

# Tilastotieteen johdantokurssi, syksy 2011

## Tilastotieteen johdantokurssi, syksy 2011



\*Kysyttävää?\*

[tilasto-johdanto\(at\)helsinki.fi](mailto:tilasto-johdanto(at)helsinki.fi)



Suoritustavat, korvaavuudet, aikataulut, tenttijärjestelyt yms.



Sisällölliset kysymykset: luennoilla ja työpajoissa.

### Luennoitsija ja osan 2 työpajaohjaajat

Kimmo Vehkalahti, VTT, dosentti  
Eija-Leena Koponen, VTK  
Miika Paavola, fil.yo  
Heidi Ryytänen, VTK  
Tomi Salminen, LuK

### Laajuus ja koodit

Osa 1: 4 op (78200)

Osa 2: 6 op (78201)

### Tyyppi

Sovelluspainotteinen, perustavanlaatuisen perusopinnotason kurssi tilastotieteen sivuaineopiskelijoille erityisesti yhteiskuntatieteiden näkökulmasta. Katso [tutkintovaatimukset 2011-2012](#).

### Kohderyhmä ja rajoitukset

Kurssi on ensisijaisesti suunnattu **valtiotieteellisen tiedekunnan opiskelijoille**.

Muiden tiedekuntien opiskelijoille suositellaan heidän omien tiedekuntiansa järjestämiä tilastotieteen (tai biometrian, epidemiologian ym.) peruskursseja.

Statistik **på svenska**? Se: [Soc&kom \(svenska social- och kommunalhogskolan\)](#)

Statistics **in English**? See: [Introduction to Statistics, fall 2011](#)

### Tavoitteet ja sisältö

Tilastotiede on tärkeä osa yhteiskunta- ja muita tieteitä, joissa kerätään, mitataan ja analysoidaan tietoja. **Tilastollisilla menetelmillä** on mahdollista löytää vastauksia lukuisiin eri alojen tutkimuskysymyksiin.

Tavoitteena on oppia **tilastollisen tutkimuksen perusteet**. Näkökulma on **yhteiskuntatieteellinen** ja lähestymistapa **sovelluspainotteinen**. Jotta perusteet voi oppia, on tärkeää ymmärtää **keskeiset käsitteet**.

- Osassa 1 tarkastellaan, miten yhteiskuntatieteellisiä **aineistoja** kerätään ja mitä kaikkea **kyselylomakkeen** laatimisessa tulee ottaa huomioon. Lisäksi tutkitaan, millä tavoin tietoja tiivistetään tilastollisiksi **kuviksi** ja **tunnuksiksi** sekä perehdytään erilaisiin **vaihtelun** ja **riippuvuuden** kuvaustapoihin.
  - *Osan 1 aikana korostuu itsenäinen työskentely. Monet käsitteistä saattavat olla entuudestaan tuttuja, mutta niihin perehdytään syvällisemmin kuin esimerkiksi lukion kursseissa.*
- Osassa 2 tutustutaan **todennäköisyytlaskentaan** ja **tilastolliseen päättelyyn** sekä sovelletaan niitä **luottamusvälien** tarkasteluun sekä eräisiin yhteiskuntatieteissä keskeisiin tilastollisiin perusmenetelmiin kuten **t-testiin** ja **regressioanalyysiin**.
  - *Osan 2 käsitteet ovat vaativampia ja edellyttävät täsmällisempää työstämistä ja pohdiskelua. Lyhyt matematiikka riittää pohjaksi, mutta tällöin kurssia varten on syytä varata enemmän aikaa. Työpajoihin osallistuminen on suositeltavaa, jos harjoitustehtävät alkavat tuntua vaikeilta.*

**Tilastotiede ei ole ensisijaisesti matematiikkaa**, vaikka sen teoriapohja matematiikkaan nojaakin. **Tilastollisen ajattelun** omaksuminen vaatii väijäämättä jonkin verran matemaattisuontoista päähäilyä, laskeskelua ja pohdiskelua.

Tyypillisiin tilastotieteen peruskursseihin verrattuna tällä kurssilla on huomattavasti vähemmän **laskemista**. Sen sijaan keskitytään miettimään niitä kysymyksiä, joita tilastollisessa tutkimuksessa oikeasti kohdataan:

- Mitä tietoja kerätään, keneltä tai mistä ja millä tavalla?
- Miten ilmiötä ja ulottuvuuksia mitataan?
- Millaisia menetelmiä missäkin tilanteessa voidaan soveltaa?
- Mitä tietokone pitäisi laittaa laskemaan ja miten?
- Millä tavoin tuloksia esitetään ja tulkitaan?
- Mitä johtopäätöksiä voidaan tehdä ja mitä ei?

