

# Tutkintovaatimukset tilastotieteen pääaineopiskelijalle 2011–2012

## Tilastotiede

### Yleistä

### Luonnontieteiden kandidaatti / Valtiotieteiden kandidaatti

### Pääaineopinnot

### Sivuaineopinnot

### Muut opinnot

### Filosofian maisteri / Valtiotieteiden maisteri

### Jatkotutkinnot tilastotieteessä

## Tutkintovaatimukset tilastotieteessä

Tilastotieteen koulutusohjelman tutkintovaatimukset ovat matemaattis-luonnontieteelliselle ja valtiotieteelliselle tiedekunnalle rakenteeltaan yhteiset. Joitakin tiedekuntakohtaisia eroja esiintyy erityisesti maisteriopinnoissa, ja ne mainitaan erikseen tekstissä.

Tutkintovaatimuksissa on pyritty joustavuuteen, jotta kahden tiedekunnan pääaineopiskelijoiden sekä kaikkien sivuaineopiskelijoiden tarpeet voidaan huomioida mahdollisimman hyvin. Kurseja koskevat seuraavat yhteiset huomautukset. Kurssit perustuvat luentomateriaaliin tai luennoijan kanssa sovittavaan kirjallisuuteen. Pakolliset kurssit on osoitettu (P)-merkinnällä. Poikkeustapauksissa voidaan sopia korvaavuuksista (esim. opiskelijan vaihtaessa pääaineensa tilastotieteeksi). Monia kurseja voi suorittaa eri laajuisina, mutta saman kurssin eri osat tulee suorittaa saman lukuvuoden aikana kuin kurssi on pidetty, ellei opettaja ilmoita toisin. Opintokokonaisuuksiin ei voi sisällyttää oppiainekseltaan päällekkäisiä kurseja. Kurssi saatetaan luennoida englanniksi, vaikka kurssin kuvaus on alla suomeksi. Kurssien tarkemmat kuvaukset löytyvät kunkin kurssin verkkosivulta.



#### **HUOM!**

Nämä tutkintovaatimukset koskevat kaikkia 1.8.2011 tai sen jälkeen aloittaneita opiskelijoita. Aiemmin opintonsa aloittaneet voivat noudattaa joko näitä tutkintovaatimuksia tai 1.8.2010 voimaan tulleita tutkintovaatimuksia.

## Luonnontieteiden kandidaatti / Valtiotieteiden kandidaatti

Tilastotieteen koulutusohjelmassa suoritetaan valtiotieteellisessä tiedekunnassa valtiotieteiden kandidaatin (VTK) tutkinto, matemaattis-luonnontieteellisessä tiedekunnassa luonnontieteiden kandidaatin (LuK) tutkinto.

Seuraavassa esitetään pääaineopintojen, sivuaineopintojen ja yleisopintojen vähimmäislaajuudet. Kandidaatintutkinnon opintojen laajuuden pitää olla yhteensä vähintään 180 op, vaikka pääaineopintojen, sivuaineopintojen ja muiden opintojen vaadittujen vähimmäislaajuuksien summa jäisikin tämän alle.

### Pääaineopinnot

#### **57505 / 78015 TILASTOTIETEEN PERUSOPINNOT (25 op)**

Perusopinnot ovat vähintään 25 op, joista pakollisia (P) opintoja 10 op. Perusopintoja voi suorittaa vaadittua vähimmäismäärää enemmän, mutta ylimääräiset perusopinnot eivät pienennä aineopintojen 55 op:n vaatimusta. Perusopintoihin voidaan sisällyttää myös aineopintojen kurseja.

#### **78300 Minustako tilastotieteilijä? 5 op**

Kurssi on orientoiva perehdyttämiskäsi jaksoksi tilastotieteen uusille pääaineopiskelijoille. Sen aikana tarkastellaan käytännönläheisesti erilaisia tilastotieteilijän toimenkuvaan kuuluvia aiheita, tilastotieteen merkitystä yhteiskunnassa sekä tilastotieteen historiaa. Mitä tilastotieteilijä työssään tekee? Miten tilastoaineistoja kerätään? Miten tilastotiede liittyy kyselylomakkeisiin? Missä tilastollisia menetelmiä sovelletaan? Mitä hyvän tilastollisen kuvan piirtäminen edellyttää? Miten tilastollisia tunnuslukuja tulkitaan? Miten hahmotetaan, tiivistetään ja visualisoidaan tilastoaineistoihin kätkeytyvät tiedot erilaisten ilmiöiden vaihtelussa ja riippuvuuksissa? Mikä on tietojenkäsittelyn ja ohjelmistojen rooli tilastotieteessä? Mistä nämä asiat juontavat juurensa? Kuka kehitti, mitä ja milloin?

#### **57045 Johdatus todennäköisyyslaskentaan 5 op (P)**

Kurssilla esitellään todennäköisyyden ja satunnaismuuttujan käsitteet, tärkeimmät diskreetit ja jatkuvat jakaumat sekä niiden tunnusluvut. Kurssilla käsitellään myös kombinatoriikkaa sekä suurten lukujen lakia ja keskeistä raja-arvolausetta.

#### **57046 Johdatus tilastolliseen päättelyyn, 5 op (P)**

Kurssilla perehdytään tilastollisen mallin käsitteeseen ja siihen liittyvän päättelyn kysymyksenasetteluihin perinteisen frekventistisen paradigman näkökulmasta. Pääpaino on parametristimoinnissa. Lisäksi käsitellään mm. hypoteesintestausta lähinnä normaalijakauman tapauksessa.

#### **57703 Data-analyysi R-ohjelmistolla, 5 op**

Kurssilla perehdytään data-analyysin perusteisiin käyttäen avoimen lähdekoodin R-ohjelmistoa (<http://www.r-project.org>), joka on maailmanlaajuisesti merkittävä data-analyysin ja tilastollisen mallinnuksen työkalu. R on sekä matemaattisiin sovelluksiin suunnattu ohjelmointikieli että laajennettavissa oleva laskentaympäristö, jonka käyttöä erityissovelluksissa tukee suuri määrä näitä varten luotuja ohjelmakirjastoja. Kurssilla tarkastellaan erilaisia havaintoaineistoja sekä graafisesti että tilastollisten menetelmien avulla.

#### **78116 Data-analyysi (osa 1), 5 op**

#### **78132 Data-analyysi (osa 2), 5 op**

Kurssilla opiskellaan data-analyysin perusteet ja perehdytään jonkin tilastollisen tietojenkäsittelyjärjestelmän itsenäiseen käyttöön analysoimalla kyselyaineistoja ja muita yhteiskuntatieteellisiä aineistoja. Kurssi rakentuu kahdesta periodin pituisesta osasta. Osa 1 koostuu luennoista ja harjoituksista, ja se suoritetaan työraporteilla sekä itse- ja vertaisarvioinneilla. Osassa 2 tehdään ohjatusti harjoitustyö, jossa sovelletaan osassa 1 opittuja asioita uuteen aineistoon.

### 57517 / 78033 TILASTOTIETEEN AINEOPINNOT (55 op)

Aineopinnot ovat vähintään 55 op. Aineopintoihin voi sisällyttää myös syventävien opintojen kurseja (esim. maisteriopinnot eri linjojen kurseja).

Matemaattis-luonnontieteellisessä pakollisia aineopintosuorituksia on 32 op (Proseminaarista 1 op kirjataan tilastotieteen aineopintoihin ja 3 op äidinkielen opintoihin). Valtiotieteellisessä pakollisia aineopintosuorituksia on 35 op (Proseminaarista kirjataan tilastotieteen aineopintoihin).

#### **57705 Todennäköisyyslaskenta, 10 op (P)**

Kurssilla opetellaan tilastollisen mallintamisen kannalta keskeisiä osia todennäköisyyslaskentaa, kuten esimerkiksi moniulotteisten jakaumien käsittelyä. Opittavia valmiuksia tarvitaan useilla myöhemmillä tilastotieteen kursseilla. Esitietona edellytetään perustietoja todennäköisyyslaskennasta (esim. kurssia Johdatus todennäköisyyslaskentaan).

#### **57701 Tilastollinen päättely, 10 op (P)**

Kurssilla esitetään uskottavuuspäättelyn perusteet, joita tarvitaan tilastotieteen erikoiskurssien opiskelua varten. Kurssi edellyttää Todennäköisyyslaskennan kurssin tiedot. Keskeisiä aiheita ovat suurimman uskottavuuden periaate, hypoteesien testaus ja luottamusvälien muodostus.

#### **78134 Lineaariset mallit, 5 op (P)**

Kurssilla esitetään lineaaristen mallien perusteoria ja lineaaristen mallien rakentamisen pääperiaatteet. Kurssi on tarkoitettu suorittamaan heti Tilastollisen päättelyn kurssin jälkeen.

#### **78172 Proseminaarit, 1-4 op (P)**

Proseminaarissa keskitytään tieteellisen kirjoittamisen opetteluun. Matemaattis-luonnontieteellisessä Proseminaarista kirjataan 1 op aineopinnoiksi ja 3 op äidinkielen opinnoiksi. Valtiotieteellisessä Proseminaarista kirjataan 4 op:n laajuisena aineopinnoiksi. Proseminaarit on tarkoitettu vain pääaineopiskelijoille. Edeltävät opinnot: Todennäköisyyslaskennan, Tilastollisen päättelyn ja Lineaaristen mallien kurssit.

#### **78431 (valt.) / 57595 (ml) Kandidaatintutkielma, 6 op (P) (ja kypsyysnäyte, 0 op)**

Kandidaatintutkielma kirjoitetaan Proseminaarin yhteydessä. Opiskelijan kypsyysnäytteen tarkastaa äidinkielen opettaja.

#### **57753 Bayes-päättely, 5op**

Perustiedot Bayes-lähestymistavan soveltamisesta tilastolliseen päättelyyn. Kurssi edellytetään esitietona useilla muilla kursseilla. Keskeistä sisältöä on posteriorin määrääntyminen uskottavuusfunktioista sekä priorijakaumasta. Posterioripäättelyä tarkastellaan sekä konjugaattiperheiden tapauksessa että yleisemmissä tilanteissa,

joissa joudutaan soveltamaan laskennallisia menetelmiä.

