa opiskelija ennakoiden käyttää aikaa uuteen aiheeseen tutustumiseen. [2] Kurssilla materiaali on valmiina verkossa, joten tämän tekniikan käyttö on kurssilla mahdollistettu [12]. Osallistumisella tarkoitetaan tiedon keräämistä aiheesta, jota voi olla esimerkiksi muistiinpanojen tekeminen luennolla. Opittua tietoa käyttämällä pyritään vahvistamaan muistijälkeä ja ymmärtämään syvemmin opittua asiaa. Tehtävien tai esitelmien tekeminen aiheesta soveltuu hyvin tämän tekniikan käyttöön. Arvioinnilla tarkoitetaan sellaisia oppimistekniikoita, joissa reflektoidaan juuri opittua asiaa ja pohditaan tarkemmin mitä se tarkoittaa. Juuri opitun asian muotoileminen omiksi sanoiksi on hyvä esimerkki tästä opiskelutekniikasta. [2]

Aktiivisen oppimisen tarkoituksena on kehittää sellaisia taitoja, joita ei passiivisella oppimisella saada yhtä hyvin kehitettyä. Aktiivisessa oppimisessa korostuvat usein metakognitiiviset taidot kuten ongelmanratkaisukyky, itseohjautuvuus ja yhteistyökyky. Matematiikassa ongelmanratkaisukykyä voidaan pitää hyvin tärkeänä taitona. Monet matemaattiset tehtävät perustuvat juuri ongelmatilanteisiin, joihin tarvitaan ratkaisu matematiikan avulla. [10]

Aktiiviseen oppimiseen liitetään usein oppimisympäristön vaikutukset oppimiseen [1]. Erilaiset työtilat voivat auttaa monen opiskelijan kohdalla opiskeluun keskittymistä. Moderneissa opetustiloissa onkin usein erilaisia pöytä- ja tuoliratkaisuita kuin perinteisessä luokkahuoneessa, jossa kaikki oppilaat istuvat omilla tuoleillaan pulpettien ääressä.

Pitkällä aikavälillä aktiivisen oppimisen on nähty antavan parempia oppimistuloksia kuin passiivisen oppimisen. Roedigerin ja Karpicken [17] tutkimuksessa verrattiin passiivisen ja aktiivisen oppimisen eroja tekemällä muistitestejä aiheen opiskelun jälkeen. Passiivisen oppimisen ryhmä luki tekstiä useaan kertaan läpi, kun taas aktiivisen oppimisen ryhmä pyrki kirjoittamaan tekstistä muistamaansa ylös. Muistitestissä lyhyellä aikavälillä passiivisten oppijoiden tulos oli parempi, mutta tulos kääntyi aktiivisen oppimisen eduksi, kun muistitestiä toistettiin kahden päivän sekä viikon päästä opiskelusta. Pitkäaikaisessa opiskelussa, kuten yliopiston matematiikan opiskelussa, voidaan aktiivista oppimista pitää parempana opiskelumenetelmänä, koska siitä jää pidempi muistijälki opiskelijoille, josta on hyötyä kun uutta tietoa rakennetaan myöhemmillä kursseilla aiemmin opitun tiedon päälle.

3.3 Arviointimenetelmät

Erilaiset arviontitavat voivat vaikuttaa oppimiseen ja tässä alaluvussa esitellään arviointimenetelmien eroavaisuuksia yleisesti. Luvassa käydään myös läpi kurssilla käytössä olleita arviointimenetelmiä ja sitä miten niitä hyödynnettiin kurssilla. Kurssin arviointi osallistumispainotteisella suoritustavalla koostui useasta eri tekijästä. Yhtenä osana kurssin suoritusta ja arviointia toimivat opiskelijoiden vertaisarvioinnit viikoittaisista tehtävistä. Kurssilla jokainen opiskelija teki sekä vertais- että itsearvioinnin edellisellä viikolla palautetuista tehtävistä. Tämän arvioinnin pisteet olivat osa kurssilla kerätyistä pisteistä. Lisäksi kurssin verkkotehtävät antavat usein suoraan pisteet palautuksen kohdalla ja oikeat vastaukset tulivat opiskelijoille näkyviin palautuksen jälkeisellä viikolla, jolloin niiden vastauksista pystyi saamaan palautetta osaamisestaan. Toisaalta tenttipainotteisella suoritustapaa käyttävät oppilaat saivat arvioinnin ainoastaan yhdestä tenttisuorituksesta. [12]

3.3.1 Formatiivinen arviointi

Formatiiviseksi eli jatkuvaksi arvioinniksi kutsutaan sellaisia arviointimenetelmiä, missä oppilas saa kurssin aikana palautetta tai arviointi suorituksistaan. Formatiivinen arviointi voidaan jakaa formaaliin ja informaaliin arviointiin. Formaalissa arvioinnissa käytetään apuna testejä tai tarkastettavia tehtäviä. Informaali arviointi ei ole ennalta suunniteltua, vaan se voi tapahtua opettajan reagoidessa oppilaan väärään vastaukseen. [19] Vertais- ja itsearviointia voidaan pitää hyvänä esimerkkinä formatiivisista arviointimenetelmistä, sillä niissä opiskelija saa palautetta jo kurssin aikana. Formatiivisen arvioinnin tavoitteena on antaa opiskelijalle käsitys siitä, miten hyvin hän osaa kurssin asioita sillä hetkellä. Tällöin opiskelija pystyy reagoimaan palautteeseen opiskelussa kurssin aikana ja mahdollisesti myös oppimaan jotain saadusta palautteesta. Formatiivisesta arvioinnista on tehty useita tutkimuksia, joissa formatiivista arviointia saaneet ryhmät ovat saanut huomattavasti parempia oppimistuloksia kuin tutkimusten verrokkiryhmät [19]. Tällä kurssilla opiskelijat saivat suoritetuista tehtävistään pisteitä viikoittain. Tämä auttaa opiskelijoita tietämään miten hyvin he osaavat kurssin aiheita sillä hetkellä.

Vertais- ja itsearviointi

Vertaisarviointi ja itsearviointi ovat suosittuja arviointimenetelmiä usealla eri koulutustasolla. Useat korkeakoulujen kurssit hyödyntävät niitä osana arviointia [12]. Erityisesti opiskelijamääriltään suurilla kursseilla vertaisarvioinnin avulla pystytään antamaan palautetta kaikille opiskelijoille, mikä olisi usein mahdotonta, jos opettaja joutuisi antamaan palautteen yksilöllisesti kaikille. Vertaisarviointia käytetäänkin apuna viikoittaisten tehtävien tarkistamisessa tutkittavalla kurssilla. Kurssilla jokainen opiskelija antaa palautetta viikkotehtävistä kahdelle muulle opiskelijalla. Näin jokainen opiskelija saa viikoittain myös vertaisarvioinnin kahdelta muulta opiskelijoilta suorittamistaan tehtävistä [12]. Samalla vertaisarviointi toimii tukena oppimiseen vertaisarviointia tekeville opiskelijoille ja auttaa heitä ymmärtämään aiheita paremmin.

Alemmilla koulutusasteilla vertais- ja itsearviointi toimivat tärkeinä opetus- ja arviointimenetelminä. Lukion opetussuunnitelmassa itsearvioinnin merkitys ja käyttö osana arviointia tulee esille useassa kohtaa. Opetussuunnitelmassa itsearvioinnin ja vertaisarvioinnin sanotaan tukevan ja kehittävän opiskelijan omaa osaamista aiheessa. [13] Vastaavasti myös peruskoulun opetussuunnitelma puhuu itsearvioinnin käytöstä osana oppilaan arviointia. Peruskoulussa itsearvioinnin tarkoitus on auttaa oppilasta löytämään oman osaamisensa tason. Kun oppilas tunnistaa oman osaamisensa tason, hänen on helpompi huomata edistymisensä ja vaikuttaa siihen. [14] Realististen tavoitteiden asettaminen opinnoissa helpottaa oikeisiin asioihin keskittymistä.

3.3.2 Summatiivinen arviointi

Summatiivisella arvioinnilla tarkoitetaan sellaisia arviointimenetelmiä, jotka arvioivat oppilaan osaamista vasta kurssin lopussa. Tentti tai loppukoe ovat tyypillisiä esimerkkejä summatiivisesta arvioinnista. [19]

Kurssin tenttipohjaiselle suoritukselle arviointipohjana toimi ainostaan kurssin lopputentti. Pelkkä tentti tai loppukoe arviointimenetelmänä on menettänyt suosiotaan monilla eri kouluasteilla. Tentin pääasiallinen tarkoitus on mitata opiskelijan osaamisen tasoa kurssin aiheissa. Hyvin rakennetun tentin tulisi sisältää tehtäviä useimmista kurssin aikana käydyistä aiheista. Tällöin se antaa kaikille opiskelijoille mahdollisimman reilun vertailun siitä, miten hyvin he osaavat kurssin asiat. Matemaattisissa aineissa koe- tai tenttisuoritus mittaakin usein hyvin aiheen osaamista, sillä tehtävät pystytään kohdentamaan usein tarkasti opiskeltuihin aiheisiin.

4. REGRESSIOANALYYSI

Tässä luvussa esitellään tilastotieteen menetelmiä, jotka toimivat matemaattisena teoriana ja työkaluna tutkimuksen toteutuksessa. Erityisesti tutustutaan tarkemmin regressioanalyysin käyttöön työkaluna tutkimuksessa. Lisäksi määritellään tarkemmin, mitä regressioanalyysin malleja tutkimuksessa käytetään ja johdetaan keskeisia kaavoja tukemaan esiteltäviä määritelmiä.