**Kurssin jälkeen** tulisi osata vastata ainakin useimpiin näistä kysymyksistä, kun ne tulevat eteen todellisessa tilanteessa. Monet kysymyksistä kohdataan jo **oman alan tutkimuskirjallisuutta** lukiessa, puhumattakaan opinnäytetöiden tekemisestä.

## Ajat ja paikat

Opetustilaisuuksien ohella on varattava aikaa **itsenäiseen työskentelyyn**, joka tarkoittaa mm. pohtimista ja päähkäilyä joko itsekseen tai kaveri(e)n kanssa (jälkimmäinen on erittäin suositeltavaa!). Jos ei pääse luennoille (esim. päällekkäisyyksien vuoksi), itsenäisen työn merkitys korostuu entisestään. **Palautte en perusteella kurssista saa eniten irti käymällä aktiivisesti luennoilla ja tekemällä mahdollisimman paljon harjoitustehtäviä.**

### Osa 1 (periodi I)

- viikko 36:
  - luento to 8-10 (Yliopistonkatu 3, Porthania, PI)
- viikot 37-41:
  - luento ke 8-10 (Unioninkatu 40, Metsätalo, sali 1)
  - luento to 8-10 (Yliopistonkatu 3, Porthania, PI)
- viikko 42:
  - **kurssikoe to 20.10 klo 12-14** (ks. Suoritustavat)

### Väliviikko

- viikko 43:
  - (ei opetusta)

### Osa 2 (periodi II)

- viikko 44:
  - luento to 8-10 (Yliopistonkatu 3, Porthania, PI)
- viikot 45-48:
  - työpajat ma 14-16, ma 16-18, ti 10-12 ja ti 12-14 (Fabianinkatu 24, sh 135) (ks. Työpajat)
  - luento ke 8-10 (Unioninkatu 40, Metsätalo, sali 1)
  - luento to 8-10 (Yliopistonkatu 3, Porthania, PI)
- viikko 49:
  - työpajat ma 14-16 ja ma 16-18 (Fabianinkatu 24, sh 135) (ks. Työpajat)
  - **ti: itsenäisyyspäivä** (ei opetusta)
  - luento ke 8-10 (Unioninkatu 40, Metsätalo, sali 1)
  - **to: ei luento**
- viikko 50:
  - **kurssikoe to 15.12 klo 10-14** (ks. Suoritustavat)

### Joululoma

- viikot 51-52:
  - (ei opetusta)

## Luennot ja harjoitukset

Kurssin molemmat osat koostuvat **viidestä teemasta**, joihin perehdytään torstain luennoilla. **Perusmateriaalia (ks. alla) täydennetään luennoilla monenlaisilla esimerkeillä, keskustelulla ja kommentailla.** Teemoihin kuuluu myös **harjoitustehtäviä**, joita työstetään osassa 1 itsenäisesti. Osassa 2 tehtävien tekemistä tukevat **työpajat**. Tehtävien ratkaisuja tarkastellaan yhdessä keskiviikon luennoilla, jotka voi katsoa jälkepäin verkosta.

Luentomateriaali ja harjoitukset tulevat saataville oheisiin luetteluihin. Harjoitukset julkaistaan viikoittain, joko ennen torstain luentoa tai pian sen jälkeen. Osan 2 harjoituksista julkaistaan myös yksityiskohtaisempia neuvoja (etukäteen) sekä luonnosmaisia ratkaisuehdotuksia omien pohdiskelujen tukemiseksi (jälkikäteen).

### Osa 1

#### Kyselytutkimus ja mittaus

- Teema 1: Tilastotiede ja kyselytutkimus
- Teema 2: Mittaus yhteiskuntatieteissä

#### Tiedonkeruu, tilastoaineiston kuvailu ja analysointi

- Teema 3: Tiedonkeruu, kuvat ja tunnusluvut
- Teema 4: Vaihtelu ja riippuvuus
- Teema 5: Ristiintaulukointi

### Osa 2

#### Todennäköisyyslaskennan perusteet

- Teema 6: Todennäköisyys ja satunnaisuus
- Teema 7: Todennäköisyyksien laskentaa

### Tilastollisen päättelyn perusteet

- Teema 8: Parametrien estimointi ja luottamusvälit
- Teema 9: Tilastollinen hypoteesin testaus
- Teema 10: Regressio- ja varianssianalyysi

## Työpajat (osa 2)

Harjoitustehtävien tekeminen on oppimisen kannalta erittäin olennaista. Osassa 1 tehtävät tehdään itsenäisesti. Osassa 2 oppimista tukevat **vapaaehtoiset työpajat**, joissa tehtäviä työstetään ohjaajien johdolla pienryhmissä. Tehtäviin perehdytään itsenäisesti jo ennen työpajan kokoontumista. Pajoihin ei ilmoitauduta, vaan saa käydä siinä ryhmässä, joka itselle parhaiten sopii. Lisäpisteitä ei pajoista saa (pisteet saa kokeista, ks. **Suoritustavat**).