### **57xxx Data-analyysi SAS-ohjelmistolla, 5op**

Perehtyminen tilastollisten menetelmien hyödyntämiseen SAS-ohjelmiston komentokielen ja makrojen avulla. SAS on eräs maailman johtavista ohjelmistoista tietokantojen ja data-analytiikan alueella. Sitä käytetään erityisen laajasti mm bio- ja lääketieteen tutkimus- ja kehitystyössä.

### **Valinnaisia aineopintojen kursseja niin että aineopintojen minimilaaajuus 55 op täyttyy:**

Valinnaiset kurssit perehdyttävät opiskelijan hänen suuntautumistaan vastaavaan tilastotieteen erikoisalaan. Valinnaisia kursseja on suoritettava niin paljon, että kandidaatin tutkinnossa vaadittavien pääaineopintojen vähimmäismäärä täyttyy. Suositeltavia kursseja ovat aineopintojen Bayes-päätely sekä matematiikasta Stokastiset prosessit ja Todennäköisyysteoria. Sopimuksen mukaan voidaan suorittaa esimerkiksi kirjatentteinä myös muita kursseja. Valinnaisten kurssien mahdolliset harjoitustyöt voidaan laajentaa kandidaatin tutkielmaksi. Aineopintojen erikoiskursseiksi soveltuvat esimerkiksi seuraavat syventävien opintojen kurssit:

- 57733 Computational statistics, 8 op
- 78144 Ei-parametriset ja robustit menetelmät, 6-8 op
- 57748 Genetic analysis and molecular evolution, 4-6 op
- 78126 Lineaaristen mallien sovellukset, 6-8 op
- 57734 Modelling inheritance in pedigrees and populations, 6-10 op
- 78145 Monimuuttujamenetelmät, 6-8 op
- 78405 Otanta-aineiston analyysi, 6-8 op
- 78143 Otantamenetelmät, 6-8 op
- 57729 Phylogeny inference and data analysis, 4-6 op,
- 78180 Regressioanalyysin jatkokurssi, 6-10 op
- 57770 Stationaariset aikasarjat, 5-10 op
- 57755 Statistical methods in gene mapping, 6-10op
- 57739 Statistical methods in medicine and epidemiology, 6-10 op
- 78187 Survey-metodiikka, 8-10 op

### **Sivuaineopinnot**

Sivuaineopintoja tulee suorittaa vähintään 55 op, jotka koostuvat matematiikan 30 op laajuisesta kokonaisuudesta (vaatimukset alla) sekä yhden tai useamman muun sivuaineen vähintään 25 op:n laajuisista opintokokonaisuuksista. Opiskelija voi myös suorittaa kaikki sivuaineopintonsa matematiikassa, jolloin matematiikan opintokokonaisuuden tulee olla vähintään 55 op laajuinen.

Matematiikasta vaaditaan vähintään 30 op laajuinen sivuaineopintokokonaisuus, joka sisältää ainakin seuraavat pakolliset kurssit:

- 57016 Analyysi I, 10 op tai 78121 Matemaattisen analyysin kurssi, 10 op
- 57017 Analyysi II, 10 op tai 78122 Matemaattisen analyysin jatkokurssi, 10 op
- 57043 Lineaarialgebra ja matriisilaskenta I, 5 op
- 57047 Lineaarialgebra ja matriisilaskenta II, 5 op.

Muista matematiikan kursseista erityisesti Stokastiset prosessit, Stokastinen analyysi ja Todennäköisyysteoria ovat sopivia käytettäväksi sivuaineopinnoissa. Ne voi sisällyttää myös tilastotieteen pääaineopintoihin. Muita sivuaineita voi valita varsin vapaasti oman kiinnostuksen mukaan omasta tai muista tiedekunnista niiden asettamien opinto-oikeusrajoitusten puitteissa. Sopivia sivuaineita matematiikan ohella ovat esimerkiksi tietojenkäsittelytiede (ohjelmoinnin perusteet, tietokannat), biotieteet, kansantaloustiede, kansanterveystiede, psykologia ja sosiologia.

### **57550 Muut opinnot**

Pakolliset muut opinnot (17-19 op):

- Henkilökohtainen opintosuunnitelma, 1 op.
- Äidinkieli, 3 op. Matemaattis-luonnontieteellisessä suoritetaan Proseminaarilla; valtiotieteellisessä Äidinkielen kirjallisen käytön kursseilla.
- Toinen kotimainen kieli, 3-4 op. Laajuus on matemaattis-luonnontieteellisessä 3 op ja valtiotieteellisessä 4 op.
- Vieras kieli, 3-4 op. Laajuus on matemaattis-luonnontieteellisessä 4 op ja valtiotieteellisessä 3 op.
- Tieto- ja viestintätekniikan (TVT) opinnot, 5 op, joista 3 op muodostuu pakollisesta TVT-ajokortista. Laitoksen Latex-kurssi sopii erityisen luontevasti tähän; valtiotieteellisessä myös oman tiedekunnan tilasto-ohjelmistokurssit. Yliopiston tietotekniikkaosaston ja maatalous-metsätieteellisen tiedekunnan tilasto-ohjelmistokurssit käyvät myös. Huom! Valtiotieteellisen tiedekunnan opiskelijan tulee pyytää kurssi kirjattavaksi TVT:n valinnaiseksi täydennyskurssiksi.
- Työelämäorientaatio, 1-3 op (työn soveltuvuuden ja vaativuuden mukaan).

## **Filosofian maisteri / Valtiotieteiden maisteri**

Tilastotieteen koulutusohjelmassa suoritetaan valtiotieteellisessä tiedekunnassa valtiotieteiden maisterin (VTM) tutkinto, matemaattis-luonnontieteellisessä tiedekunnassa filosofian maisterin (FM) tutkinto. Maisterin tutkinnon edellytyksenä on soveltuva kandidaatin tutkinto. Maisterin tutkinnossa tilastotiede pääaineena valmistutaneen erikoistuneena jollekin seuraavista linjoista:

Aikasarja-analyysin ja ekonometrian linja  
Biometrian ja bioinformatiikan linja  
Psykometriikan linja  
Yhteiskuntatilastotieteen linja  
Yleisen tilastotieteen linja  
Tilastollisen koneoppimisen linja

Maisterin tutkinnossa on suoritettava pääaineen syventävät opinnot, joiden vaatimukset esitetään alla. Pakollinen on myös henkilökohtainen opintosuunnitelma (1 op). Maisterin tutkintoon tulee sisällyttää myös työharjoittelu: matemaattis-luonnontieteellisessä tiedekunnassa väh. 1 op, valtiotieteellisessä väh. 5 op.

Maisterin tutkinnossa opintojen laajuuden tulee olla yhteensä vähintään 120 op. Tarvittaessa voi pääaineen syventävät opinnot suorittaa vähimmäisvaatimuksia laajempina tai täydentää kandidaatin tutkinnon sivuaineopintoja.

### **TILASTOTIETEEN SYVENTÄVÄT OPINNOT (85 op)**

**Pakolliset kurssit on merkitty P-kirjaimella.**

#### **57713 Tilastollisen päättelyn jatkokurssi, 5 op (P)**

Kurssilla syvennetään ja laajennetaan aineopintojen tilastollista päättelyä ja sen teoriaa.

#### **78185 Yleistetyt lineaariset mallit (osa 1), 5 op (P)**

#### **78186 Yleistetyt lineaariset mallit (osa 2), 3 op**

Kurssilla esitetään yleistettyjen lineaaristen mallien perusteoria ja sen erikoistapauksina erityyppisille vastemuuttujille soveltuvia malleja, kuten logistinen ja log-lineaarinen regressiomalli. Kurssin ensimmäinen osa (78185, 5 op) on pakollinen; sitä voi täydentää vapaaehtoisella harjoitustyöllä (78186, 3 op). Kurssi edellyttää Lineaaristen mallien kurssin tiedot. Kirjallisuus: McCullagh, P. and Nelder, J. A. (1989) Generalized Linear Models (Second edition). London: Chapman & Hall. Dobson, A. J. and Barnett, A. (2008) An Introduction to Generalized Linear Models (Third edition). London: Chapman & Hall, tai muu tentaattorin tai luennoijan osoittama materiaali.

#### **78172 Tutkielmaseminaari, 2 op (P)**

Osallistuminen seminaariin, esitelmien pito ja opponointi.

#### **784130 (valt.)/57620 (ml) Pro gradu -tutkielma, 40 op ja kypsyysnäyte, 0 op (P)**

Pro gradu -tutkielman laatiminen.

### **Valinnaisia syventäviä kursseja niin että syventävien opintojen minimilaaajuus 85 op täyttyy.**

Valinnaiset kurssit valitaan linjavälän mukaan. Erikoiskurssit perustuvat luentomateriaaliin tai luennoijan kanssa sovittavaan kirjallisuuteen.

## **57560 / 78080 Aikasarja-analyysin ja ekonometrian linja**

Linjan ydin muodostuu kursseista Stationaariset, Moniulotteiset ja Epästationaariset aikasarjat. Vaihtoehtoisesti linjalla voi keskittyä ekonometriaan eli taloudellisten muuttujien yleiseen tilastotieteelliseen analyysiin. Keskeiset kurssit ovat tällöin Regressioanalyysin jatkokurssi ja Ekonometrian kurssi. Kaikkia mainittuja kursseja ei luennoida säännöllisesti. Muita kursseja luennoidaan ajoittain (esimerkiksi Epälineaarinen aikasarja-analyysi ja Epälineaarinen regressio). Bayes-päätely kurssi on suositeltava esitietovaatimus, samoin Todennäköisyysteoria ja Stokastiset prosessit.