Regressioanalyysin on tilastotieteen menetelmä, jolla voidaan tutkia usean selitettävän muuttujan suhdetta toiseen selitettävään muuttujaan [7]. Tässä työssä regressioanalyysiä käytetään esimerkiksi tutkimaan erilaisten kurssin osallistumistapojen vaikutusta oppimismenestykseen. Regressioanalyysin avulla voidaan selvittää kumpi osallistumistapa antaa parempia oppimistuloksia ja miten paljon oppimistulokset eroavat toisistaan eri osallistumistapojen välillä. Tässä työssä regressioanalyysilla vertaillaan kahden eri ryhmän välistä opintomenestystä kurssilla. Vertailtavana on esimerkiksi miesten ja naisten menestys kurssilla.

Regressioanalyysin tulkinnassa käytetään usein kuvaajia, joihin piirretty regressiosuora osoittaa muuttujien välisen riippuvuuden voimakkuutta ja riippuvuuden suuntaa. Regressiosuoran suunta kertoo onko muuttujien välillä negatiivinen vai positiivinen riippuvuus. Suoran kulmakerroin eli regressiokerroin kuvaa miten vahvaa tämä muuttujien välinen riippuvuus on. [7]

4.1 Lineaarinen regressioanalyysi

Lineaarisella regressioanalyysillä tarkoitetaan regressioanalyysin erikoistapausta, jossa arvioidaan aineistossa olevan muuttujien välistä lineaarista riippuvuutta. [7] Lineaarinen regressioanalyysi voidaan jakaa vielä tarkemmin yhden selittäjän lineaariseen regressioanalyysiin ja monen selittäjän lineaariseen regressioanalyysiin.

Monen selittäjän lineaarisessa regressioanalyysissa tarkasteltavaa muuttujaa pyritään selittämään usean muuttujan avulla [15]. Tässä työssä tutkitaan usean eri asian vaikutusta opintomenestykseen kurssilla. Tällä mallilla voidaan tutkia esimerkiksi miten opintomenestykseen vaikuttavat yhtäaikaisesti osallistumistavan valinta, opiskelijan sukupuoli ja opiskelun aloitusvuosi. Tässä esimerkissä opintomenestys on selitettävä muuttuja ja kurssin osallistumistapa, opiskelijan sukupuoli ja opiskelun aloitusvuosi ovat selittäviä muuttujia.

Monen muuttujan lineaarinen regressiomalli voidaan kirjoittaa muodossa

$$Y = b_0 + b_1X_1 + \cdots + b_kX_k + \epsilon, \tag{4.1}$$

missä Y on selitettävä muuttuja, b_0 on vakiotekijä, $X_1, \ldots, X_k$ ovat selittävät muuttujat ja $b_1, \ldots, b_2$ ovat niitä vastaavat regressiokertoimet. [15]

Tässä työssä satunnaistermi ϵ oletetaan normaalisti jakautuneeksi odotusarvolla 0 ja varianssilla σ^2. Yksittäiset satunnaistermit oletetaan myös toisistaan ja selittävästä muuttujasta riippumattomiksi eli niiden välille korrelaatiota ei muodosteta. Ilman tätä oletusta analyysien tulokset vääristyvät. [15]

Tässä tutkimuksessa yleinen malli pelkistyy kahden muuttujan tapaukseen, jolloin se voidaan kirjoittaa muodossa

$$Y = b_0 + b_1X_1 + b_2X_2 + \epsilon \tag{4.2}$$

missä edelleen Y on selitettävän muuttujan ennustettu arvo, b_0 on vakiotekijä, X_1 ja X_2 ovat selittävät muuttujat ja b_1 sekä b_2 ovat niitä vastaavat regressiokertoimet. [15] [7]

Regressiomallin avulla voidaan tehdä toistokoe, jossa realisoituvat selitettävä muuttuja Y ja selittävät muuttujat $X_1, X_2, ..X_k$. Selitettävä muuttuja voidaan esittää tällöin muodossa [16]

$$Y = b_0 + b_1X_1 + \cdots + b_kX_k + \epsilon. \tag{4.3}$$

Usein selittävät muuttujat voidaan valita vapaasti jolloin ne voidaan tulkitata deterministisiksi. Tällöin päästään lineaariseen deterministiseen malliin, jossa selitettävä muuttuja voidaan kirjoittaa ilman satunnaismuuttujaa ϵ muodossa [16]

$$Y = b_0 + b_1X_1 + \cdots + b_kX_k. \tag{4.4}$$

Satunnaismuuttuja ϵ voidaan tällöin olettaa normaalijakautuneeksi.

Käyttämällä vektoreita

$$x = \begin{bmatrix} 1 \\ X_1 \\ \vdots \\ X_m \end{bmatrix} \quad \text{ja} \quad b = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_m \end{bmatrix} \tag{4.5}$$

regressiomalli voidaan kirjoittaa lyhemmässä muodossa [16]

$$Y = x^T b + \epsilon. \tag{4.6}$$

Tuntemattomat parametrit $b_0, b_1, \ldots, b_k$ ja σ^2 pyritään määrittämään realisoimalla koe k kertaa, $k \geq m$. Koerealisaatiota numero i voidaan kuvata mallilla [16]

$$Y_i = b_0 + b_1 x_{i2} + \cdots + b_k x_{ik} + \epsilon_i = x_i^T b + \epsilon_i. \tag{4.7}$$

Kun koetta on toistettu k kertaa, se voidaan esittää muodossa

$$Y = \hat{X} b + \epsilon, \tag{4.8}$$

missä

$$\hat{X} = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_k^T \end{bmatrix} \text{ja } \epsilon \sim N_k(0, \sigma^2 I_k). \tag{4.9}$$

$\hat{X}$ on nyt $k \times m$-matriisi, jota kutsutaan myös datamatriisiksi. Tässä tehdään oletus, että $rank(X) = m$. Lisäksi matriisissa $\hat{X}$ on m saraketta ja $k > m$ eli pystyrivien määrä matriisissa on suurempi kuin vaakarivien. Tällöin matriisi on ϵ on multinormaalijakautunut parametreillä 0 ja $\sigma^2 I_k$ [16, s. 58].

Regressiokertoimien estimointi tehdään yleensä pienimmän neliösumman menetelmällä. Estimaatti a saadaan minimoimalla sovitusvirhevektorin kakkosnormin neliö [16]

$$a = \arg\min ||Y - \hat{X} b||^2. \tag{4.10}$$

Residuaalivektori $Y - \hat{X} a$ on ortogonaalinen matriisin $\hat{X}$ virittämällä aliavaruudella eli

$$\hat{X}^T (Y - \hat{X} a) = 0_p. \tag{4.11}$$

Tämä voidaan todistaa laskemalla funktion

$$||Y - \hat{X}a||^2 = (Y - \hat{X}a)^T(Y - \hat{X}a) \tag{4.12}$$

derivaatta muuttujan a suhteen [16]. Tällöin sen erotusosamäärä voidaan kirjoittaa muodossa

$$\begin{aligned}&\frac{(Y - \hat{X}(a+h))^T(Y - \hat{X}(a+h)) - (Y - \hat{X}a)^T(Y - \hat{X}a)}{||h||}\\ =&2(\hat{X}^T\hat{X}a - \hat{X}^T Y)^T\frac{h}{||h||} + O((||h||)).\end{aligned} \tag{4.13}$$

Tällöin $2(\hat{X}^T\hat{X}a - \hat{X}^T Y)^T$ on derivaatta muuttujan a suhteen ja on minimipisteessään nollavektori. Tämän seurauksena pätee yhtälö (4.11).

Tästä pystytään edelleen ratkaisemaan regressiokertoimien b pienimmän neliösumman estimaatti [16]

$$a = (\hat{X}^T\hat{X})^{-1}\hat{X}^T Y. \tag{4.14}$$

Kun kertoimet sijoitetaan malliin, saadaan realisoituneiden ulostulosuureen arvojen Y ja mallista saatujen arvojen erotuksesta residuaalivektori

$$e = Y - \hat{X}a. \tag{4.15}$$

Kun estimaatti a on laskettu, residuaalivektori e on tunnettu vektori. [16]

Regressiokertoimien ja residuaalivektorin tilastollinen luonne

Mittaustuloksista saatu vektori Y on satunnaisvektori sattunnaisuureen ϵ vuoksi. Tällöin myös vektorit a ja e ovat satunnaisuureita, koska ne on muodostettu satunnaisvektorin Y avulla. Kun oletamme käytetyn datan taustalle ideaalin mallin, saadaan [16]

$$Y = \hat{X}b + \epsilon, u \sim N_n(0, \sigma^2 I_n). \tag{4.16}$$

Sijoitetaan tämä Y satunnaisuureeseen a ja saadaan

$$\begin{aligned}a &= (\hat{X}^T\hat{X})^{-1}\hat{X}^T Y\\ &= (\hat{X}^T\hat{X})^{-1}\hat{X}^T(\hat{X}b + \epsilon)\\ &= b + (\hat{X}^T\hat{X})^{-1}\hat{X}^T\epsilon.\end{aligned} \tag{4.17}$$