Työpajat järjestetään osoitteessa [Fabianinkatu 24, sh 135](#). Ohjaajat ovat tilastotieteen pääaine- tai sivuaineopiskelijoita. Pajat pyörivät edellä olevan viikkoaikataulun mukaisina päivinä seuraavasti:

Päivä	Aika	Ohjaaja(t)
ma	14 - 16	Eija-Leena Koponen ja Miika Paavola
ma	16 - 18	Miika Paavola ja Eija-Leena Koponen
ti	10 - 12	Heidi Ryyänen ja Tomi Salminen
ti	12 - 14	Tomi Salminen ja Heidi Ryyänen

### Ilmoittaudu osaan 1

**i** Osan 1 kurssikokeen tulokset on raportoitu ilmoittautuneille ja viety Oodiin.

### Ilmoittaudu osaan 2

**i** Osan 2 kurssikokeen tulokset on raportoitu ilmoittautuneille ja viety Oodiin.

## Suoritustavat

Kurssi suoritetaan joko kurssikokeilla tai erilliskokeilla.



Suosittelavin suoritustapa on kurssikoe.

### Kurssikoe (vain kurssille ilmoittautuneille)

Osan 1 kurssikoe järjestetään I periodin lopussa ja osan 2 kurssikoe II periodin lopussa.

- Osan 1 koealue kattaa Teemat 1-5 ja Harjoitukset 1-5.
- Osan 2 koealue kattaa Teemat 6-10 ja Harjoitukset 6-10.

### Kurssikokeen salijako ilmoitetaan (sähköpostilla) ennen koepäivää.

- Osa 1: **to 20.10 klo 12-14** (Yliopistonkatu 3, Porthania, PI ja [Fabianinkatu 33, päärakennus, sali 1](#))
- Osa 2: **to 15.12 klo 10-14** (Yliopistonkatu 3, Porthania, PI ja [Fabianinkatu 33, päärakennus, sali 1](#))

**i** Kokeeseen osallistuminen edellyttää kurssin vastaavaan osaan ilmoittautumista.

**i** Kurssikokeiden ajat kannattaa merkitä heti kalenteriin, sillä varsinaisia uusintakokeita ei järjestetä (uusinnat suoritetaan erilliskokeissa, ks. alla).



Mitä mukaan kokeisiin, mitä ei?

- **Opiskelijakortti** (tai muu henkilötodistus) on oltava mukana.
- **Laskin** saa olla mukana (tavallinen [neilaskin](#) riittää, mutta [funktiolaskimesta](#) voi olla hyötyä, kunhan opettelee sen käytön jo ennen koetta).
- **Taulukkokirjaa** ei saa olla, vaan tarvittavat **taulukot** ja **kaavat** annetaan kokeissa. Niiden käyttöä ei kuitenkaan kokeissa neuvota (sitä varten ovat työpajat).

## Erilliskoe (erillissuorituksia ja uusintoja varten)

Erilliskokeet ovat [valtiotieteellisen tiedekunnan lauantaissa tiedekuntatenteissä 21.1.2012 ja 28.4.2012](#) sekä [kesätenteissä ti 19.6.2012 ja pe 17.8.2012](#).

Tiedekuntatentit ja kesätentit järjestetään keskustakampuksella (tarkemmat tiedot yllä olevista linkeistä lähempänä tenttipäiviä). Niitä koskevat seuraavat ohjeet, joista ei poiketa:

- Ilmoittautuminen WebOodissa viimeistään 10 pv ennen tenttiä.
  - **Tiedekunta EI hyväksy myöhästyneitä ilmoittautumisia.**
- Yhdessä tentissä voi suorittaa joko osan 1 **tai** osan 2 **tai** molemmat osat.
- Kuhunkin tentittävään osaan on ilmoitettava erikseen.
- **Tentittävä kirjallisuus** on kurssin luentomateriaali ja harjoitustehtävät (ks. edellä).
- **Tenttikysymykset ovat suomeksi**, men du kan svara på svenska.  
(Valitettavasti Oodin ilmoittautumissivulla olevia, tämän kurssin osalta turhia kohtia (kirjat ja kieli) ei saa helposti poistettua. Niillä ei siis ole tässä mitään merkitystä.)