#### **57770 Stationaariset aikasarjat, 5-10 op**

Aikasarja-analyysin peruskäsitteet, empiirisesti paljon käytetty ARMA-mallin ominaisuudet, valinta, estimointi, testaus ja ennustaminen sekä ehdollisen heteroskedastisuuden mallintaminen kuuluvat opiskeltaviin asioihin. Edeltävät opinnot: Tilastollinen päätely ja Lineaariset mallit. Kurssi sopii kaikille tilastotieteen opiskelijoille linjavalmista riippumatta. Kirjallisuus: James Hamiltonin Time Series Analysis, luvut 1-5 ja 21 (aineopinnot 8 op) tai muu tentaattorin tai luennonjoijan osoittama materiaali. Kurssi voidaan täydentää harjoitustyöllä syventäviksi opinnoiksi (10 op).

#### **57771 Moniulotteiset aikasarjat, 5-10 op**

Aiheita ovat monen aikasarjan yhtäaikainen mallintaminen, asymptoottisen teorian soveltaminen ja empiirisesti suositut VAR-mallit. Edeltävät opinnot: Tilastollinen päätely, Lineaariset mallit ja perustiedot stationaarista aikasarjoista. Kirjallisuus: James Hamiltonin Time Series Analysis, luvut 7-11 (syventävät opinnot; 10 op) tai muu tentaattorin tai luennonjoijan osoittama materiaali.

#### **57772 Epästationaariset aikasarjat, 5-10 op**

Monien aikasarjojen keskeinen piirre on trendimäisyys. Sen huomioiminen on mullistanut taloudellisten aikasarjojen analyysin teorian viime vuosikymmeninä. Kurssilla opiskellaan aikasarjaekonometrian nykyisen valtavirran tekniikoita. Edeltävät opinnot: Tilastollinen päätely, Lineaariset mallit, Stationaariset aikasarjat ja Moniulotteiset aikasarjat. Kirjallisuus: James Hamiltonin Time Series Analysis, luvut 15-20 (syventävät opinnot; 10 op) tai muu tentaattorin tai luennonjoijan osoittama materiaali.

#### **78180 Regressioanalyysin jatkokurssi, 6-10 op**

Kurssilla syvennetään lineaarisen regressiomallin tuntemusta. Kurssi sopii kaikille tilastotieteen opiskelijoille linjavalmista riippumatta. Edeltävät opinnot: Tilastollinen päätely ja Lineaariset mallit. Kirjallisuus: Russell Davidsonin ja James MacKinnonin Econometric Theory and Methods, luvut 1-6 (aineopinnot, 10 op) tai 1-7 (syventävät opinnot; 10 op) tai muu tentaattorin tai luennonjoijan osoittama materiaali. Kurssi voidaan täydentää vaihtoehtoisesti harjoitustyöllä syventäviksi opinnoiksi.

#### **78173 Ekonometria, 10 op**

Aiheena ovat sekä klassiset että modernit ekonometrian menetelmät. Edeltävät opinnot: Tilastollinen päätely, Lineaariset mallit ja Regressioanalyysin jatkokurssi. Kirjallisuus: Russell Davidsonin ja James MacKinnonin Econometric Theory and Methods, luvut 7 (mikäli ei aiemmin suoritettu), 8-12 ja 15 (syventävät opinnot; 10 op) tai muu tentaattorin tai luennonjoijan osoittama materiaali.

### **57561 / 78081 Biometrian ja bioinformatiikan linja**

Bio- ja lääketieteiden kvantitatiivinen data-analyysi ja tilastollinen menetelmäkehitys näiden sovellusalueiden tarpeisiin. Opetus on englanninkielistä, koska useimmat kurssit kuuluvat kansainvälisiin maisteriohjelmiin, Master's Degree Programme in Bayesian Statistics and Decision Analysis (EuroBayes) ja Master's Degree Programme in Bioinformatics (MBI). Opetusohjelmassa voi olla muitakin kuin alla mainittuja kursseja. Osa kursseista järjestetään vuosittain, osa keskimäärin joka toinen vuosi.

#### **57059 Markovian modelling and Bayesian learning, 5 op**

Kurssilla esitellään aluksi erilaisten Markov-mallien toimintaperiaatteet. Lähes kaikki moderni todennäköisyyspohjainen mallinnus hyödyntää markovilaisia riippumattomuusoletuksia eri muodoissa. Kurssilla perehdytään mallien rakentamiseen erilaisille ilmiöille tavallisten ja ns. piilo-Markovin oletusten perusteella. Bayes-oppimisen perusteet ja niiden hyödyntäminen sovelluksissa esitellään kurssin jälkimmäisessä osassa. Esitietoina edellytetään Bayes-päätely (tai vastaavat tiedot).

#### **57733 Computational statistics, 8 op**

Kurssilla käsitellään tiettyjä moderneja tilastomenetelmiä, jotka vaativat paljon laskentaa. Kurssin aiheita ovat esimerkiksi todennäköisyysjakaumien simulointi, Monte Carlo -integrointi sekä EM-algoritmi. Yhden tärkeän aihepiirin muodostavat MCMC-menetelmät, jotka ovat erityisesti Bayes-päätelyyn soveltuvia laskentamenetelmiä, jotka perustuvat Markovin ketjujen simulointiin. Esitietoina edellytetään Bayes-päätely (tai vastaavat tiedot).

#### **57744 Bayesian theory with applications, 5 op**

Bayesläistä lähestymistapaa käytetään nykyään laajalti todennäköisyysmallinnuksessa useimmilla tieteenaloilla ja se edustaa myös hyvin tärkeää työkalua insinööritieteiden teknologioissa sovelluksissa. Kurssilla perehdytään bayesläisen mallinnuksen teoreettisiin perusteisiin sekä tarkastellaan yksityiskohtaisesti miten ja miksi sen avulla pystytään tehokkaasti ratkaisemaan erilaisia sovellusongelmia. Esitietoina edellytetään Bayes-päätely (tai vastaavat tiedot).

#### **57745 WinBUGS/OpenBUGS with applications, 6-10 op**

Kurssilla perehdytään WinBUGS/OpenBUGS/JAGS ohjelmistoihin, jotka edustavat laajalti käytössä olevia stokastisen bayes-laskennan työkaluja. Sovellusesimerkeissä perehdytään mm riskien arviointiin elintarvikevalvonnassa, mikrobiologiassa, diagnostiikassa, eläinlääketieteessä ja epidemiologiassa. Kurssin keskeiset tavoitteet ovat oppia miten hierarkkisia todennäköisyysmalleja määritellään syklittömien suunnattujen graafien ja BUGSin formaalin syntaksin avulla, simulaatiotulosten analysointi sekä ohjelmien tiettyjen erikoisominaisuuksien hyödyntäminen. Esitietoina edellytetään Bayes-päätely (tai vastaavat tiedot).

#### **57742 Hierarchical models, 6-8 op**

Tilastolliset tutkimusaineistot voidaan usein luontevasti tulkita kerroksittaisena rakenteena, jossa kukin karkeamman luokittelutason ryhmä voidaan osittaa useaan hienomman jaon mukaiseen alaryhmään. Hierarkkinen mallintaminen on yleinen tapa hallita tällaisia tutkimusaineistoja. Hierarkkisten mallien kehikossa voidaan esittää useita mm. biometriassa, psykometriassa ja ekonometriassa käytettyjä malleja, jotka ovat yleisen teorian erikoistapauksia.

#### **57740 Nonparametric statistics, 6 op**

Parametrisen ja epäparametrisen tilastollisen päätelyn käsitteet. Empiiriset jakaumat, Kolmogorov-Smirnov -testi, luottamusvälit ja kvanttilit, merkkitesti, Wilcoxon-Mann-Whitney -testi, Kruskal-Wallis -testi, riippumattomuustestit, korrelaation epäparametriset mitat, asymptoottinen tehokkuus, epäparametrisuus Bayes-kehikossa.

#### **57746 Longitudinal data-analysis, 6-10 op**

Kurssilla käsitellään lineaaristen ja yleistettyjen lineaaristen mallien soveltamista tilanteissa, jossa samoista yksilöistä tehdään mittauksia useassa eri aikapisteessä. Tämä mahdollistaa yksilössä ajan tai esimerkiksi iän mukana tapahtuvien muutosten erottamisen yksilöistä eroista. Kurssilla käsiteltäviä asioita ovat mm. satunnaistekijämallit, marginaalimallit ja latentit prosessimallit. Kurssi keskittyy bio- ja lääketieteen sovelluksiin.

#### **57741 Event-history analysis, 6-8 op**

Elinäika-analyysissa ja tapahtumahistoria-analyysissa tarkastellaan todennäköisyysmallien avulla satunnaisten tapahtumien toteutumista tietyn seuranta-ajan kuluessa. Yksinkertaisimmillaan tutkitaan yhden tapahtuman toteutumisriskiä, mutta yleisemmin voidaan mallittaa useiden toisistaan riippuvien tapahtumien muodostamien ketjujen toteutumastodennäköisyyksiä. Tapahtumahistoria-analyysin matemaattisena perustana on pisteprosessien, erityisesti laskuri-prosessien teoria.

#### **57739 Statistical methods in medicine and epidemiology, 6-10 op**

Kurssilla tutustutaan tautien yleisyyden mittaamisessa sekä terveyteen liittyvien altistevaikutusten arvioimisessa käytettyihin tutkimusasetelmiin ja tilastollisiin menetelmiin. Perusaineksina ovat erilaiset todennäköisyysmallit ja niiden käyttö epidemiologian ongelmissa. Kurssilla käsitellään mm. kohortti- ja tapausverkkitutkimusten suunnittelua ja tilastollista analyysia.

#### **57747 Mathematical models of molecular evolution, 5 op**

Molekyyli evoluution mallit ovat tärkeässä roolissa monilla tutkimusaloilla joilla hyödynnetään molekyylibiologiaa ja bioinformatiikkaa. Kurssilla esitellään molekyyli evoluution matemaattisten mallien keskeiset ominaisuudet ja perehdytään menetelmiin joilla malleja sovitetaan havaintoaineistoihin. Esitietoina edellytetään Bayes-päätely (tai vastaavat tiedot).