Matriisi $(\hat{X}^T\hat{X})^{-1}\hat{X}^T$ on $p \times n$ matriisi ja $\mathrm{rank}((\hat{X}^T\hat{X})^{-1}\hat{X}^T) = p < n$. Tämän seurauksena vektori a on multinormaalijakuatunut. Vektorin ϵ odoutusarvon 0_n avulla saadaan $E(a) = b$. Kovarianssimatriisiksi saadaan [16]

$$\begin{aligned} V(a) &= (\hat{X}^T\hat{X})^{-1}\hat{X}^T V(\epsilon)\hat{X}(\hat{X}^T\hat{X})^{-1} \\ &= (\hat{X}^T\hat{X})^{-1}\hat{X}^T(\sigma^2 I_n)\hat{X}(\hat{X}^T\hat{X})^{-1} \\ &= \sigma^2(\hat{X}^T\hat{X})^{-1}(\hat{X}^T\hat{X})(\hat{X}^T\hat{X})^{-1} \\ &= \sigma^2(\hat{X}^T\hat{X})^{-1}. \end{aligned} \tag{4.18}$$

Täten saadaan satunnaisuureelle a jakaumatulokseksi

$$a \sim N_p(b, \sigma^2(\hat{X}^T\hat{X})^{-1}). \tag{4.19}$$

Vastaavasti voidaan selvittää residuaalivektorille $e = Y - \hat{X}a$ odotusarvo ja kovarianssimatriisi. Merkitään

$$P = I_n - \hat{X}(\hat{X}^T\hat{X})^{-1}\hat{X}^T, \tag{4.20}$$

jolloin e voidaan kirjoittaa

$$\begin{aligned} e = Y - \hat{X}a &= (\hat{X}a + \epsilon - x(b + (\hat{X}^T\hat{X})^{-1}\hat{X}^T\epsilon) \\ &= (I_n - \hat{X}(\hat{X}^T\hat{X})^{-1}\hat{X}^T)\epsilon = P\epsilon. \end{aligned} \tag{4.21}$$

Koska tiedetään, että

$$\begin{aligned} P^T &= (I_n - \hat{X}(\hat{X}^T\hat{X})^{-1}\hat{X}^T)^T \\ &= I_n^T - (\hat{X}^T)^T((\hat{X}^T\hat{X})^{-1})^T\hat{X}^T \\ &= I_n - \hat{X}(\hat{X}^T\hat{X})^{-1}\hat{X}^T = P \end{aligned} \tag{4.22}$$

niin matriisi P on symmetrinen. Lisäksi koska

$$\begin{aligned} PP &= (I_n - \hat{X}(\hat{X}^T\hat{X})^{-1}\hat{X}^T)(I_n - \hat{X}(\hat{X}^T\hat{X})^{-1}\hat{X}^T) \\ &= I_n - \hat{X}(\hat{X}^T\hat{X})^{-1}\hat{X}^T - \hat{X}(\hat{X}^T\hat{X})^{-1}\hat{X}^T + -\hat{X}(\hat{X}^T\hat{X})^{-1}\hat{X}^T\hat{X}(\hat{X}^T\hat{X})^{-1}\hat{X}^T \\ &= I_n - \hat{X}(\hat{X}^T\hat{X})^{-1}\hat{X}^T - \hat{X}(\hat{X}^T\hat{X})^{-1}\hat{X}^T + \hat{X}(\hat{X}^T\hat{X})^{-1}\hat{X}^T \\ &= I_n - \hat{X}(\hat{X}^T\hat{X})^{-1}\hat{X}^T \\ &= P \end{aligned} \tag{4.23}$$

niin matriisi P on idempotentti. Tällöin residuaalivektorille e saadaan odotusarvoksi

$$E(e) = E(P\epsilon) = PE(\epsilon) = 0_n \tag{4.24}$$

ja kovarianssimatriisiksi

$$V(e) = PV(\epsilon)P^T = P(\sigma^2 I_n)P = \sigma^2 P. \tag{4.25}$$

Residuaalivektorilla e ei kuitenkaan ole ei multinormaalijakaumaa.

Varianssin σ^2 estimointi

Parametrin σ^2 estimmattori määritellään residuaalivektorin $e = P\epsilon$ avulla. Tällöin [16]

$$e^T e = \epsilon^T P^T P\epsilon = \epsilon^T P\epsilon \tag{4.26}$$

Oletetaan että satunnaismuuttuja x on $N_n(0_n, \sigma^2 I_N)$-jakauma ja että symmetrinen $n \times n$ -matriisi A on idempotentti ja $\mathrm{rank}(A) = r$. Tällöin

$$\frac{1}{\sigma^2} x^T Ax \sim \chi^2(r) \tag{4.27}$$

Selvitetään matriisin P aste $\mathrm{rank}(P)$. Koska matriisin P kuva-avaruuden dimensio on n. Matriisin $\hat{X}$ sarakkeet virittävät matriisin nolla-avaruuden, sillä

$$P\hat{X} = (I_n - \hat{X}(\hat{X}^T \hat{X})^{-1}\hat{X}^T)\hat{X} = \hat{x} - \hat{X} = 0. \tag{4.28}$$

Tästä seuraa, että kaikki matriisin $\hat{X}$ sarakkeet kuuluvat matriisin P nolla-avaruuteen. Lisäksi havaitaan, että jos vektori v kuuluu matriisin $\hat{X}^T$ nolla-avaruuteen, niin $Pv = v$ eli vektori kuuluu matriisin P kuva avaruuteen. Tällöin matriisin P nolla-avaruuteen ei kuulu muuta kuin matriisin $\hat{X}$ sarakkeiden virittämät vektorit. Koska matriisi $\hat{X}$ on p lineaarisesti riippumatonta saraketta, niin matriisin P aste on $n - p$.

Matriisin P symmetrisyyden ja idempotenttiuden sekä edellisten oletuksen seurauksena. Kun $\epsilon \sim N_n(0_n, \sigma^2 I_n)$, niin

$$\frac{1}{\sigma^2}\epsilon^T P\epsilon \sim \chi(n-p). \tag{4.29}$$

Nyt $\epsilon^T P\epsilon = e^T e$, joten varianssin harhaton estimaattori on

$$\tilde{\sigma^2} = s_u^2 = \frac{1}{n-p} e^T e. \tag{4.30}$$

Nyt olemme estimoineet parametrit, kerroinvektorin b ja varianssin σ^2. Myös estimaattien käyttäytyminen satunnaisuureina tunnetaan niiden jakaumien avulla.

Regressiokertoimien luotettavuusvälit ja testaus

Seuraavaksi esittelemme regressiokertoimien luottamusvälit. Regressomallissa estimoidut kertoimet a_i ovat satunnaismuuttujia. Tämän vuoksi regressiokertoimille b_i muodostetaan luotettavuusvälit. [16]

Osoitetaan estimaattivektorin a ja residuaalivektorin e riippumattomuus. Tiedämme, että [16]

$$\begin{aligned} a &= b + (\hat{X}^T\hat{X})^{-1}\hat{X}^T\epsilon, \\ e &= P\epsilon = (I_n - \hat{X}(\hat{X}^T\hat{X})^{-1}\hat{X}^T)\epsilon \text{ja} \\ \epsilon &\sim N_n(0, \sigma^2 I_n). \end{aligned} \tag{4.31}$$

Normaalijakauman tapauksessa satunnaismuuttujat a ja e ovat riippumattomia, jos $cov(a, e) = 0$. Lasketaan kyseinen kovarianssimatriisi.

$$\begin{aligned} cov(a, e) &= cov((\hat{X}^T\hat{X})-1\hat{X}^T\epsilon, P\epsilon \\ &= (\hat{X}^T\hat{X})-1\hat{X}^T cov(\epsilon, \epsilon)P^T \\ &= \sigma^2(\hat{X}^T\hat{X})-1\hat{X}^T(I_n - \hat{X}(\hat{X}^T\hat{X})-1\hat{X}^T)^T \\ &= \sigma^2(\hat{X}^T\hat{X})-1\hat{X}^T - \sigma^2(\hat{X}^T\hat{X})-1\hat{X}^T(\hat{X}^T\hat{X})-1\hat{X}^T \\ &= \sigma^2(\hat{X}^T\hat{X})-1\hat{X}^T - \sigma^2(\hat{X}^T\hat{X})-1\hat{X}^T = 0 \end{aligned} \tag{4.32}$$

