

TILASTOTIEDE, KURSSIEN KUVAUKSET 13.10.2010

**Kuvaileva tilastotiede ja aineiston hankinta (3622210, 4 op)**

**Descriptive Statistics and Data Acquisition**

Vastuuhenkilö: Esko Valtonen

**Sisältö:** Havaintoaineiston ominaispiirteiden esittämisessä käytettäviä keski- ja hajontalukuja ja kuvioita. Kahden muuttujan välisen assosiaation tarkastelu ristiintaulukoinnin ja sirontakuvion avulla. Regressiomallin käyttö kahden kvantitatiivisen muuttujan välisen riippuvuuden kuvaamiseen. Tutkimusasetelmia, kokeelliseen tutkimuksen periaatteita ja yksinkertaisen satunnaisotoksen poiminta.

**Content:** Graphical and numerical description of data. Crosstabulation and scatterplot. Simple linear regression. Research designs and principles of experiments. Simple random sampling.

**Tavoitteet:** Opintojakson tavoitteena on toisaalta perehdyttää havaintoaineiston tarkasteluun tunnuslukujen ja kuvioiden avulla ja toisaalta tutustuttaa tulosten yleistettävyyden asettamiin vaatimuksiin aineiston hankinnalle.

**Objectives:** Goal of the course is to introduce basic methods to describe data by statistics and graphics. And discuss e.g. on research designs and principles of experiments.

**Oppimateriaali:** Moore-McCabe, Introduction to practice of statistics, 5<sup>th</sup> edition, 2006. Luvut 1-3

**Study materials:** Moore-McCabe, Introduction to practice of statistics, 5<sup>th</sup> edition, 2006. Chapters 1-3

**Toteutustavat:** Luentoja 24 t, harjoituksia 12 t.

**Teaching methods:** Lectures 24 h, exercises 12 h.

**Suoritustavat:** Tentti

**Modes of study:** Exam

**Arvostelu:** 0-5

**Assessment:** 0-5

## **Tilastolliset mallit ja testaus (3622211, 4 op)**

### **Statistical Models and Testing**

Vastuuhenkilö: Esko Valtonen

**Sisältö:** Normaali- ja binomijakauma havaintojen mallina. Otantajakauma ja päättelymenetelmiä populaation keskiarvolle ja populaatio-osuudelle.

**Content:** Normal and binomial distributions. Sampling distributions and some methods of inference on mean and on proportion.

**Tavoitteet:** Tavoitteena on tutustuttaa mallipohjaiseen tilastolliseen päättelyyn muutamien sovellusten avulla.

**Objectives:** Goal of the course is to give introduction to parametric statistical inference.

**Oppimateriaali:** Moore-McCabe, Introduction to practice of statistics, 5<sup>th</sup> edition, 2006. Luvut 4-9.

**Study materials:** Moore-McCabe, Introduction to practice of statistics, 5<sup>th</sup> edition, 2006. Chapters 4-9.

**Toteutustavat:** Luentoja 24 t ja harjoituksia 12 t.

**Teaching methods:** Lectures 24 h, exercises 12 h

**Suoritustavat:** tentti

**Modes of study:** exam

**Arvostelu:** 0-5

**Assessment:** 0-5

**Edellytykset:** Kuvaileva tilastotiede ja aineiston hankinta.

**Prerequisites:** Descriptive Statistics and Data Acquisition

**Lisätiedot:** Tämä kurssi ja Kuvaileva tilastotiede ja aineiston hankinta –kurssi yhdessä vastaavat kurssia Tilastotieteen peruskurssi 1 (J550107)

## **Todennäköisyysmallit, päättely ja epäparametriset menetelmät (3622212, 5 op)**

### **Probability Models, Inference and Nonparametric Methods**

Vastuuhenkilö: Esko Valtonen

**Sisältö:** Todennäköisyyslaskennan peruskäsitteitä; satunnaismuuttuja ja sen jakauma, yhteisjakauma ja riippumattomuus. Tilastollisessa päättelyssä käytettäviä jakaumia. Epäparametrisia testejä populaatioiden keskiarvojen ja jakaumien vertailuun.

**Content:** Basic concepts of probability theory: random variable and distribution, joint distribution and independence. Some common distributions for statistical inference. Nonparametric tests to compare location of population distributions.

**Tavoite:** Opiskelija tutustuu todennäköisyyslaskennan käsitteistöön ja laskusääntöihin, tilastollisessa päättelyssä usein käytettäviin jakaumiin sekä yksinkertaisiin epäparametrisiin päättelymenetelmiin.

**Objectives:** To introduce some key concepts in probability theory and computation, some probability distributions used in statistical inference and non-parametric testing.

**Oppimateriaali:** Noether E., Introduction to statistics – the nonparametric way. 1990 Valittuja kohtia.

Moore- McCabe, Introduction to the practice of statistics, 5<sup>th</sup> edition, 2006. Luvut 4, 15.