#### **57748 Genetic analysis and molecular evolution, 4-6 op**

Kurssi on räätälöity Bioinformatiikan maisteriohjelman opiskelijoille, joilla ei ole bio-opintojen taustaa. DNA:n ja proteiinien ominaisuudet tilastollisina ja laskennallisina haasteina, genomien osien molekulaarinen rakentuminen sekä keskeisimmät polymorfismit ja niiden tilastollisen analyysin perusteet, eliökunta ja sen evoluutiohistoria molekyylien valossa, sekvenssitietokantatyöskentelyn perusteet. Kurssi on avoin myös muille kuin maisteriohjelman opiskelijoille.

#### **57780 Molecular genetics reading group, 2op**

Lukupiiri on räätälöity Bioinformatiikan maisteriohjelman opiskelijoille, joilla ei ole bio-opintojen tasua. Lukupiiri ja tentti: Griffiths et al. Genetic analysis, luvut 1,7,8,9,13. Lukupiiri on avoin myös muille kuin maisteriohjelman opiskelijoille.

#### **57730 Sequence evolution from molecules to genomes, 4-6 op**

Kurssilla perehdytään sekvenssitietokantoihin ja niistä kerättävän, geenejä ja muita genomien rakenteita, koskevan datan analysointimenetelmiin.

#### **57734 Modelling inheritance in pedigrees and populations, 4-6 op**

Periytyminen sukupuu-tasolla ja pidemmällä aikavälillä populaatiogenetiikan koalesenssiteorian kautta simulointien ja haplotyyppi-tietokannan materiaalia käyttäen. Käytännön perehtyminen PHASE- ja HapView-ohjelmistoihin sekä keskeisiin datan simulointiohjelmistoihin. Esitietoina edellytetään Bayes-päätely (tai vastaavat tiedot).

#### **57729 Phylogeny inference and data analysis, 4-6 op**

Kurssi painottuu tietokone-luokassa tapahtuvaan sekvenssitietokannoista kerättävän datan analysointiin keskeisimmillä fylogenia-ohjelmistoilla. Painopiste on Bayes-menettelyssä. Esitietoina edellytetään Bayes-päätely (tai vastaavat tiedot).

#### **57755 Statistical methods in gene mapping, 4-6 op**

Kurssilla perehdytään geneettisen epidemiologian ja geenikartoituksen tilastollisiin menetelmiin, erityisesti kytkentä- ja assosiaatioanalyysiin. Esitietoina edellytetään Bayes-päätely (tai vastaavat tiedot).

#### **57743 Genome-wide association mapping, 4-6 op**

Kurssilla perehdytään genomilajuisiin tilastollisiin assosiaatiokartoitusmenetelmiin, toistettavuuteen, sekoitaviin tekijöihin ja meta-analyysiin. Esitietovaatimuksina todennäköisyyslaskennan ja uskottavuuspäätelyn perusteet. Molekyylibiologian perusteita ei edellytetä.

### **78083 Psykometriikan linja**

Linjalla keskitytään käyttäytymistieteissä tarvittavien tilastollisten menetelmien opiskeluun. Suositeltavia valinnaisia erikoiskursseja on lueteltu alla. Soveltuvia erikoiskursseja ovat myös yhteiskuntatieteiden linjan kurssit Survey-metodiikka, Otantamenetelmät ja Otanta-aineiston analyysi.

#### **78164 Mittaaminen ja mittausvirheiden tilastollinen hallinta, 6-8 op**

Kurssilla perehdytään mittausvirheiden vaikutuksen arviointiin sekä tässä tarvittaviin tilastollisiin menetelmiin erityisesti yhteiskunta- ja käyttäytymistieteissä. Tarkasteltavia aiheita ovat mittaamisen tasot, niiden vaikutus analyysimenetelmiin, mittaamisen laadun arviointi (reliabiliteetti ja validiteetti), mittausmalli ja rakennevaliditeetti, mittauksen reliabiliteetin arviointi, mittausasteikkojen laatiminen ja ennustevaliditeetti. Kirjallisuus: Duane Alwin: Margins of Error, a Study of Reliability in Survey Measurement; Wayne Fuller: Measurement Error Models; Kimmo Vehkalahti: Kyselytutkimuksen mittarit ja menetelmät (luku 4), tai muu tentaattorin tai luennoijan osoittama materiaali.

#### **78145 Monimuuttujamenetelmät, 6-8 op**

Kurssilla perehdytään sovelluspainotteisesti usean muuttujan data-analyysiin sekä aineistojen moniulotteiseen kuvailuun ja mallintamiseen erityisesti yhteiskunta- ja käyttäytymistieteissä. Tarkasteltavia menetelmiä ovat faktorianalyysi, erilaiset ryhmittelymenetelmät, erotteluanalyysi, moniulotteinen skaalaus sekä korrespondenssianalyysi. Kirjallisuus: Richard Johnson & Dean Wichern: Applied Multivariate Statistical Analysis; Brian Everitt: Multivariable Modeling and Multivariate Analysis for the Behavioral Sciences; Michael Greenacre: Biplots in Practice; Robert Cudeck & Robert MacCallum: Factor Analysis at 100, Historical Developments and Future; Kimmo Vehkalahti: Kyselytutkimuksen mittarit ja menetelmät (luvut 4, 6 ja 7); Seppo Mustonen: Tilastolliset monimuuttujamenetelmät, tai muu tentaattorin tai luennoijan osoittama materiaali.

#### **78177 Rakenneyhtälömallit, 6-8 op**

Kurssilla tarkastellaan useilla aloilla, mm. yhteiskunta- ja käyttäytymistieteissä, sovellettavia rakenneyhtälömalleja, jotka ovat yleistyksiä tavallisista faktori- ja regressioanalyysistä. Olennainen osa mallien rakentelua ja testausta on jonkin soveltuvan ohjelmiston (mm. Mplus, Amos, Lisrel, Calis) opettelu. Kenneth Bollen: Structural Equations with Latent Variables; Randall Schumacker & Richard Lomax: A Beginner's Guide to Structural Equation Modeling; Mikko Ketokivi: Tilastollinen päätely ja tieteellinen argumentointi, tai muu tentaattorin tai luennoijan osoittama materiaali.

#### **78178 Monitasomallit, 6-8 op**

Kurssilla tarkastellaan useilla aloilla, mm. yhteiskunta- ja käyttäytymistieteissä sekä organisaatiotutkimuksessa sovellettavia malleja, jotka on tarkoitettu hierarkkisesti rakentuneiden aineistojen analyysiin. Mallit ovat yleistyksiä tavallisista regressio- ja varianssianalyysistä ja niitä kutsutaan monitasomalleiksi, lineaarisiksi sekamalleiksi tai hierarkkiseksi malleiksi. Olennainen osa mallien rakentelua ja testausta on jonkin soveltuvan ohjelmiston (mm. Mplus, MLwiN, HLM, SAS/Mixed, Glimmix) opettelu. Kirjallisuus: Harvey Goldstein: Multilevel Statistical Models (3. tai 4. painos), Hox, J. J. (1995) Applied Multilevel Analysis. Amsterdam: TT-Publikaties, <http://www.sozioogie.uni-halle.de/langer/multilevel/books/hox95m1a.pdf> ja Snijders, T. and Bosker, R. (2002) Multilevel Analysis. An Introduction to Basic and Advanced Multilevel Modeling. London: Sage Publications, Ronald Heck & Scott Thomas: An Introduction to Multilevel Modeling Techniques, tai muu tentaattorin tai luennoijan osoittama materiaali.

#### **78126 Lineaaristen mallien sovellukset, 6-8 op**

Sovelluspainotteinen kurssi keskittyy regressioanalyysin käyttöön yhteiskunta- ja käyttäytymistieteellisessä tutkimuksessa. Tavoitteena on oppia analysoimaan käytännön tutkimusaineistoja, arvioimaan tilastollista mallia havaintoaineiston informaation tiivistäjänä sekä esittämään tutkimusongelmat ja tulokset selkeästi graafisia menetelmiä ja vastaavia tilastollisia esitystapoja käyttäen. Kirjallisuus: Sanford Weisberg: Applied Linear Regression; Dennis Cook & Sanford Weisberg: Applied Regression Including Computing and Graphics; Kimmo Vehkalahti: Kyselytutkimuksen mittarit ja menetelmät (luku 5), tai muu tentaattorin tai luennoijan osoittama materiaali.

#### **78144 Ei-parametriset ja robustit menetelmät, 6-8 op**

Kurssilla syvennycyään merkkeihin ja järjestyslukuihin perustuviin tilastollisiin testeihin, estimaatteihin ja luottamusväleihin. Käsiteltäviä aiheita ovat mm. keskiarvotyyppiset menetelmät, mediaanityyppiset menetelmät, Hodges-Lehmann-tyyppiset menetelmät, robustisuus ja menetelmien ominaisuuksien tutkiminen tietokonesimulointien avulla. Kirjallisuus: Hettmansperger, T. P. (1984). Statistical Inference Based on Ranks. New York: Wiley. Huber, P. J. (1981). Robust Statistics. New York: Wiley, tai muu tentaattorin tai luennoijan osoittama materiaali.

### **57562 / 78082 Yhteiskuntatieteiden linja**

Linjalla keskitytään yhteiskuntatutkimuksessa ja tilastoinnissa tarvittavien tilastollisten menetelmien opiskeluun. Linjan pakollisia opintoja ovat Survey-metodiikka ja Otantamenetelmät. Suositeltavia valinnaisia erikoiskursseja on lueteltu alla. Soveltuvia erikoiskursseja ovat myös biometriin linjan kurssit Hierarchical models, Event-history analysis ja Longitudinal data-analysis sekä aikasarja-analyysin ja ekonometrian linjan kurssit. Teoreettisemmin suuntautuneille sopivia kursseja ovat matematiikasta esimerkiksi todennäköisysteorian, matriisilaskennan ja stokastisten prosessien kurssit.