Joten satunnaismuuttujat a ja e ovat toisistaan riippumattomat. Myös vektorin a komponentit ovat vektorista e riippumattomia. Lisäksi

$$b \sim N_p(b, \sigma^2(\hat{X}^T\hat{X})^{-1}), \tag{4.33}$$

joten sen komponenteille pätee

$$b_i \sim N(b_i, \sigma^2 c_{ii}) \tag{4.34}$$

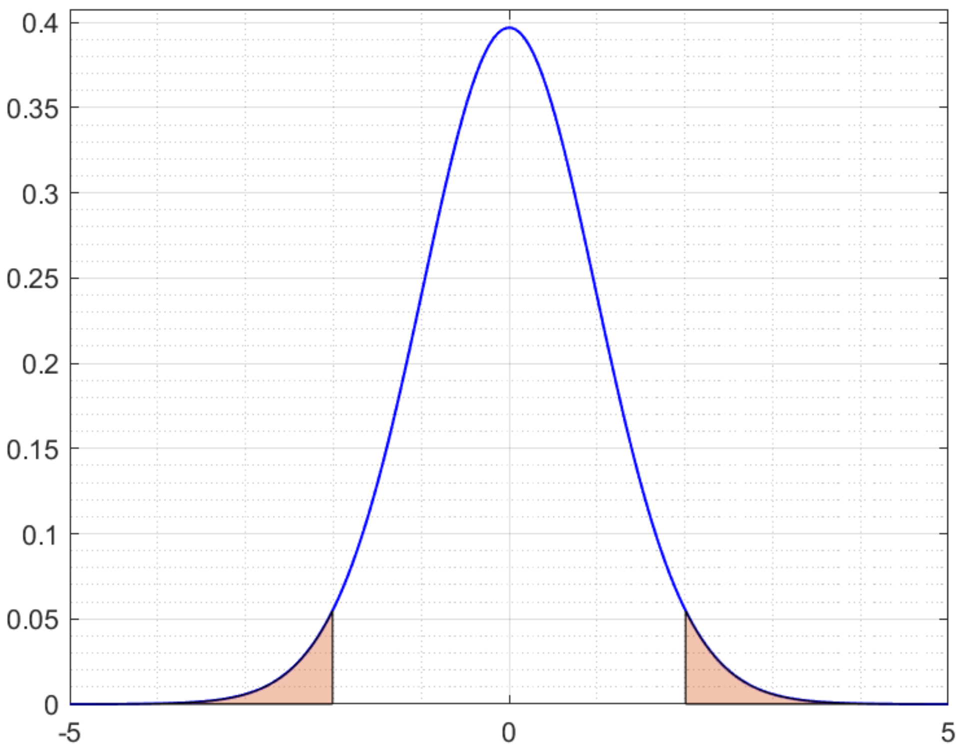
missä c_{ii} on matriisin $(\hat{X}^T\hat{X})^{-1}$ lävistäjäalkio. Tiedämme myös, että

$$\frac{1}{\sigma^2}e^T e \sim \chi^2(n-p). \tag{4.35}$$

Tällöin tilastollisen riippumattomuuden seuraksena voidaan määritellä

$$T = \frac{\frac{a_i - b_i}{\sigma\sqrt{c_{ii}}}}{\frac{\sqrt{e^T e}}{\sigma}}\sqrt{n-p} = \frac{a_i - b_i}{s_u\sqrt{c_{ii}}} \sim t(n-p). \tag{4.36}$$

Kuvassa 4.1 esitellään t-jakauman tiheysfunktio 95 $\%$ luottamusvälillä.



Kuva 4.1. *t-jakauman vapausastein 50 ja 95% kriittinen alue*

Tämän jälkeen etsimme $t(n-p)$-jakaumalle vakiot t_1 ja t_2 kuvan mukaisesti 95 prosentin luottamusvälillä. Luottamusväli esitellään tarkemmin vähän myöhemmin tässä alaluvussa. Regressiokertoimien b_i luottamusväliksi saadaan tällöin [16]

$$[a_i - t_2 s_u \sqrt{c_{ii}}, a_i - t_1 s_u \sqrt{c_{ii}}]. \tag{4.37}$$

Yhden selittäjän lineaarisessa regressiossa tarkastellaan tilannetta, jossa selitettävää muuttujaa Y pyritään selittämään yhden selittävän muuttujan X avulla. Tällaiselle kahden muuttujan regressioanalyysille voidaan kirjoittaa kaava

$$Y = b_0 + b_1 X + \epsilon, \tag{4.38}$$

missä Y on selitettävä muuttuja, X on selittävä muuttuja ja ϵ on satunnaistermi. b_0 ja b_1 ovat parametrejä, jotka yhtälöstä halutaan ratkaista. Parametriä b_0 voidaan nimittää vakiotermiksi ja b_1 regressiokertoimeksi. [15] Satunnaistermin ϵ odotusarvo oletetaan nollaksi. Samoja merkintöjä käytetään koko luvun ajan.

Regressiosuoran sovitus

Yhden selittäjän lineaarisessa regressioanalyysissa yhtälöstä muodostuu suora, jossa regressiokerroin on suoran kulmakerroin ja vakiotermi on suoran leikkauspiste. Jotta mallia kuvaava suora voidaan piirtää, tehdään käsiteltävälle aineistolle sovitus. Usein lineaarisen regressiomallin suoran sovitus tehdään pienimmimän neliösumman menetelmällä minimoimalla parametrit b_0 ja b_1 neliöiden summasta. [15]

$$\min_{b_0,b_1} \sum [y_i - (b_0 + b_1 x_i)]^2 = \min_{b_0,b_1} \sum [y_i - b_0 - b_1 x_i)]^2 \tag{4.39}$$

Parametrit b_0 ja b_1 valitaan siten, että lauseke saa niillä pienimmän mahdollisen arvonsa. [15] Käytännössä tämä voidaan toteuttaa niin, että Matlab-ohjelmalla etsitään parametreille sopivat arvot, joilla minimi voidaan laskea.

Luottamusvälillä tarkoitetaan parametrin arvon sellaista vaihteluväliä, johon sisältyy valitun luottamustason mukainen osuus, esim. 95% kaikista mahdollisista arvoista, joita parametrille voi realisoitua riippumattomassa toistokokeessa. Luottamustason voi esittää prosenttiosuuden lisäksi myös välillä $[0, 1]$ olevana lukuna $1 - \alpha$, missä α on se osuus realisoituvista arvoista, jotka eivät kuulu luottamusvälille. [15] Luottamusväli lineaarisen regression selitettävälle muuttujan odotusarvolle voidaan esittää yhden muuttujan tapauksessa. Usean muuttujan tapauksessa luottamusvälin täsmällinen esittäminen on haastavampaa. Yhden muuttujan tapauksessa luottamusväli luottamustasolla $1 - \alpha$ odotusarvolle $E(Y_u)$ on

$$E(Y_u) \pm t_{1-\alpha/2}(n - k - 1)s\sqrt{v_u}, \tag{4.40}$$

missä s on keskihajonta, luku $n - k - 1$ esittää t-jakauman vapausasteita ja v_u on varianssin komponentti. Alaindeksi u viittaa aineistoon kuulumattomaan havaintomateriaaliin. [15, s. 39]

4.2 Tilastollinen testaus regressioanalyysissä

Tilastollisia tunnuslukuja tarkastellessa on tärkeää arvioida niiden merkitsevyyttä. Tilastollisten tunnuslukujen merkistystä voidaan arvioida luottamusvälien ja erilaisten tilastollisten testausmenetelmien avulla. [15] Yksi näistä testausmenetelmistä on nimeltään t-testi.

T-testi on regressioanalyysissä työkalu, jonka avaulla selvitetään ovatko saadut tulokset järkeviä. Yleisesti t-testiä voidaan käyttää tilastotieteen tutkimusmenetelmänä, kun tutkittava aineisto on normaalijakautunut [4]. Tässä tutkimuksessa eri ryhmien arvosanajakaumat oletetaan normaalijakautuneiksi, joten t-testiä voidaan käyttää. T-testillä on useita eri muotoja riippuen testattavista ryhmistä. Yleisesti tunnettu t-testi on nimeltään Studentin t-testi [5]. Tässä testissä normaalijakautuneen muuttujan X testisuureena on

$$T = \sqrt{n}\frac{\bar{X} - \mu_0}{S_X}, \tag{4.41}$$

missä n on otoksen suuruus, $\bar{X}$ on otoskeskiarvo, μ_0 on odotusarvo ja S_X on keskivirhe otoskeskiarvolle. [22]

Jotta päästään tutkimaan tässä tutkimuksessa käytettävää aineistoa, tarvitaan kahden otoksen testiä. Tässä työssä käytettävää kahden otoksen t-testiä voidaan käyttää kahden toisistaan riippumattoman ryhmän keskiarvojen vertailuun [4]. Tämän testausmenetelmän avulla pystytään tutkimaan kaikkia tutkimuskysymyksiä, sillä kaikissa niissä opintomenestyksen vertailu tapahtuu vain kahden ryhmän välillä. Vertailtavien ryhmien välille oletetaan myös riippumattomuus. Kahden riippumattoman muuttujan t-testissä vertailtavien aineistojen variansseille ei vaadita yhtäsuuruutta t-testiä käytettäessä [4]. Kahden otoksen t-testille saadaan tällaisessa tilanteessa testisuureeksi:

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{m} + \frac{S_Y^2}{n}}}, \tag{4.42}$$

missä n ja m ovat otosten suuruudet, $\bar{X}$ ja $\bar{Y}$ ovat otoskeskiarvot ja S_X ja S_Y ovat otoskeskiarvojen keskivirheet. [22]

T-jakaumalla tarkoitetaan todennäköisyysjakaumaa, jota voidaan käyttää normaalijakautuneen aineiston tutkimisessa. Studentin t-testin yleisen muodon t-jakauman tiheysfunktio satunnaismuuttujalle t on

$$f_T(t) = \frac{\Gamma(\frac{v+1}{2})}{\sqrt{v\pi}\Gamma(\frac{v}{2})}(1 + \frac{t^2}{v})^{-\frac{v+1}{2}}, \tag{4.43}$$

missä v ilmaisee jakauman vapausasteet ja Γ on gammafunktio. [5]

Nollahypoteesilla H_0 tarkoitetaan tutkimuksessa tehdylle oletukselle tehtävää vastaoletusta, jonka mukaan vertailtavien ryhmien välillä ei ole eroa. Esimerkiksi ensimmäisessä tutkimuskysymyksessa oletetaan, että osallistuvalla tavalla kurssin suorittavat opiskelijat saavat parempia tuloksia, kuin tenttipainotteisella opetustavalla kurssia suorittavat opiskelijat. Tälle oletukselle voidaan tilastollisessa testauksessa asettaa vastaoletus eli nollahypoteesi, joka voisi olla, että tenttipainotteisella tavalla kurssin suorittaneet opiskelijat saavat vähintään yhtä hyviä arvosanoja kuin muutkin. Nollahypoteesin vastaoletuksena toimii vaihtoehtoinen hypoteesi, joka on tutkimuksessa on tehty oletus, että vertailtavien ryhmien välillä on eroa. Tätä oletusta merkitään usein H_1. [6]

Merkitään kurssia tenttipainotteisella tavalla suorittavia opiskelijoita μ_n ja kurssia osallistuvalla tavalla suorittavia opiskelijoita μ_m. Tällöin testille saataisiin nollahypoteesiksi ja sen vastaoletukseksi

$$\begin{aligned} H_0 &: \mu_n \geq \mu_m \\ H_1 &: \mu_n < \mu_m \end{aligned} \tag{4.44}$$

Testauksen avulla pyritään kumoamaan nollahypoteesi, jolloin tutkimuksessa tehty alkuperäinen hypoteesi pitäisi paikkaansa.