**Study materials:** Noether E., Introduction to statistics – the nonparametric way. 1990. Selected sections.

Moore- McCabe, Introduction to the practice of statistics, 5<sup>th</sup> edition, 2006. Chapters 4, 15.

**Toteutustavat:** Luentoja 28 t ja harjoituksia 14 t.

**Teaching methods:** Lectures 28 h, exercises 14 h.

**Suoritustavat:** Kirjallinen kuulustelu

**Modes of study:** Written exam

**Arvostelu:** 0-5

**Assessment:** 0-5

**Edellytykset:** Kuvaileva tilastotiede ja Tilastolliset mallit ja testaus.

**Prerequisites:** Descriptive Statistics and Data Acquisition, Statistical Models and Testing.

## **Regressiotekniikat (3622213, 4 op)**

### **Regression Techniques**

Vastuuhenkilö: Esko Valtonen

**Sisältö:** Yksi- ja kaksisuuntainen varianssianalyysi; yhden ja useamman selittäjän lineaarinen regressiomalli; logistinen regressio. Regressiomallin soveltaminen aikasarjan trendin ja kausivaihtelun analysointiin.

**Content:** One-way analysis of variance and two-way analysis of variance. Simple linear regression, multiple regression and logistic regression. Application of regression to trend estimation of time series and analysis of seasonal variation.

**Tavoitteet:** Opiskelija tutustuu muutamiin tilastollisiin malleihin, joissa analyysi tehdään varianssi- tai regressioanalyysin avulla.

**Objectives:** To give an introduction to some statistical models for analysis of variance and for regression analysis

**Oppimateriaali:** Moore-McCabe, Introduction to practice of statistics, 5<sup>th</sup> edition, 2006. Luvut 10-13,16.

**Study materials:** Moore-McCabe, Introduction to practice of statistics, 5<sup>th</sup> edition, 2006. Chapters 10-13,16.

**Toteutustavat:** Luentoja 24 t ja harjoituksia 12 t

**Teaching methods:** Lectures 24 h, exercises 12 h

**Suoritustavat:** Harjoitustyöt ja tentti

**Modes of study:** Exercises and exam.

**Arvostelu:** 0-5

**Assessment:** 0-5

**Edellytykset:** Kuvaileva tilastotiede ja aineiston hankinta, Tilastolliset mallit ja testaus sekä Todennäköisyysmallit, päättely ja epäparametriset menetelmät.

**Prerequisites:** Descriptive Statistics and Data Acquisition, Statistical Models and Testing, Probability Models and Nonparametric Methods

**SPSS-kurssi (3622222, 2 op)**

**Basic Course in SPSS**

Vastuuhenkilöt: Esko Valtonen ja Eero Korpelainen

**Sisältö:** SPSS-ohjelmiston käytön alkeita yksinkertaisten esimerkkien avulla.

**Content:** Introduction to SPSS-program.

**Tavoitteet:** Tutustuttaa SPSS-ohjelmistoon tilastollisten analyysien apuvälineenä.

**Objectives:** The course gives introduction to SPSS as a tool to do statistical analysis.

**Oppimateriaali:** verkkosivut

**Study materials:** web pages

**Toteutustavat:** itseopiskelu web-kurssina

**Teaching methods:** self-study web course

**Suoritustavat:** testit verkossa

**Modes of study:** tests on web

**Arviointi:** hyväksyty/hylätty

**Assessment:** pass/fail

**Edellytyksenä:** Kuvaileva tilastotiede ja aineiston hankinta sekä Tilastolliset mallit ja testaus

**Prerequisites:** Descriptive Statistics and Data Acquisition, Statistical Models and Testing

**SAS-kurssi (3622221, 2 op)**

**Basic Course in SPSS**

Vastuuhenkilöt: Eero Korpelainen

**Sisältö:** SAS-ohjelmiston käytön alkeita yksinkertaisten esimerkkien avulla.

**Content:** Introduction to SAS-program.

**Tavoitteet:** Tutustuttaa SAS-ohjelmistoon tilastollisten analyysien apuvälineenä.

**Objectives:** The course gives introduction to SAS as a tool to do statistical analysis.

**Oppimateriaali:** SAS-verkkosivut

**Study materials:** Web pages of SAS

**Toteutustavat:** Luentoja 20 t, harjoitustyö

**Teaching methods:** Lectures 20 h, exercises

**Suoritustavat:** Luennot ja harjoitustyöt

**Modes of study:** Lessons and exercises

**Arviointi:** hyväksyty/hylätty

**Assessment:** pass/fail

**Edellytyksenä:** Kuvaileva tilastotiede ja aineiston hankinta sekä Tilastolliset mallit ja testaus

**Prerequisites:** Descriptive Statistics and Data Acquisition, Statistical Models and Testing

**R-kurssi (3622223, 2 op)**

**Basic Course in R**

Vastuuhenkilö: Osmo Kolehmainen

**Suoritustavat:** Erillistä kurssia ei järjestetä, vaan kurssi suoritetaan harjoitustöillä sellaisten kurssien yhteydessä, joilla käytetään R-kieltä.