## Pisteet ja arvosanat

Molemmista osista saa erillisen arvosanan joko kurssikokeen tai erilliskokeen perusteella (harjoitustehtävistä ei saa pisteitä). Arvosanat määräytyvät koepisteiden perusteella oheisen taulukon mukaisesti:

Pisteet	Arvosana
28-30	erinomainen (5)
24-27	kiitettävä (4)
20-23	hyvä (3)
16-19	tydyttävä (2)
12-15	välttävä (1)
0-11	hylätty (0)

**i** Samaa pisteytystä on sovellettu kaikissa kurssin kokeissa vuodesta 2008 lähtien.

## Aineistoja

Luentomateriaalissa, harjoituksissa ja koetehtävissä tarkastellaan lukuisia yhteiskuntatieteellisiä, enimmäkseen kysely- ja haastattelututkimusten aineistoja. Tällä kurssilla tarkastelu tapahtuu lähinnä kyselylomakkeiden, kuvien, tulosteiden ja tunnuslukujen avulla. Tarkemmin aineistoihin pääsee käsiksi mm. **Data-analyyisin** kursseilla.

- [EVAn kansallinen asennetutkimus 2010](#)
  - [\[Aineiston kuvaus | Kyselylomake\]](#)
- [ISSP 2010: ympäristö III: Suomen aineisto](#)
  - [\[Aineiston kuvaus | Kyselylomake\]](#)
- [Kehitysyhteistyötutkimus 2008](#)
- [Nuorisobarometri 2007](#)
- [Suomalaisten arvot ja uskonnollisuus 2008](#)
- [European Social Survey 2002/2003: Suomen aineisto](#)
- [Kaupunkipalvelututkimus 2001](#)
- [KuntaSuomi 2004: kuntatyöyhteistyötutkimus 2003](#)
- [Käsitykset ilmastonmuutoksesta 2006](#)
- [NYT-liitteen moraalikysely 2001](#)
- [Prices and Earnings 2006, 2009](#)
- [Tiedebarometri 2007](#)

## Kirjallisuutta


**Luentomateriaali ja harjoitukset (ks. edellä) kattavat kokeissa vaadittavan sisällön.**

Alla luetellut teokset soveltuvat **oheislukemistoksi** kurssin molempiin osiin.

- Alkula, Tapani & Pöntinen, Seppo & Ylöstalo, Pekka (1994). Sosiaalitutkimuksen kvantitatiiviset menetelmät. WSOY.
- Fowler, Jr., Floyd J. (1995). Improving Survey Questions: Design and Evaluation. Sage.
- Heikkilä, Tarja (2004). Tilastollinen tutkimus, 5. painos. Edita.
- Howell, David C. (2002). Statistical Methods for Psychology, 5th ed., Duxbury.
- Kuusela, Vesa (2000). Tilastografiikan perusteet. Edita.
- Magnello, Eileen & Van Loon, Borin (2009). [Introducing Statistics: a Graphic Guide](#). Icon Books.
- Nummenmaa, Lauri (2004). Käyttämistieteiden tilastolliset menetelmät. Tammi.
- Paulos, J. A. (1991). Numerotaidottomuus: matemaattinen lukutaidottomuus ja sen seuraukset (suom. Klaus Vala), Art House.
- Robbins, Naomi B. (2005). Creating More Effective Graphs. Wiley.
- Taleb, Nassim Nicholas (2010). The Black Swan, 2nd ed., Penguin Books.
- Vehkalahti, Kimmo (2008). Kyselytutkimuksen mittarit ja menetelmät. Tammi.
- Wild, Christopher J. & Seber, George A. F. (2000). Chance Encounters: a First Course in Data Analysis and Inference. Wiley.

- Ziliak, Stephen T. & McCloskey, Deirdre N. (2008). The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives. University of Michigan Press.

Yhteiskuntatieteellisen tietoarkiston (FSD) **kvantitatiivisten tutkimusmenetelmien oppimisympäristössä** on esillä [laaja valikoima oppikirjoja ja esimerkkejä erilaisin menetelmin tehdyistä tutkimuksista](#). Myös monet niistä soveltuvat **oheislukemistoksi** kurssin molempiin osiin.

 Useimmat teokset menevät kurssin sisältöjä pidemmälle.