P-arvo kuvaa tutkittavalle tapahtumalle todennäköisyyttä, että se olisi tapahtunut sattumalta. P-arvon avulla tutkitaan onko tapahtumalla tilastollista merkitsevyyttä. Tilastollisesti merkitsevän p-arvon suuruus riippuu käytetystä luottamusvälistä, mutta yleisesti käytössä olevalle 95 prosentin luottamusvälille saadaan taulukossa 4.1 esitellyt merkitsevyystasot. [6]

p-arvo	Tilastollinen merkitsevyys
> 0.05	ei tilastollisesti merkitsevä
< 0.05	tilastollisesti melkein merkitsevä
< 0.01	tilastollisesti merkitsevä
< 0.001	tilastollisesti erittäin merkitsevä

Taulukko 4.1. *P-arvon tilastollisen merkitsevyyden tulkinta*

Tässä tutkimuksessa tehtävissä tilastollisissa analyyseissa käytetään 95 prosentin luottamusväliä, joten Taulukon 4.1 p-arvojen merkitsevyystasojen avulla tehdään tulkintaa tilastollisesta merkitsevyydestä koko tutkimuksen ajan.

5. TUTKIMUKSEN TOTEUTUS

Tutkimus toteutettiin Tampereen yliopistolta saadun tutkimusluvan pohjalta. Tutkimuksessa tarkastellaan opintomenestystä Johdatus todennäköisyyslaskentaan ja tilastolliseen päättelyyn -kurssilla. Erityisesti opintomenestystä tarkastellaan kurssia osallistavalla tavalla suorittavien ja kurssia tenttipainotteisella tavalla suorittavien opiskelijoiden välillä. Tutkimuksen aiheen valinta perustuu kiinnostukseen tutkia käänteisen oppimisen ja käänteisen opetuksen eli flippauksen oppimistuloksia. Kurssin osallistava osallistumistapa sisälsi useita erilaisia käänteisen opetuksen menetelmiä, kuten ennen yhteisiä tapaamisia videomateriaaleihin tutustumista. [12]

Aiemmin Tampereen yliopistossa on tehty myös muita tutkimuksia flippaukseen liittyen. Kuokkasen, Kangaslammen ja Hirvosen [8] artikkelissa Opetusmenetelmän ja oppimistehtävien yhteys korkeakouluopiskelijoiden matematiikan oppimiseen tutkitaan käänteisen oppimisen vaikutusta oppimistuloksiin korkeakouluissa. Tutkimuksen tuloksissa huomattiin käänteisen oppimisen mallilla opiskelleiden opiskelijoiden saaneen huonompia tuloksia kuin vertailuryhmän, joka opiskeli luentopohjaisella mallilla. Rämön ym. [18] artikkelissa "Engineering students’ approaches to learning in two student-centred instructional models before and during the COVID-19 pandemic" vertaillaan kahden eri opiskelijakeskeisen opetusmallin eroja insinööriopiskelijoiden keskuudessa koronapandemian aikana ja ennen pandemiaa. Tutkimuksessa pyrittiin selvittämään poikkesivatko opiskelijoiden lähestymistavat matematiikan opiskeluun eri opetusmalleja käytettäessä ja muuttuivatko opiskelijoiden lähestystavat insinöörimatematiikan opiskeluun. Vertailuryhmänä tutkimuksessa oli ryhmä, jonka luentokeskeiseen opetukseen lisättiin opiskelijakeskeisijä menetelmiä. Toinen ryhmä oli interventioryhmä, jonka opetuksessa perinteisiä opetusmalleja muutettiin vahvemmin käyttämällä esimerkiksi käänteisen opetuksen menetelmiä. Tuloksissa huomattiin, että invertioryhmän lähestymistavat opiskeluun laajentuivat enemmän kuin vertailuryhmällä.

5.1 Tutkimusaineisto

Tutkimuksen aineisto koostuu pääosin vuonna 2022 järjesteystä Johdatus todennäköisyyslaskentaan ja tilastolliseen päättelyyn -kurssilta kerätyistä tiedoista. Kurssi oli pääosin verkossa järjestetty massakurssi ja siihen osallistui yhteensä 607 tutkimusluvan antanutta opiskelijaa. Kurssin arvosanat ja kyselytulokset sisältävät aineisto on pseudonymisoitu. Aineiston käyttö perustuu tutkimusta varten haettuun tutkimuslupaan, joka antaa oikeuden käsitellä oppimistuloksia pseudonymisoidusti. Tämän lisäksi aineistoon sisältyy tällä ja aiemmilla saman lukuvuoden matematiikan kursseilla on järjestetty kysely, jolla on kartoitettu opiskelijoiden aiempaa matematiikan opintomenestystä yliopistossa. Tutkimuskyselyyn osallistui kaikkiaan noin 600 opiskelijaa.

Tarkastelemalla kurssin suorittaneiden arvosanoja tutkimusaineistosta saadaan vastauksia ensimmäiseen oppimiskysymykseen. Kurssista loppuarvosanan sai 574 opiskelijaa, joista 141 oli oman ilmoituksensa mukaan naisia ja 431 oli miehiä. Loput opiskelijat eivät halunneet ilmoittaa sukupuoltaan tai kertoivat olevansa muunsukupuolisia. Muunsukupuolisten tai sukupuolen ilmoittamatta jättäneiden opiskelijoiden oppimismenestystä ei vertailla miesten ja naisten välillä, koska heitä oli liian vähän pseudonymiteetin säilymiseen tutkimuksessa. Opiskelijoiden iät ja syntymävuodet saadaan selville myös tutkimusaineistosta.

Kurssista loppusuorituksen saaneiden välillä tutkimusaineistoista saadaan selville kummalla osallistumistavalla opiskelija on saanut kurssista suorituksen. Suoritustapa määräytyy virallisesti sen mukaan kummasta osallistumistavasta opiskelija on saanut enemmän pisteitä. Kurssin arvosanan määrävien pisteiden saaminen on esitelty luvussa 2. Vertailtavat ryhmät olivat keskenään lähes yhtä suuria. Yhteensä 572 opiskelijaa on tarkasteltavana tähän tutkimuskysymykseen liittyen aineistosta saatujen tietojen avulla. Loppuarvosanan kurssilta sai osallistuvalla suoritustavalla 250 oppilasta. Tenttipainotteisella suoritustavalla loppuarvosanan sai 322 oppilasta.

Kurssilla järjestettyyn kyselyyn opintojen aloitusvuodesta vastasi 570 opiskelijaa. Tätä tietoa käytetään apuna kolmannessa tutkimuskysymyksessä, kun opintomenestystä verrataan opiskeluaikaan yliopistossa. Tärkeimpänä tietona tutkimuksessa on opiskeluvuosien määrä kurssia suorittaessa. Työssä tarkastelu päädyttiin tekemään tässäkin kysymyksessä kahden ryhmän välillä. Ensimmäinen ryhmä, johon kuuluvat ensimmäisen vuoden opiskelijat, pitää sisällään suurimman osan kurssilta suorituksen saaneista opiskelijoista. Tähän ryhmään kuului yhteensä 510 opiskelijaa. Toisessa vertailtavassa ryhmässä ovat aiemmin opiskelunsa aloittaneet opiskelijat. Heitä oli yhteensä 60 opiskelijaa. Toiseen ryhmään kuului useana eri vuonna opiskelut aloittaneita opiskelijoita. Heitä ei tutkimuksessa jaettu useampaan vertailtavaan ryhmään, sillä se ei ollut järkevää heidän pienen lukumääränsä vuoksi.

5.2 Tutkimusmenetelmät

Tutkimuksen aineistoa käsitellään luvussa 4 esittellyin tilastollisin menetelmin. Taulukkomuodossa olevaa aineistoa käsitellään tutkimuksessa Matlab-ohjelmiston avulla. Opintomenestyksen mittarina tutkimuksessa toimii kurssin loppuarvosana.