#### **78187 Survey-metodiikka, 8-10 op (P)**

Kurssilla esitellään empiirisen kvantitatiivisen yhteiskuntatutkimuksen eli ns. surveyn keskeiset käsitteet, menetelmät ja suoritusvaiheet. Tarkasteltavia aiheita ovat tutkimusasetelmat, otannan perusteet, aineiston keruu ja lomakesuunnittelu, poikkileikkaus- ja pitkäaikaisaineistot, aineiston puhdistaminen (tilastollinen editointi, imputointi ja uudelleenpainotus), estimoinnin ja analyysin perusteet sekä aineiston jakelu ml. tilastolliset tietosuojamenetelmät. Kirjallisuus: Laaksonen, Seppo (2010) Survey-metodiikka. 214 sivua. Ventus Publishing ApS. Ilmainen nettikirja. <http://bookboon.com/fi/student/statistics/surveymetodiikka>. Leeuw, Edith, Hox, Joop and Dillman, Don (2008) International Handbook of Survey Methodology. 549 pages. Lawrence Erlbaum Associates. Taylor&Francis Group. New York, London, tai muu tentaattorin tai luennoijan osoittama materiaali.

#### **78143 Otantamenetelmät, 6-8 op (P)**

Kurssilla annetaan yleiskuva tilastollisista otantamenetelmistä ja niiden käytöstä eri tieteenalojen empiirisessä tutkimuksessa. Otannan perusmenetelmien lisäksi käsitellään otoskoon määrittelyn perusteita ja lisäinformaation käyttöä otannassa ja estimoinnissa sekä vastaavia tilastollisia ohjelmistoja. Kirjallisuus: Lehtonen, Risto and Pahkinen, Erkki (2004) Practical Methods for Design and Analysis of Complex Surveys. Second Edition. Chichester: Wiley (luvut 1-4), tai muu tentaattorin tai luennoijan osoittama materiaali. Verkkosovellus VLISS: <http://mathstat.helsinki.fi/VLISS/>.

#### **78405 Otanta-aineiston analyysi, 6-8 op**

Kurssilla perehdytään otanta-aineistojen tilastolliseen analyysiin erityisesti tilanteissa, joissa aineisto on kerätty jollakin mutkikkaalla otanta-asetelmalla. Keskiössä ovat menetelmät, joilla otanta-asetelman ominaispiirteitä (ositus, ryvastyminen, painokertoimet) voidaan ottaa huomioon tilastollisen analyysin yhteydessä. Pääpaino on ns. asetelmaperusteisissa (design-based) menetelmissä; lisäksi tarkastellaan suppeasti malliperusteisia menetelmiä (ml. sekamallit). Kirjallisuus: Lehtonen, Risto and Pahkinen, Erkki (2004) Practical Methods for Design and Analysis of Complex Surveys. Second Edition. Chichester: Wiley (luvut 5, 7-9), tai muu tentaattorin tai luennoijan osoittama materiaali. Verkkosovellus VLISS: <http://mathstat.helsinki.fi/VLISS/>.

#### **78189 Pienalue-estimointi, 6-8 op**

Kurssilla perehdytään perusjoukon osajoukkoja koskevan estimoinnin (small area estimation, SAE) teoriaan, malleihin, laskentamenetelmiin ja sovelluksiin. Asetelmaperusteisten malliluvusteisten menetelmien (yleistetyt regressioestimaattorit ja kalibrointimenetelmät) ohella tarkastellaan malliperusteisia menetelmiä (synteettiset, EBLUP- ja EBP-estimaattorit) sekä estimointiin soveltuvia tilastollisia ohjelmistoja. Sovellukset ovat pääasiassa yhteiskuntatieteellisiltä ja terveystieteellisiltä aloilta. Kirjallisuus: Rao, J.N.K. (2003) Small Area Estimation. New York: Wiley, Lehtonen, Risto and Pahkinen, Erkki (2004) Practical Methods for

Design and Analysis of Complex Surveys. Second Edition. Chichester: Wiley (luku 6), Lehtonen, Risto and Veijanen, Ari (2009) Design-based methods of estimation for domains and small areas. In: C. R. Rao and D. Pfeffermann (eds.), Handbook of Statistics, vol. 29B. Sample Surveys: Theory, Methods and Inference. Elsevier, tai muu tentaattorin tai luennoijan osoittama materiaali.

#### **78126 Lineaaristen mallien sovellukset, 6-8 op**

Sovelluslainotteinen kurssi keskittyy regressioanalyysin käyttöön yhteiskunta- ja käyttäytymistieteellisessä tutkimuksessa. Tavoitteena on oppia analysoimaan käytännön tutkimusaineistoja, arvioimaan tilastollista mallia havaintoaineiston informaation tiivistäjänä sekä esittämään tutkimusongelmat ja tulokset selkeästi graafisia menetelmiä ja vastaavia tilastollisia esitystapoja käyttäen. Kirjallisuus: Sanford Weisberg: Applied Linear Regression; Dennis Cook & Sanford Weisberg: Applied Regression Including Computing and Graphics; Kimmo Vehkalahti: Kyselytutkimuksen mittarit ja menetelmät (luku 5), tai muu tentaattorin tai luennoijan osoittama materiaali.

#### **78144 Ei-parametriset ja robustit menetelmät, 6-8 op**

Kurssilla syvennytään merkkeihin ja järjestyslukuihin perustuviin tilastollisiin testeihin, estimaatteihin ja luottamusväleihin. Käsiteltäviä aiheita ovat mm. keskiarvotyypiset menetelmät, mediaanityypiset menetelmät, Hodges-Lehmann-tyyppiset menetelmät, robustisuus ja menetelmien ominaisuuksien tutkiminen tietokonesimulointien avulla. Kirjallisuus: Hettmansperger, T. P. (1984). Statistical Inference Based on Ranks. New York: Wiley. Huber, P. J. (1981). Robust Statistics. New York: Wiley, tai muu tentaattorin tai luennoijan osoittama materiaali.

#### **78137 Robustit regressiomenetelmät, 6-8 op**

Kurssilla perehdytään erilaisiin robusteihin regressiomalleihin. Robustin regressioanalyysin pääasiallisena tarkoituksena on sovittaa havaintoaineistoon malli, joka ei ole herkkä poikkeaville havainnoille. Yleisesti käytössä olevassa lineaarisessa regressioanalyysissä mallin parametrit estimoidaan pienimmän neliosumman menetelmän avulla. Yksinä poikkeava havainto saattaa vaikuttaa merkittävästi pienimmän neliosumman menetelmällä saataviin estimaatteihin ja tuloksena saatava malli ei tällöin edusta suurinta osaa havaintoaineistosta. Rousseeuw, P. J. and Leroy, A. M. (1987). Robust Regression & Outlier Detection. New York: Wiley. Huber, P. J. (1981). Robust Statistics. New York: Wiley, tai muu tentaattorin tai luennoijan osoittama materiaali.

#### **78145 Monimuuttujamenetelmät, 6-8 op**

Kurssilla perehdytään sovelluspainotteisesti usean muuttujan data-analyysiin sekä aineistojen moniulotteiseen kuvailuun ja mallintamiseen erityisesti yhteiskunta- ja käyttäytymistieteissä. Tarkasteltavia menetelmiä ovat faktorianalyysi, erilaiset ryhmittelymenetelmät, erotteluanalyysi, moniulotteinen skaalaus sekä korrespondenssianalyysi. Kirjallisuus: Richard Johnson & Dean Wichern: Applied Multivariate Statistical Analysis; Brian Everitt: Multivariable Modeling and Multivariate Analysis for the Behavioral Sciences; Michael Greenacre: Biplots in Practice; Robert Cudeck & Robert MacCallum: Factor Analysis at 100, Historical Developments and Future; Kimmo Vehkalahti: Kyselytutkimuksen mittarit ja menetelmät (luvut 4, 6 ja 7); Seppo Mustonen: Tilastolliset monimuuttujamenetelmät, tai muu tentaattorin tai luennoijan osoittama materiaali.

#### **78177 Rakenneyhtälömallit, 6-8 op**

Kurssilla tarkastellaan useilla aloilla, mm. yhteiskunta- ja käyttäytymistieteissä, sovellettavia rakenneyhtälömalleja, jotka ovat yleistyksiä tavallisista faktori- ja regressioanalyysistä. Olennainen osa mallien rakentelua ja testausta on jonkin soveltuvan ohjelmiston (mm. Mplus, Amos, Lisrel, Calis) opettelu. Kenneth Bollen: Structural Equations with Latent Variables; Randall Schumacker & Richard Lomax: A Beginner's Guide to Structural Equation Modeling; Mikko Ketokivi: Tilastollinen päättely ja tieteellinen argumentointi, tai muu tentaattorin tai luennoijan osoittama materiaali.

#### **78178 Monitasomallit, 6-8 op**

Kurssilla tarkastellaan useilla aloilla, mm. yhteiskunta- ja käyttäytymistieteissä sekä organisaatiotutkimuksessa sovellettavia malleja, jotka on tarkoitettu hierarkkisesti rakentuneiden aineistojen analyysiin. Mallit ovat yleistyksiä tavallisista regressio- ja varianssianalyysistä ja niitä kutsutaan monitasomalleiksi, lineaarisiksi sekamalleiksi tai hierarkkiksiksi malleiksi. Olennainen osa mallien rakentelua ja testausta on jonkin soveltuvan ohjelmiston (mm. Mplus, MLwiN, HLM, SAS/Mixed, Glimmix) opettelu. Kirjallisuus: Harvey Goldstein: Multilevel Statistical Models (3. tai 4. painos), Hox, J. J. (1995) Applied Multilevel Analysis. Amsterdam: TT-Publikaties. <http://www.soziolegie.uni-halle.de/langer/multilevel/books/hox95m1a.pdf> ja Snijders, T. and Bosker, R. (2002) Multilevel Analysis. An Introduction to Basic and Advanced Multilevel Modeling. London: Sage Publications, Ronald Heck & Scott Thomas: An Introduction to Multilevel Modeling Techniques, tai muu tentaattorin tai luennoijan osoittama materiaali.