Tilastollista testausta varten tutkimuksen aineistosta rajataan ryhmiä, joista saatavat tiedot vastaavat tutkimuskysymysten asettelua. Esimerkiksi kurssin kahden eri osallistumistavan välistä opintomenestystä pystytään vertaamaan lineaarisen regressioanalyysin avulla. Tutkimusaineistosta haetaan niiden opiskelijoiden arvosanat, jotka ovat suorittaneet kurssin tenttipainotteisella suoritustavalla ja niitä verrataan osallistuvalla suoritustavalla kurssin käyneiden opiskelijoiden arvosanoihin. Arvosanoista luodaan Matlab-ohjelmistolla yleistetty lineaarinen malli. Lineaarisen mallin avulla saadaan yksi tilastollinen tulkinta siitä, miten paljon osallistumistapa vaikuttaa arvosanaan. Lineaarisen mallin analyysia käytetään tutkimuksessa myös, kun vertaillaan sukupuolien välistä eroa opintomenestyksessä. Molemmissa tapauksisa tilastollinen testaus tehdään 95 prosentin luottamusvälillä.

Yleistetty lineaarisen monen muuttujan regressionalyysin malli sopii hyvin tilanteisiin, joissa käsiteltäviä ryhmiä on vain kaksi. Tätä mallia voidaan käyttää, kun tarkastellaan kurssin opintomenestystä kahden ryhmän välillä. Ensimmäisen tutkimuskysymyksen kohdalla opiskelijat jaetaan kahteen ryhmään kurssin osallistumistapojen välillä. Toisessa tutkimuskysymyksessä opiskelijat jaetaan kahteen ryhmään sukupuolen perusteella. Myös kolmannen tutkimuskysymyksen kohdalla voidaan tehdä kahtiajako ensimmäisen vuoden opiskelijoiden ja muiden kuin ensimmäisen vuoden opiskelijoiden välillä.

6. TUTKIMUKSEN TULOKSET

Tässä kappaleessa tutkimuskysymyksiin pyritään vastaamaan vertailemalla tuloksia eri ryhmien välillä. Kurssin tuloksia vertaillaan Matlab-ohjelmistolla tilastollisten tunnuslukujen ja regressioanalyysin avulla.

Tutkittava ryhmä koostuu kurssista arvosanan saaneista opiskelijoista. Kussista loppuarvosanan sai yhteensä 572 opiskelijaa. Opiskelijoista 458 sai hyväksytyn arvosanan kurssista ja 114 opiskelijaa sai hylätyn arvosanan. Kurssista loppuarvosanan osallistumispainotteisella suoritustavalla sai 250 opiskelijaa ja tenttipainotteisella suoritustavalla 322 opiskelijaa.

6.1 Tulokset tilastollisten tunnuslukujen avulla

Tässä luvussa pyritään vastaamaan yhdestä näkökulmasta ensimmäiseen, toiseen ja kolmanteen tutkimuskysymykseen. Tilastollisten tunnuslukujen avulla vastataan tutkimuskysymyksiin eri osallistumistapojen, sukupuolien ja opiskeluvuosien määrän vaikutuksesta kurssin tuloksiin. Tilastollisista tunnusluvuista käytämme keskiarvoa ja keskiarvon luottamusväliä määrittämään ryhmän opintomenestystä. Tuloksista tarkastellaan keskiarvojen eron lisäksi tilastolliselle testaamiselle tyypilliseen tapaan tuloksen tilastollista merkitsevyyttä p-arvon avulla.

6.1.1 Eri osallistumistapojen vaikutus opintomenestykseen

Tutkimuksen ensimmäiseen tutkimuskysymykseen saadaan vastauksia tulkitsemalla eri osallistumistapojen välistä menestystä kurssilla.

Arvosanojen välisen eron testaamista varten käsitellään tuloksia tilastollisella testaamisella. Kurssin eri osallistumistapojen välisistä tuloksista lasketaan Matlabin avulla tilastollisia tunnuslukuja. Kun tarkastellaan arvosanan keskiarvoja kummallakin osallistumistavalla, saadaan taulukossa 6.1 esitellyt tulokset.

Osallistumistapa	Keskiarvo	Keskiarvon luottamusväli
Osallistuva	2.84	[2.66, 3.03]
Tenttipainotteinen	2.32	[2.12, 2.52]

Taulukko 6.1. *Arvosanojen keskiarvot eri osallistumistapojen välillä*

Tulosten vertailussa on käytetty 95 prosentin luottamusväliä. Kurssin osallistuvalla suoritustavalla suorittaneet opiskelijat ovat saaneet selvästi parempia arvosanoja kurssista kuin tenttipainotteisella suoritustavalla kurssin suorittaneet opiskelijat. P-arvoksi tälle tulokselle saadaan 0.0002. Erittäin pieni p-arvo kertoo tuloksen olevan tilastollisesti erittäin merkitsevä. Keskiarvojen eron ja pienen p-arvon perusteella voidaan todeta, että kurssilla tenttipainotteinen opiskelutapa tuotti merkittävästi heikompia tuloksia kuin osallistuva opiskelutapa. Tulos on ennakko-oletuksen mukainen ja vastaa osaltaan ensimmäiseen tutkimuskysymykseen osallistumistapojen eroista opintomenestyksessä.

6.1.2 Sukupuolen vaikutus opintomenestykseen

Sukupuolen vaikutusta opintomenestykseen kurssilla tutkittaessa tutkimusaineistosta valittiin opiskelijat, jotka ovat saaneet kurssista suorituksen ja antaneet tutkimusluvan. Heidän joukostaan poistettiin opiskelijat, jotka olivat ilmoittaneet sukupuolensa muuksi tai eivät halunneet kertoa sitä. Tämän tutkittavan ryhmän kooksi jäi 572 henkilöä, joista 431 oli miehiä ja 141 naisia.

Tutkittavasta ryhmästä saatiin arvosanojen keskiarvot ja niiden luottamusvälit sukupuolittain. Nämä tulokset on esitelty taulukossa 6.2.

Sukupuoli	Keskiarvo	Keskiarvon luottamusväli
Miehet	2.45	[2.29, 2.61]
Naiset	2.86	[2.60, 3.12]

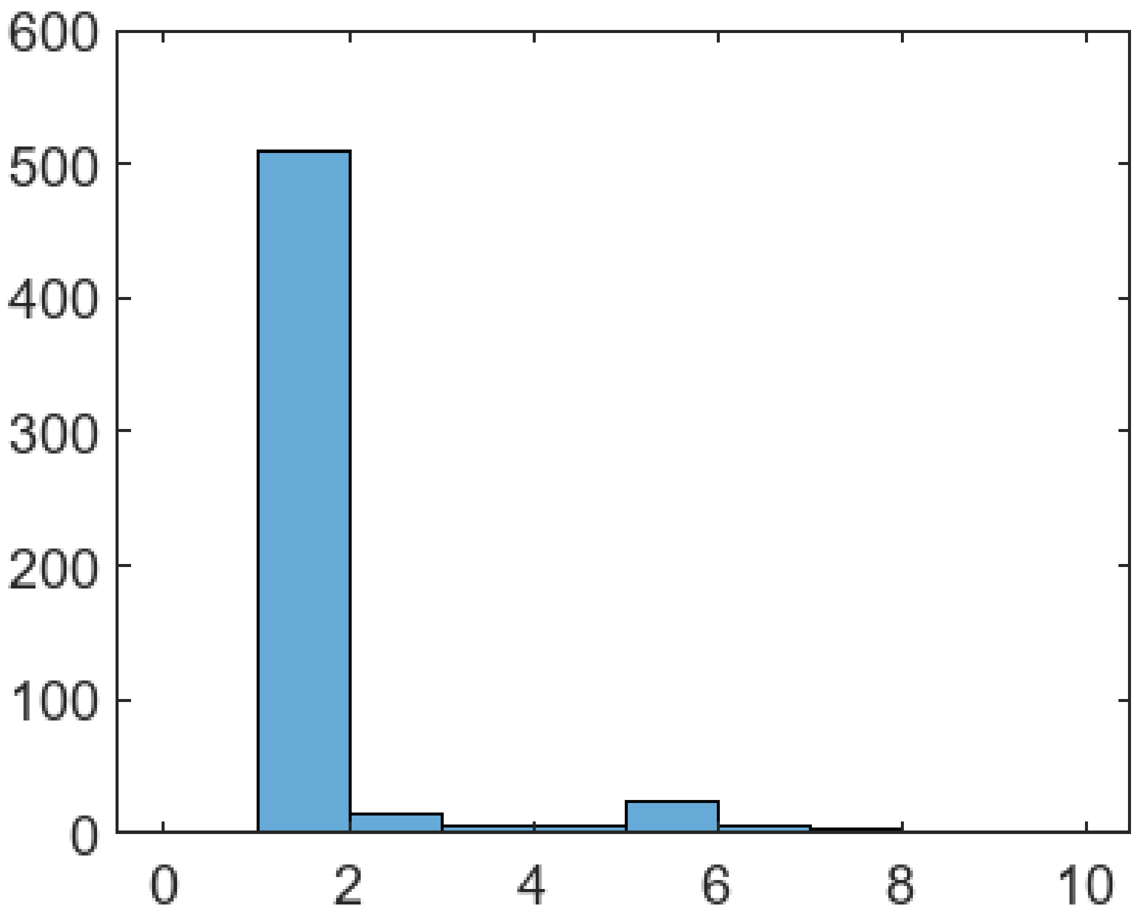
Taulukko 6.2. *Arvosanojen keskiarvot eri sukupuolten välillä*

Naisten huomataan saaneen kurssista selvästi parempia tuloksia kuin miesten. P-arvoksi tälle tulokselle saadaan 0.0126, jota voidaan pitää melko pienenä ja tulosta voidaan pitää tilastollisesti melkein merkitsevänä.

Saatuja tuloksia voidaan pitää toisen tutkimuskysymykseen kannalta olennaisina. Ennakko-oletuksen mukainen ero miesten ja naisten välillä saadaan oletetusti naisten eduksi.