### **57xxx / 78084 Yleinen tilastotiede**

Tilastotieteen yleisellä linjalla painotetaan keskeisiä teoreettisia ja laskennallisia valmiuksia ilmiöiden stokastiseen mallinnukseen eri sovellusalueilla. Valinnaiset kurssit voivat koostua erikoistumislinjojen kurssien lisäksi esimerkiksi soveltuvista matematiikan, tietojenkäsittelytieteen ja fyysikan kursseista. Kurssien soveltavuudesta on syytä neuvotella etukäteen tilastotieteen opettajien kanssa. Bayes-päätely kurssi on suositeltava esitietovaatimus, samoin Todennäköisyysteoria ja Stokastiset prosessit. Linjan ydin muodostuu seuraavista kursseista:

#### **57733 Computational statistics, 8 op**

Kurssilla käsitellään tiettyjä moderneja tilastomenetelmiä, jotka vaativat paljon laskentaa. Kurssin aiheita ovat esimerkiksi todennäköisyysjakaumien simulointi, Monte Carlo -integrointi sekä EM-algoritmi. Yhden tärkeän aihepiirin muodostavat MCMC-menetelmät, jotka ovat erityisesti Bayes-päätelyyn soveltuvia laskentamenetelmiä, jotka perustuvat Markovin ketjujen simulointiin. Esitietoina edellytetään Bayes-päätely (tai vastaavat tiedot).

#### **57059 Markovian modelling and Bayesian learning, 5 op**

Kurssilla esitellään aluksi erilaisten Markov-mallien toimintaperiaatteet. Lähes kaikki moderni todennäköisyyspohjainen mallinnus hyödyntää markovilaisia riippumattomuusoletuksia eri muodoissa. Kurssilla perehdytään mallien rakentamiseen erilaisille ilmiöille tavallisten ja ns. piilo-Markovin oletusten perusteella. Bayes-oppimisen perusteet ja niiden hyödyntäminen sovelluksissa esitellään kurssin jälkimmäisessä osassa. Esitietoina edellytetään Bayes-päätely (tai vastaavat tiedot).

#### **57744 Bayesian theory with applications, 5 op**

Bayesläistä lähestymistapaa käytetään nykyään laajalti todennäköisyysmallinnuksessa useimmilla tieteenaloilla ja se edustaa myös hyvin tärkeää työkalua insinöritieteiden teknologisissa sovelluksissa. Kurssilla perehdytään bayesläisen mallinnuksen teoreettisiin perusteisiin sekä tarkastellaan yksityiskohtaisesti miten ja miksi sen avulla pystytään tehokkaasti ratkaisemaan erilaisia sovellusongelmia. Esitietoina edellytetään Bayes-päätely (tai vastaavat tiedot).

#### **57770 Stationaariset aikasarjat, 5-10 op**

Aikasarja-analyysin peruskäsitteet, empiirisesti paljon käytetyn ARMA-mallin ominaisuudet, valinta, estimointi, testaus ja ennustaminen sekä ehdollisen heteroskedastisuuden mallintaminen kuuluvat opiskeltaviin asioihin. Edeltävät opinnot: Tilastollinen päättely ja Lineaariset mallit. Kurssi sopii kaikille tilastotieteen opiskelijoille linjavalmista riippumatta. Kirjallisuus: James Hamiltonin Time Series Analysis, luvut 1-5 ja 21 (aineopinnot 8 op) tai muu tentaattorin tai luennoijan osoittama materiaali. Kurssi voidaan täydentää harjoitustyöllä syventäviksi opinnoiksi (10 op).

#### **78180 Regressioanalyysin jatkokurssi, 6-10 op**

Kurssilla syvennetään lineaarisen regressiomallin tuntemusta. Kurssi sopii kaikille tilastotieteen opiskelijoille linjavalmista riippumatta. Edeltävät opinnot: Tilastollinen päättely ja Lineaariset mallit. Kirjallisuus: Russell Davidsonin ja James MacKinnonin Econometric Theory and Methods, luvut 1-6 (aineopinnot, 10 op) tai 1-7 (syventävät opinnot; 10 op) tai muu tentaattorin tai luennoijan osoittama materiaali. Kurssi voidaan täydentää vaihtoehtoisesti harjoitustyöllä syventäviksi opinnoiksi.

#### **57742 Hierarchical models, 6-8 op**

Tilastolliset tutkimusaineistot voidaan usein luontevasti tulkita kerroksittaisena rakenteena, jossa kukin kerroksittaisen rakenteen ryhmä voidaan soittaa useaan pienemmän jaon mukaiseen alaryhmään. Hierarkkinen mallintaminen on yleinen tapa hallita tällaisia tutkimusaineistoja. Hierarkkisten mallien kehikossa voidaan esittää useita mm. biometriassa, psykometriassa ja ekonometriassa käytettyjä malleja, jotka ovat yleisen teorian erikoistapauksia.

#### **57740 Nonparametric statistics, 6 op**

Parametrisen ja epäparametrisen tilastollisen päättelyn käsitteet. Empiiriset jakaumat, Kolmogorov-Smirnov -testi, luottamusvälit ja kvantiilit, merkkitesti, Wilcoxon-Mann-Whitney -testi, Kruskal-Wallis -testi, riippumattomuustestit, korrelaation epäparametriset mitat, asymptoottinen tehokkuus, epäparametrisuus Bayes-kehikossa.

### **57547 Tilastollisen koneoppimisen linja**

Linjalle voidaan erikoistua myös soveltavassa matematiikassa. Vaadittavat aineopinnot noudattavat joko soveltavan matematiikan tai tilastotieteen tutkintovaatimuksia. Suositeltavia kursseja ovat tietojenkäsittelytieteen kurssit, erityisesti Introduction to Machine Learning. Tilastollisella koneoppimisella viitataan yleisesti todennäköisyysmallien nojaviin tietokonealgoritmeihin jotka kykenevät oppimaan tarkastellulle ilmiölle tyypillisiä piirteitä opetusaineistosta ja suorittamaan tämän perusteella esimerkiksi luokittelua ja ennustetehtäviä uusille havaintoaineistoille. Usein keskeistä näiden menetelmien toimivuuden kannalta on niiden kyky suodattaa pois opetusaineistosta mahdollisesti olevaa epäolennaista satunnaisvaihtelua ja löytää piirteitä jotka yleistyvät opetusaineiston ulkopuolisiin havaintoihin.

### 57733 Computational statistics, 8 op

Kurssilla käsitellään tiettyjä moderneja tilastomenetelmiä, jotka vaativat paljon laskentaa. Kurssin aiheita ovat esimerkiksi todennäköisyysjakaumien simulointi, Monte Carlo -integrointi sekä EM-algoritmi. Yhden tärkeän aihepiirin muodostavat MCMC-menetelmät, jotka ovat erityisesti Bayes-päätelyyn soveltuvia laskentamenetelmiä, jotka perustuvat Markovin ketjujen simulointiin. Esitietoina edellytetään Bayes-päätely (tai vastaavat tiedot).

### 57744 Bayesian theory with applications, 5 op

Bayesläistä lähestymistapaa käytetään nykyään laajalti todennäköisyysmallinnuksessa useimmilla tieteenaloilla ja se edustaa myös hyvin tärkeää työkalua insinöritieteiden teknologisissa sovelluksissa. Kurssilla perehdytään bayesläisen mallinnuksen teoreettisiin perusteisiin sekä tarkastellaan yksityiskohtaisesti miten ja miksi sen avulla pystytään tehokkaasti ratkaisemaan erilaisia sovellusongelmia. Esitietoina edellytetään Bayes-päätely (tai vastaavat tiedot).

## Jatkotutkinnot tilastotieteessä

Jatkotutkintona voi valtiotieteellisessä tiedekunnassa suorittaa valtiotieteiden lisensiaatin (VTL) tai valtiotieteiden tohtorin (VTT) tutkinnon pääaineena tilastotiede, ja matemaattis-luonnontieteellisessä tiedekunnassa filosofian lisensiaatin (FL) tai filosofian tohtorin (FT) tutkinnon pääaineena tilastotiede. Jatko-opintojen edellytyksenä on soveltuva ylempi korkeakoulututkinto.

## Lisensiaatin tutkinto

Lisensiaatin tutkinnon suorittamiseksi on

- 1) suoritettava laitoksen hyväksymässä opintosuunnitelmassa määrätyt tiedekuntakohtaiset vähintään 60 op:n laajuiset jatko-opinnot, jotka koostuvat vähintään 50 op:n tieteenalan opinnoista sekä vähintään 10 op:n yleisistä jatko-opinnoista (näihin tulee sisältyä tieteenfilosofian, tutkimusetiikan ja yleiseen asiantuntijuuteen valmentavia opintoja sekä kansainvälistä tieteellistä toimintaa); sekä
- 2) laadittava lisensiaatintutkimus, jonka tiedekunta hyväksyy.

## Tohtorin tutkinto

Tohtorin tutkinnon suorittamiseksi on 1) suoritettava lisensiaatin tutkinto pääaineena tilastotiede, tai edellä esitetty lisensiaatin tutkinnon vaatimusten kohta 1, sekä 2) laadittava ja julkaistava väitöskirja, jonka tiedekunta julkisen tarkastuksen jälkeen hyväksyy.

**Päivitys 7.2.2011**

## Tilastotieteen opintojen ajoitusmalli

Yleisopinnoissa on tiedekuntakohtaisia eroja (VT = valtiotieteellinen, ML = matemaattis-luonnontieteellinen). (P) tarkoittaa pakollisia opintoja.

### Kandidaatin tutkinto (VTK/LuK), 180 op

#### 1. syksy

Minustako tilastotieteilijä? 5 op  
Analyysi I, 10 op (P) tai Matemaattisen analyysin kurssi, 10 op (P)  
Lineaarialgebra ja matriisilaskenta I, 5 op (P)  
Lineaarialgebra ja matriisilaskenta II, 5 op (P)  
TVT-ajokortti, 3 op (P)  
Ohjaajatuutorointiin osallistuminen, 1 op  
Pääaineopintoja, 10 op

#### 1. kevät

Johdatus todennäköisyyslaskentaan, 5 op (P)  
Johdatus tilastolliseen päätelyyn, 5 op (P)

Analyysi II, 10 op (P) tai Matemaattisen analyysin jatkokurssi, 10 op (P)  
HOPS, 1 op (P)  
Ohjaajatuutorointiin osallistuminen, 1 op  
Pääaineopintoja, 5 op  
Vieras kieli (VT: 3 op, ML: 4 op) (P)

## 2. syksy

Todennäköisyyslaskenta, 10 op (P)  
Tilastollinen päättely (*jatkuu keväälle*) (P)  
Pääaineopintoja, 10 op  
Sivuaineopintoja, 10 op

## 2. kevät

Tilastollinen päättely (*jatkuu syksyltä*), 10 op (P)  
Lineaariset mallit, 5 op (P)  
Toinen kotimainen kieli (VT: 4 op, ML: 3 op) (P)  
Sivuaineopintoja, 10 op

## 3. syksy

Proseminaari (VT: 4 op, ML: 1 op + 3 op äidinkielen opinnot, ks. alla) (P)  
Työelämäorientaatio, 1–3 op (P)  
Pää- ja sivuaineopintoja, 20 op (mahdollisesti jo maisteriopintojen linjaa tukevia)

## 3. kevät

Kandidaatintutkielma, 6 op (P)  
Äidinkielen opinnot, 3 op (VT: kirjallisen käytön kurssi, ML: proseminaari) (P)  
Pää- ja sivuaineopintoja, 20 op (mahdollisesti jo maisteriopintojen linjaa tukevia)

## Maisterin tutkinto (VTM/FM), 120 op

## 4. syksy

Tilastollisen päättelyn jatkokurssi, 5 op (P)  
Yleistetyt lineaariset mallit (osa 1), 5 op (P)  
HOPS, 1 op (P)  
Pää- ja sivuaineopintoja, 20 op

## 4. kevät

Pää- ja sivuaineopintoja, 30 op  
(osa pakollisia linjasta riippuen)

## 5. syksy ja kevät

Tutkielmaseminaari, 2 op (P)  
Pro gradu –tutkielma, 40 op (P)  
Harjoittelu (VT: vähintään 5 op, ML: vähintään 1 op) (P)  
Pro gradu -tutkielmaa tukevia pää- ja sivuaineopintoja, 20 op

# Tutkintovaatimukset tilastotieteen sivuaineopiskelijalle

## Tilastotieteen sivuainekokonaisuudet

[Perusopinnot 25 op](#)

[Aineopinnot 35 op](#)

[Soveltavan tilastotieteen perusopinnot 25 op](#)

[Syventävät opinnot 60 op](#)

[Tutkimusmenetelmät yhteiskuntatieteissä](#)

[Menetelmätieteiden sivuaine](#)



### HUOM!

Nämä tutkintovaatimukset koskevat kaikkia 1.8.2011 (?) tai sen jälkeen sivuaineopintonsa aloittaneita opiskelijoita.

## TILASTOTIETEEN OPINNOT SIVUAINEOPISELIJOILLE

Perusopintojen (25 op) suorittamiseksi voi valita joko teoreettisemman tai soveltavamman vaihtoehdon. Valinta riippuu mm. tavoiteltavien sivuaineopintojen laajuudesta, opiskelijan pääaineen mahdollisista erityisvaatimuksista ja opiskelijan matemaattisista valmiuksista. Perusopintoihin voidaan sisällyttää aineopintojen opintojaksoja.

### Vaihtoehto 1: 57505 / 78015 Tilastotieteen perusopinnot (25 op)

Vaihtoehto 1 tarjoaa teoriapainotteiset pohjatiedot tilastotieteen aineopintoihin tai muihin perusopintoja laajempiin sivuaineopintoihin. Suositeltava kokonaisuus koostetaan seuraavista kursseista, joista kaksi ensin mainittua vaaditaan aineopintoja varten:

#### 57045 Johdatus todennäköisyyslaskentaan, 5 op

Kurssilla esitellään todennäköisyyden ja satunnaismuuttujan käsitteet, tärkeimmät diskreetit ja jatkuvat jakaumat sekä niiden tunnusluvut. Kurssilla käsitellään myös kombinatoriikkaa sekä suurten lukujen lakia ja keskeistä raja-arvolauseetta.

#### 57046 Johdatus tilastolliseen päättelyyn 5 op

Kurssilla perehdytään tilastollisen mallin käsitteeseen ja siihen liittyvän päättelyn kysymyksenasetteluihin perinteisen frekventistisen paradigman näkökulmasta. Pääpaino on parametristimoinnissa. Lisäksi käsitellään mm. hypoteesintestausta lähinnä normaalijakauman tapauksessa.

#### 57703 Data-analyysi R-ohjelmistolla, 5 op

Kurssilla perehdytään data-analyysin perusteisiin käyttäen avoimen lähdekoodin R-ohjelmistoa (<http://www.r-project.org>), joka on maailmanlaajuisesti merkittävä data-analyysin ja tilastollisen mallinnuksen työkalu. R on sekä matemaattisiin sovelluksiin suunnattu ohjelmointikieli että laajennettavissa oleva laskentaympäristö, jonka käyttöä erityissovelluksissa tukee suuri määrä näitä varten luotuja ohjelmakirjastoja. Kurssilla tarkastellaan erilaisia havaintoaineistoja sekä graafisesti että tilastollisten menetelmien avulla.

#### 78116 Data-analyysi (osa 1), 5 op

#### 78132 Data-analyysi (osa 2), 5 op

Kurssilla opiskellaan data-analyysin perusteet ja perehdytään jonkin tilastollisen tietojenkäsittelyjärjestelmän itsenäiseen käyttöön analysoidulla kyselyaineistoja ja muita yhteiskuntatieteellisiä aineistoja. Kurssi rakentuu kahdesta periodin pituisesta osasta. Osa 1 koostuu luennoista ja harjoituksista, ja se suoritetaan työraporteilla sekä itse- ja vertaisarvioinneilla. Osassa 2 tehdään ohjastusti harjoitustyö, jossa sovelletaan osassa 1 opittuja asioita uuteen aineistoon.

### Vaihtoehto 2: 57505/ 78015 Tilastotieteen perusopinnot (25 op)

Vaihtoehto 2 tarjoaa sovelluspainotteiset pohjatiedot useille mm. yhteiskunta- ja käyttäytymistieteissä tarvittaville menetelmäopinnoille. Tilastotieteen johdantokurssi (10 op) on vaihtoehdossa 2 pakollinen. Lisäksi on suoritettava valinnaisia kursseja niin että kokonaisuuden minimilaaajuus 25 op täyttyy. Valinnaisia kursseja ovat esimerkiksi Tilastotieteen jatkokurssi, Tilastotiede käytännön tutkimuksessa ja Data-analyysi (osat 1 ja 2). Vaihtoehdon 2 perusopintojen pohjalta voi sivuaineopintoja laajentaa suorittamalla Soveltavan tilastotieteen perusopinnot (25 op), ks. alempana.

#### 78200 Tilastotieteen johdantokurssi (osa 1), 4 op

#### 78201 Tilastotieteen johdantokurssi (osa 2), 6 op

Kurssilla opitaan sovelluspainotteisesti tilastollisen tutkimuksen perusteet erityisesti yhteiskuntatieteiden näkökulmasta. Osassa 1 tarkastellaan, miten aineistoja kerätään ja mitä kyselylomakkeessa tulee mittauksen kannalta ottaa huomioon. Lisäksi tutkitaan, millä tavoin tietoja tiivistetään tilastollisiksi kuviksi ja tunnusluvuiksi sekä perehdytään erilaisiin vaihteluun ja riippuvuuden kuvaustapoihin. Osassa 2 tutustutaan todennäköisyyslaskentaan ja tilastolliseen päättelyyn sekä sovelletaan niitä luottamusvälien tarkasteluun sekä eräisiin yhteiskuntatieteissä keskeisiin tilastollisiin menetelmiin kuten t-testiin ja regressioanalyysiin.

#### 78203 Tilastotieteen jatkokurssi (osa 1), 5 op

#### 78204 Tilastotieteen jatkokurssi (osa 2), 5 op

Kursseilla syvennetään johdantokurssilla annettuja tietoja tilastollisten menetelmien soveltamisesta erilaisissa tutkimusasetelmissä. Kurssi perustuu luentomateriaaliin tai luennoijan kanssa sovittavaan kirjallisuuteen. Esitietoina edellytetään Tilastotieteen johdantokurssi, osa 1 ja osa 2 10 op laajuisena (tai vastaavat tiedot).

#### 78103 Tilastotiede käytännön tutkimuksessa, 8-10 op

Tilastotieteellistä ajattelua ja menetelmiä, niiden periaatteita, käyttöä, soveltuvuutta ja ongelmia opetetaan ja havainnollistetaan käytännön esimerkkien avulla. Mahdollisia muita aihepiirejä ovat esimerkiksi käytännön tutkimusongelmien merkitys tilastotieteen kehitykselle, tilastotieteen historia, yhteiskunnallisesti merkittävä empirinen tilastotieteellinen tutkimus, Edeltävät opinnot: Tilastotieteen johdantokurssi tai vastaavat tiedot. Pääaineopiskelijat voivat sisällyttää kurssin perusopintoihin ylimääräisinä opintoina.

#### 78116 Data-analyysi (osa 1), 5 op

#### 78132 Data-analyysi (osa 2), 5 op

Kurssilla opiskellaan data-analyysin perusteet ja perehdytään jonkin tilastollisen tietojenkäsittelyjärjestelmän itsenäiseen käyttöön analysoidulla kyselyaineistoja ja muita yhteiskuntatieteellisiä aineistoja. Kurssi rakentuu kahdesta periodin pituisesta osasta. Osa 1 koostuu luennoista ja harjoituksista, ja se suoritetaan työraporteilla sekä itse- ja vertaisarvioinneilla. Osassa 2 tehdään ohjastusti harjoitustyö, jossa sovelletaan osassa 1 opittuja asioita uuteen aineistoon.