6.1.3 Aiempien opintojen määrän vaikutus opintomenestykseen

Aiempien opintojen vaikutusta opintomenestykseen tutkitaan vertailemalla ryhmittäin JTT-kurssin arvosanoja eri lukuvuosina opintonsa aloittaneiden opiskelijoiden välillä. Kurssi on tarkoitus käydä ensimmäisen vuoden aikana, joten voidaan olettaa, että suuri osa myöhemmin kurssia suorittavista opiskelijoista uusii kurssia. Tämän pohjalta voidaan olettaa ensimmäisen vuoden opiskelijoiden saavan keskimäärin parempia arvosanoja kurssista kuin myöhemmässä vaiheessa opintoja olevat opiskelijat. Tutkimuksessa vertailtiin ensimmäisen vuoden opiskelijoiden ja muiden kuin ensimmäisen vuoden opiskelijoiden arvosanoja. Useampaan kuin näihin kahteen ryhmään jakaminen tarkastelua varten ei ollut järkevää, koska aiempien vuosikurssien osallistujien määrät olivat tiettyjen opiskeluvuosien osalta melko pieniä. Kurssin jakauma opiskeluvuosien mukaan on esitelty kuvassa 6.1.



Kuva 6.1. *Histogrammi kurssin opiskelijoista opiskeluvuosien lukumäärän mukaan*

Kyselyyn vastanneita kurssin suorittaneita ensimmäisen vuoden opiskelijoita on 510 ja muun kuin ensimmäisen vuoden opiskelijoita on 60. Näiden kahden ryhmän arvosanojen keskiarvot ja niiden luottamusvälit on esitelty taulukossa 6.3.

Osallistumistapa	Keskiarvo	Keskiarvon luottamusväli
Osallistuva	2.65	[2.51, 2.80]
Tenttipainotteinen	1.78	[1.35, 2.21]

Taulukko 6.3. *Arvosanojen keskiarvot eri osallistumistapojen välillä*

Tuloksista huomataan, että ensimmäisen vuoden opiskelijat saavat odotetusti parempia tuloksia kurssista kuin pidempään opiskelleet opiskelijat. Tuloksista saatiin p-arvoksi 0.0001. Tätä tulosta voidaan pitää tilastollisesti erittäin merkitsevänä.

6.2 Tulokset regressioanalyysillä

Regressioanalyysiä käytetään tässä työssä toisena vertailukeinona tuloksille. Tilastollisten tunnuslukujen avulla saatiin tuloksia arvosanojen keskiarvoista jokaiselle ryhmälle. Lineaarisen regressioanalyysin avulla vertaillaan jokaisen tutkimuskysymyksen kohdalla oletettua arvosanojen eroa suhteessa vertailtavaan ryhmään. Vertailtavana ryhmänä käytetään opiskelijoita, joiden opintomenestys oli huonompi tunnuslukujen avulla saatujen tulosten kohdalla. Regressioanalyysi tehdään Matlab-ohjelmiston avulla.

Tutkitaan ensin kaikkien tutkimuskysymysten yksi, kaksi ja kolme vaikutusta oppimistuloksiin. Tällöin lineaarisen regressioanalyysin malliin vaikuttavat opintovuosien määrä, opiskelijan sukupuoli ja valittu kurssin osallistumistapa.

Ensimmäisessä tutkimuskysymyksessä halutaan selvittää kurssin eri osallistumistapojen vaikutusta opintomenestykseen kurssilta. Tutkittavat osallistumistavat ovat perinteinen tenttipainotteinen ja osallistuva opiskelutapa. Kahden eri osallistumistavan välille saadaan regressioanalyysillä keskimäärin seuraavia arvoja. Lineaarisen mallin avulla saaduista tuloksista huomataan, että osallistuvalla opiskelutavalla kurssin suorittaneet oppilaat saavat kurssista oletetusti arvosanan, joka on 0.49 parempi kuin tenttipainotteisella osallistumistavalla kurssin suorittaneet oppilaat. P-arvoksi tälle tulokselle saadaan 0.0005. Erittäin pienen p-arvon vuoksi tulosta voidaan pitää tilastollisesti erittäin merkitsevänä. Tämä tulos vahvistaa jo aiemmin arvosanojen keskiarvoja tutkittaessa saatua tulosta.

Toisessa tytkimuskysymyksessä käsitellään sukupuolen vaikutusta opintomenestykseen. Sukupuolen vaikutusta kurssin opintomenestykseen lineaarisella regressioanalyysillä vertailtaessa käytetään edelleen ryhminä ainoastaan miehiä ja naisia. Regressioanalyysin tuloksien perusteella tällä kurssilla naisten oletettu arvosana on 0.38 suurempi kuin miehillä. Tulos vahvistaa myös tilastollisten tunnuslukujen avulla saatua tulosta. Tälle tulokselle p-arvo on 0.016. P-arvo on tilastollisesti sen suuruinen, että tulosta voidaan pitää ainoastaan tilastollisesti melkein merkitsevänä.

Kolmannen tutkimuskysymyksen kohdalla tutkitaan aiempien yliopisto-opintojen määrän vaikutusta opintomenestykseen tällä kurssilla. Regressioanalyysin avulla tehdyssä tutkimuksessa ryhminä ovat ensimmäisen vuoden opiskelijat ja pidempään yliopistossa opiskelleet. Lineaarisen regressioanalyysin tuloksien mukaan ensimmäisen vuoden opiskelijat saavat kurssista oletetusti arvosanan, joka on 0.81 parempi kuin kauemmin yliopistossa opiskelleet opiskelijat. Tulokselle saadaan P-arvoksi 0.0004. Erittäin pieni p-arvo tekee tuloksesta tilastollisesti erittäin merkitsevän.

7. YHTEENVETO

Tutkimuksen tarkoituksena oli selvittää eroavaisuuksia erilaisten ryhmien välisessä opintomenestyksessä Johdatus todennäköisyyteen ja tilastolliseen päättelyyn -kurssilla. Tutkimusaineistona toimivat JTT-kurssin suoritetiedot ja opiskelijoiden vastaukset kurssilla teetettyyn tutkimusluvan saaneeseen kyselyyn. Kyselyyn vastanneita opiskelijoita, jotka olivat antaneet tutkimusluvan, oli kurssilla yli 600. Tutkimuksessa arvosanoja vertailtiin ainoastaan opiskelijoilta, joilta oli vähintään yksi tenttisuoritus kurssilta. Tässä luvussa tarkastellaan tarkemmin tutkimuksessa tilastollisen päättelyn ja regressioanalyysin avulla saatuja tuloksia, sekä analysoidaan niiden luotettavuutta. Luku vastaa myös siihen, miten tutkimuskysymykset toteutuivat tutkimuksen tulosten perusteella.

Vastauksia tutkimuskysymyksiin

Ensimmäisessä tutkimuskysymyksessä tutkittiin opintomenestystä kurssin kahden eri osallistumistavan välillä. Tutkittavina osallistumistapoina olivat perinteisempi tenttikeskeinen osallistumistapa ja erilaisia käänteisen opetuksen keinoja käyttävä osallistuva osallistumistapa. Osallistuvalla tavalla kurssista sai arvosanan 44% tutkimusluvan antaineista opiskelijoista ja tenttipainotteisella osallistumistavalla 56% tutkimusluvan antaneista opiskelijoista. Ensimmäisen tutkimuskysymyksen yhteydessä oletuksena oli, että osallistuvalla opiskelulla saadaan parempia tuloksia. Tätä oletusta voidaan saatujen tulosten pohjalta pitää oikeana. Tutkimuksen tulosten mukaan osallistuvalla tavalla kurssin suorittavan arvosana oli keskimäärin 0.49 suurempi kuin tenttipainotteisella tavalla kurssia suorittavan opiskelijan arvosana.

Osallistavassa tavassa käytettävien opetusmenetelmien voidaan olettaa olevan tehokkaampia ja monipuolisempia oppimisen kannalta. Toisaalta voidaan myös ajatella, että osallistava tapa suorittaa kurssi vaatii opiskelijalta useampia suoritteita kuten harjoitustehtäviä ja ryhmätöitä, jolloin opiskelija keskittyy kurssin suorittamiseen tasaisesti koko opintojakson ajan.

Toisessa tutkimuskysymyksessä käsiteltiin eroja sukupuolten välisessä menestyksessä kurssilla. Analysoitavassa materiaalissa pseudonymiteetin takaamiseksi tutkittiin ainoastaan opintomenestystä miesten ja naisten välillä. Toisen tutkimuskysymyksen yhteydessä oli oletus, että naiset saavat miehiä parempiä arvosanoja. Tutkimusluvan antaneista kyselyyn vastaajista 75% oli miehiä ja 25% oli naisia. Oletuksen pohjalla oli useita tutkimuksia, esimerkiksi suomalaisten tyttöjen ja poikien erot PISA-vertailussa [21]. Tämä oletus toteutui melko selkeästi tuloksissa, sillä naisille odotusarvo kurssin arvosanasta tutkimuksen mukaan oli 0.38 numeroa suurempi kuin miehille.