### 57518 / 78034 Tilastotieteen aineopinnot (35 op)

Aineopinnoissa vaatimukset ovat (Proseminaaria ja Kandidaatintutkielmaa lukuun ottamatta) samat kuin tilastotieteen pääaineopiskelijoilla. Pakollisten opintojen lisäksi on suoritettava valinnaisia, vähintään aineopintotasoisia kursseja niin että kokonaisuuden minimilaaajuus 35 op täyttyy.

Edeltävinä tietoina vaaditaan tilastotieteestä vaihtoehdon 1 (suositeltava) tai vaihtoehdon 2 perusopinnot (25 op) tai vastaavat tiedot.

Lisäksi vaaditaan matematiikasta vähintään 30 op laajuinen kokonaisuus, joka sisältää kurssit:

78121 Matemaattisen analyysin kurssi tai 57016 Analyysi I, 10 op

78122 Matemaattisen analyysin jatkokurssi tai 57017 Analyysi II, 10 op

57043 Lineaarialgebra ja matriisilaskenta I, 5 op

57047 Lineaarialgebra ja matriisilaskenta II, 5 op

## Tilastotieteen aineopintojen pakolliset opinnot 25 op:

### 57705 Todennäköisyyslaskenta, 10 op

Kurssilla opetellaan tilastollisen mallintamisen kannalta keskeisiä osia todennäköisyyslaskentaa, kuten esimerkiksi moniulotteisten jakaumien käsittelyä. Opittavia valmiuksia tarvitaan useilla myöhemmillä tilastotieteen kursseilla. Esitietona edellytetään perustietoja todennäköisyyslaskennasta (esim. kurssia Johdatus todennäköisyyslaskentaan).

### 57701 Tilastollinen päättely, 10 op

Kurssilla esitetään uskottavuuspäättelyn perusteet, joita tarvitaan tilastotieteen erikoiskurssien opiskelua varten. Kurssi edellyttää Todennäköisyyslaskennan kurssin tiedot.

### 78134 Lineaariset mallit, 5 op

Kurssilla esitetään lineaaristen mallien perusteoria ja lineaaristen mallien rakentamisen pääperiaatteet. Kurssi on tarkoitettu suorittamaan heti Tilastollisen päättelyn kurssin jälkeen.

## 57506 / 78017 Soveltavan tilastotieteen perusopinnot (25 op)

Soveltavan tilastotieteen perusopinnot on tarkoitettu opiskelijoille, jotka haluavat laajentaa tilastotieteen sivuaineopintoja sovelluspainotteisilla menetelmäkursseilla. Mahdollisia kursseja on lueteltu perusopintojen vaihtoehto 2:n yhteydessä. Lisäksi kokonaisuuteen soveltuvia kursseja ovat mm. seuraavat:

### 78126 Lineaaristen mallien sovellukset, 6-8 op

Sovelluspainotteinen kurssi keskittyy regressioanalyysin käyttöön yhteiskunta- ja käyttäytymistieteellisessä tutkimuksessa. Tavoitteena on oppia analysoimaan käytännön tutkimusaineistoja, arvioimaan tilastollista mallia havaintoaineiston informaation tiivistäjänä sekä esittämään tutkimusongelmat ja tulokset selkeästi graafisia menetelmiä ja vastaavia tilastollisia esitystapoja käyttäen.

### 78164 Mittaaminen ja mittausvirheiden tilastollinen hallinta, 6-8 op

Kurssilla perehdytään mittausvirheiden vaikutuksen arviointiin sekä tässä tarvittaviin tilastollisiin menetelmiin. Tarkasteltavia aiheita ovat mittaamisen tasot, niiden vaikutus analyysimenetelmiin, mittaamisen laadun arviointi (reliabiliteetti ja validiteetti), mittausmalli ja rakennevaliditeetti, mittauksen reliabiliteetin arviointi, mitta-asteikkojen laatiminen ja ennustevaliditeetti.

### 78145 Monimuuttujamenetelmät, 6-8 op

Kurssilla perehdytään sovelluspainotteisesti usean muuttujan data-analyysiin sekä aineistojen moniulotteiseen kuvailuun ja mallintamiseen erityisesti yhteiskunta- ja käyttäytymistieteissä. Tarkasteltavia menetelmiä ovat faktorianalyysi, erilaiset ryhmittelymenetelmät, erotteluanalyysi, moniulotteinen skaalaus sekä korrespondenssianalyysi.

### 78143 Otantamenetelmät, 6-8 op

Kurssilla annetaan yleiskuva tilastollisista otantamenetelmistä ja niiden käytöstä eri tieteenalojen empiirisessä tutkimuksessa. Otannan perusmenetelmien lisäksi käsitellään otoskoon määrittelyyn perusteita ja lisäinformaation käyttöä otannassa ja estimoinnissa sekä vastaavia tilastollisia ohjelmistoja.

### 78405 Otanta-aineiston analyysi, 6-8 op

Kurssilla perehdytään otanta-aineistojen tilastolliseen analyysiin erityisesti tilanteissa, joissa aineisto on kerätty jollakin mutkikkaalla otanta-asetelmalla. Keskiössä ovat menetelmät, joilla otanta-asetelman ominaispiirteitä (ositus, ryvästyminen, painokertoimet) voidaan ottaa huomioon tilastollisen analyysin yhteydessä. Pääpaino on ns. asetelmaperusteisissa (design-based) menetelmissä; lisäksi tarkastellaan suppeasti malliperusteisia menetelmiä (ml. sekamallit).

### 78187 Survey-metodiikka, 8-10 op

Kurssilla esitellään empiirisen kvantitatiivisen yhteiskuntatutkimuksen eli ns. surveyn keskeiset käsitteet, menetelmät ja suoritusvaiheet. Tarkasteltavia aiheita ovat tutkimusasetelmat, otannan perusteet, aineiston keruu ja lomakesuunnittelu, poikkileikkaus- ja pitkittäisaineistot, aineiston puhdistaminen (tilastollinen editointi, imputointi ja uudelleenpainotus), estimoinnin ja analyysin perusteet sekä aineiston jakelu ml. tilastolliset tietosuojamenetelmät.

## 57565 / 78090 Tilastotieteen syventävät opinnot sivuaineopiskelijoille (60 op)

Tilastotieteen syventävät opinnot sivuaineena voidaan suorittaa kaikilla tilastotieteen erikoistumislinjoilla. Ne voidaan sisällyttää maisterintutkintoon, jonka pääaine ei ole tilastotiede, ja niiden edellytyksenä ovat yhteensä vähintään 60 op:n laajuiset tilastotieteen perusopinnot ja aineopinnot sivuaineopiskelijalle. Syventävinä opintoina on suoritettava vähintään 60 op ko. erikoistumislinjan opintoja; nämä sisältävät tutkielman 20 op, ja kursseista on sovittava linjasta vastaavan opettajan kanssa.

## Tiedekuntaohittaiset kurssit ja opintokokonaisuudet

### 78450 Sosiaalitutkimuksen tilastolliset menetelmät, 5 op

Kyseessä on sosiaalitieteiden oppiaineiden uuden yhteisen, kandidaatin tutkintoon kuuluvan, metodikurssin luentojakso. Esitietovaatimuksena on Tilastotieteen johdantokurssi (osat 1 ja 2). Oppiainekohtaisista harjoituksista on tietoa mm. sosiaalipolitiikan, sosiaalityön ja viestinnän tutkintovaatimuksissa. Sekä luentojakso että harjoitukset on tarkoitettu ainoastaan valtiotieteellisen tiedekunnan opiskelijoille. Luentojaksolla tavoitteena on perehtyä sosiaalitutkimuksessa sovellettaviin tiedonkeruun ja analysoinnin tilastollisiin näkökohtiin ja menetelmiin. Kurssi järjestetään kevätlukukaudella. Tarkemmat tiedot löytyvät kurssin kotisivulta.

## 799450 Tutkimusmenetelmät yhteiskuntatieteissä (25 op)

Opintokokonaisuus on tarjolla valtiotieteellisessä tiedekunnassa ja liittyy tiedekunnan maisteriopintojen yhteiseen menetelmäkoriin (tarkemmat tiedot valtiotieteellisen tiedekunnan opinto-oppaassa). Opintokokonaisuus koostuu valinnaisista kursseista, joiksi kelpaavat muun muassa sivuaineopiskelijoiden (vaihtoehtoon 2 mukaisten) perusopintojen valinnaiset opinnot. Esitietoina edellytetään Tilastotieteen johdantokurssi 10 op laajuisena (tai vastaavat tiedot).

## 57509 Menetelmätieteiden perusopinnot kokonaisuus (25 op) ja 57508 aineopinnot kokonaisuus (35 op)

Menetelmätieteiden sivuainekokonaisuuden opinto-oikeus on kaikilla niillä yliopiston opiskelijoilla, joiden pääaineen tutkintovaatimuksissa on määritelty pääaineeseen sopiva menetelmätieteiden sivuainekokonaisuus. Sivuaine koostuu matematiikan, tilastotieteen ja tietojenkäsittelytieteen opintojaksoista, joista opiskelija voi koota kokonaisuuden tutkintoonsa oman koulutusohjelmansa määräämien rajojen puitteissa. Menetelmätieteiden sivuainekokonaisuutta varten on koottu matematiikan, tilastotieteen, bioinformatiikan ja tietojenkäsittelytieteen opintojaksoja tarjolle yhteen "koriin", josta sivuainetta suorittavat opiskelijat voivat joustavasti koota kokonaisuuden tutkintoonsa oman koulutusohjelmansa tutkintovaatimusten mukaisesti. Tarkemmat tiedot matemaattis-luonnontieteellisen tiedekunnan opinto-oppaassa.