Kolmannessa tutkimuskysymyksessä vertailtiin opiskelijoiden menestystä opiskeluvuosien lukumäärän perusteella. Ensimmäisen vuoden opiskelijat muodostivat kurssilla runsaan enemmistön, joten tutkimusta tehtäessä vertailtavat ryhmät olivat ensimmäisen vuoden opiskelijat ja pidempään opiskelleet. Kyselyyn vastannaista tutkimusluvan antaneista opiskelijoista 89% oli ensimmäisen vuoden opiskelijoita ja 11% oli opiskellut pidempään. Oletuksena tutkimuskysymyksessä oli, että ensimmäisen vuoden opiskelijat saavat parempia tuloksia kuin muut. Oletuksen perustana on ajatus siitä, että moni muu kuin ensimmäisen vuoden opiskelija uusii kurssia, sillä kurssi on suunnattu ensimmäisen vuoden opiskelijoille. Kurssia uusivat opiskelijat saavat todennäköisesti keskimäärin muita heikompia arvosanoja.

Tutkimuksessa saatujen tulosten ja tutkimuskysymysten yhteydessä tehtyjen oletusten välillä on yhteys, joka pitää paikkansa jokaisen tutkimuskysymyksen kohdalla. Toisen tutkimuskysymyksen tapauksessa erosta naisten ja miesten välille saadaan vähemmän tilastollista merkitsevyyttä kuin muille tutkimuskysymyksille. Tutkimuskysymysten yhteydessä tehtyjä ennakko-oletuksia voidaan pitää hyvinä, sillä niillä oli vahva vastaavuus tutkimuksen tuloksiin. Vaikka tulokset ovat oletettuja, voidaan erityisesti käänteisen oppimisen positiivista vaikutusta opintomenestykseen pitää mielenkiintoisena tuloksena. Käänteisen oppimisen tutkimusta suomalaisen yliopistomatematiikan osalta ei ole vielä kovin runsaasti, joten lähes kaikkia siihen liittyviä tutkimuksituloksia voidaan pitää kiinnostavina.

7.1 Tutkimuksen luotettavuus ja toistettavuus

Tutkimuksessa käytetty otoskoko on suuri, sillä tutkimuksessa analysoitiin lähes 600 opiskelijan opintomenestystä. Tutkimusta voidaan siis pitää kvantitatiivisena tutkimuksena. Kvantitatiivisessa tutkimuksessa tutkimuksen luotettavuutta määritellään sekä tutkimuksen validiteetin että reliabiliteetin avulla. Kvantitatiivisessä tutkimuksessa validiteetillä tarkoitetaan sitä miten hyvin mitataan haluttua asiaa. Reliabiliteetti kvantitatiivisessa tutkimuksessa tarkoittaa tutkimuksen johdonmukaisuutta. Tarkemmin se kuvaa sitä, muuttuvatko tutkimuksen tulokset ajan seuraksena, tai ovatko tehdyt mittaukset vertailukelpoisia eli mitataanko vertailtavien ryhmien välillä samaa asiaa. [11]

Tutkimuksen reliabiliteettia voidaan tarkastella tutkimuskysymysten tulosten avulla. Kaikkiin tutkimuskysymyksiin vastataan regressioanalyysin menetelmillä, tarkemmin kahden muuttujan lineaarissella regressioanalyysilla, joka on esitelty tarkemmin luvussa 3.

Tutkimuksen tulokset on kerätty käyttämällä Matlab-ohjelmistoa. Matlab-ohjelmisto on yleisesti matriisi- ja vektorilaskennassa käytetty ohjelma, jonka laskennallisia ominaisuuksia käytettiin apuna tulosten laskemisessa tässä tutkimuksessa. Matlab-ohjelmistoa voidaan pitää luotettavana tieteellisen tutkimuksen näkökulmasta laskujen tarkkuudessa ja tilastollisen tulkinnan esittämisessä. Tulosten laskemista samoilla menetelmillä samaan tai uusiin aineistoihin voidaan pitää toistettavana. Myös tutkimuksessa käytetyn aineiston kaltaisten aineiston muodostaminen yliopisto-opiskelijoiden oppimistuloksista on mahdollista. Tämä vahvistaa osaltaan tutkimuksen toistettavuutta.

Validiteetin kannalta tutkimuksen tulokset vastaavat hyvin tutkimuskysymyksiin. Tutkimuksessa saadaan selville eroja vertailtavien ryhmien välillä jokaisen tutkimuskysymyksen kohdalla. Ensimmäisen ja kolmannen tutkimuskysymyksen kohdalla tehdyistä tutkimuksista saadut tulokset ovat tilastollisesti vakuuttavia. Niistä saadut p-arvot ovat todella pieniä, jolloin saatuja tuloksia voidaan pitää tilastollisesti erittäin merkittävinä. Näistä kysymyksistä saatuihin tuloksiin virheen mahdollisuus on todennäköisyyksien valossa pieni.

LÄHTEET

[1] Paul Baepler, D. Christopher Brooks ja J. D. Walker (toim.) *Active Learning Spaces*. John Wiley & Sons, Inc, 2014. 112 s. URL: https://www.ellibslibrary.com/book/9781118870112.

[2] Elizabeth F Barkley. *Interactive Lecturing: A Handbook for College Faculty*. Newark: John Wiley & Sons, Inc, 2018.

[3] Mikael Cronhjort, Lars Filipsson ja Maria Weurlander. ”Improved Engagement and Learning in Flipped-Classroom Calculus”. *Teaching Mathematics and its Applications: An International Journal of the IMA* 37.3 (toukokuu 2017), s. 113–121. DOI: 10.1093/teamat/hrx007.

[4] Tarja Heikkilä. *Tilastollinen tutkimus 9. Uudistettu painos*. Edita Publishing Oy, 2014.

[5] Zongliang Hu ym. ”Regularized t Distribution: Definition, Properties, and Applications.” 50.4 (2023), s. 1884–1900. DOI: 10.1111/sjos.12655.

[6] Markus Kaakinen ja Noora Ellonen. ”Hypoteesien testaus. Teoksessa *Kvantitatiivisen tutkimuksen verkkokäsikirja*.” (2021). https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvanti/hypoteesi/testaus/, {viitattu 21.9.2023}.

[7] Markus Kaakinen ja Noora Ellonen. ”Regressioanalyysi. Teoksessa *Kvantitatiivisen tutkimuksen verkkokäsikirja*.” (2021). https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvanti/regressio/analyysi/, {viitattu 23.7.2022}.

[8] Tytti Kuokkanen, Riikka Kangaslampi ja Jani Hirvonen. ”Opetusmenetelmän ja oppimistehtävien yhteys korkeakouluopiskelijoiden matematiikan oppimiseen”. *Yliopistopedagogiikka* 2/2022 (2022). URL: https://lehti.yliopistopedagogiikka.fi/2022/12/22/opetusmenetelman-ja-oppimistehtavien-yhteys-korkeakouluopiskelijoiden-matematiikan-oppimiseen/.

[9] Wes Maciejewski. ”Flipping the Calculus Classroom: An Evaluative Study”. *Teaching Mathematics and its Applications: An International Journal of the IMA* 35.4 (2016), s. 187–201.

[10] Anastasia Misseyanni ym. *Active Learning Strategies in Higher Education: Teaching for Leadership, Innovation, and Creativity*. Bingley, England: Emerald Publishing, 2018. 399 s.

[11] ”Mittaaminen: Mittarin luotettavuus Teoksessa *Kvantitatiivisen tutkimuksen verkkokäsikirja*.” (2021). https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvanti/mittaaminen/luotettavuus/, {viitattu 28.1.2024}.

[12] *Moodle: MATH.APP.210 Johdatus todennäköisyyslaskentaan ja tilastolliseen päättelyyn K-2022*. https://moodle.tuni.fi/course/view.php?id=18675, {viitattu 10.5.2022}.

[13] Opetushallitus. ”Lukion opetussuunnitelman perusteet” (2019). https://www.oph.fi/sites/default/files/documents/lukion_opetussuunnitelman_perusteet_2019.pdf, {viitattu 19.6.2022}.

[14] Opetushallitus. ”Perusopetuksen opetussuunnitelman perusteet” (2014). https://www.oph.fi/sites/default/files/documents/perusopetuksen_opetussuunnitelman_perusteet_2014.pdf, {viitattu 19.6.2022}.

[15] Pekka Pere. ”Opinto-opas Tilastomenetelmien perusteet” (2021).

[16] Armo Pohjavirta ja Keijo Ruohonen. ”Laaja tilastomatematiikka” (2005).

[17] Henry L. Roediger ja Jeffrey D. Karpicke. ”Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention”. eng. *Psychological science* 17.3 (2006), s. 249–255. ISSN: 0956-7976.

[18] Johanna Rämö ym. ”Engineering students’ approaches to learning in two student-centred instructional models before and during the COVID-19 pandemic”. *European Journal of Engineering Education* 48.6 (2023), s. 1186–1212. DOI: 10.1080/03043797.2023.2206794.

[19] Marika Toivola. *Käänteinen arviointi*. Edita Publishing Oy, 2019. 119 s.

[20] Marika Toivola, Pekka Peura ja Markus Humaloja. *Flipped learning - Käänteinen oppiminen*. Edita Publishing Oy, 2017. 137 s. URL: https://www.ellibslibrary.com/book/978-951-37-7128-7.

[21] Minna Torppa ym. ”Why Do Boys and Girls Perform Differently on PISA Reading in Finland? The Effects of Reading Fluency, Achievement Behaviour, Leisure Reading and Homework Activity.” *Journal of Research in Reading* 41.1 (2018). DOI: 10.1111/1467-9817.12103.

[22] Aad van der Vaart. *An Introduction to Mathematical Statistics. 1st ed.* Edita Publishing Oy, 2017.