**Modes of study.** A separate course is not arranged. Instead the course can be completed by exercises in connection of any course in which the R language is used.

**Arvostelu:** Hyväksytty/Hylätty

**Assessment:** Pass/Fail

## **Tilastotieteen matriisilaskenta (3622310, 6 op)**

### **Matrix Algebra for Statistics**

Vastuuhenkilö: Osmo Kolehmainen

**Sisältö:** Matriisien peruslaskutoimitukset, lineaaristen yhtälöryhmien ratkaiseminen sekä ominaisarvot ja ominaisvektorit ja niihin liittyviä tuloksia.

**Content:** Basic matrix operations, solving linear equations, as well as eigenvalues and eigenvectors, with some related properties and results.

**Tavoitteet:** Kerätessä tilastollista aineistoa syntyy luonnollisesti havaintomatriiseja. Kurssilla on tarkoitus oppia tällaisten matriisien matematiikkaa.

**Objectives:** When statistical data are collected data matrices arise in a natural way. The goal is to learn the mathematics of these matrices.

**Oppimateriaali:** Searle, Matrix algebra useful for statistics, 1982. Luvut 1-9, 11, 11 A, 12.

**Study materials:** Searle, Matrix algebra useful for statistics, 1982. Chapters 1-9, 11, 11 A, 12.

**Toteutustavat:** Luentoja 40 h ja harjoituksia 20 h

**Teaching Methods:** Lectures 40 h and exercises 20 h

**Suoritustavat:** Harjoitustehtävät ja kirjallinen kuulustelu

**Modes of study:** Exercises and written exam

**Arvostelu:** 0 - 5

**Assessment:** 0 - 5

**Sovelluskurssi (3622214, 6 op)**

**Applications of Statistics**

Vastuuhenkilö: Osmo Kolehmainen

**Sisältö:** Kurssin sisältö vaihtelee.

**Content:** The content of the course varies.

**Tavoitteet:** Tarkoituksena on oppia tilastotieteen perusmenetelmiä käytännön sovellutusten avulla.

**Objectives:** The goal is to learn some basic methods of statistics via practical applications.

**Toteutustavat:** Luentoja 40 h

**Teaching Methods:** Lectures 40 h

**Suoritustavat:** Harjoitustyöt

**Modes of study:** Exercises

**Arvostelu:** Hyväksytty/Hylätty

**Assessment:** Pass/Fail

**Edellytykset:** Tilastotieteen peruskurssi 1 ja 2 tai vastaavat opinnot

**Prerequisites:** Basic Course in Statistics 1 and 2 or equivalent

## **Johdatus spatiaaliseen tilastotieteeseen (3622215, 6 op)**

### **Introduction to Spatial Statistics**

Vastuuhenkilö: Osmo Kolehmainen

**Sisältö:** Modernilla spatiaalisella data-analyysillä on kasvava rooli monilla tutkimuksen ja päätöksenteon alueilla. Tästä laajasta aihepiiristä käsitellään kolmea pääaluetta: alueellinen data, pistekuvioiden analysointi ja geostatistiikka.

**Content:** Modern spatial data analysis is playing an increasing role in many areas of research and policy making. Here we present the basic ideas of applied spatial statistics in three main areas: areal data, analysis of point patterns and geostatistics

**Tavoitteet:** Opitaan spatiaalisen tilastotieteen peruskäsitteitä sopivien R –pakettien avulla.

**Objectives:** The goal is to learn basic concepts of spatial statistics by using appropriate R packages.

**Oppimateriaali:** Bailey & Gatrell, Interactive spatial data analysis, 1995

**Study materials:** Bailey & Gatrell, Interactive spatial data analysis, 1995

**Toteutustavat:** Luentoja 40 h

**Teaching Methods:** Lectures 40 h

**Suoritustavat:** Harjoitustyöt

**Modes of study:** Exercises

**Arvostelu:** Hyväksytty/Hylätty

**Assessment:** Pass/Fail

**Edellytykset:** Tilastotieteen peruskurssi 1 ja 2 tai vastaavat opinnot

**Prerequisites:** Basic Course in Statistics 1 and 2 or equivalent

**Spatiaalisen tilastotieteen harjoitustyöt(3622216, 2 op)**

**Applied Project in Spatial Statistics**

Vastuuhenkilö: Osmo Kolehmainen

**Sisältö:** Tavallisia harjoitustehtäviä vaativampia itsenäisiä harjoitustöitä.

**Content:** Course consists of applied assignments that are more demanding than usual exercises.

**Suoritustavat:** Harjoitustyöt

**Modes of study:** Exercises

**Arvostelu:** Hyväksytty/Hylätty

**Assessment:** Pass/Fail

## TILASTOTIETEEN AINEOPINTOJA

Subject Studies in Statistics

### **Todennäköisyyslaskenta ja tilastotiede (3622311, 9 op)**

#### **Probability and Statistics**

Vastuuhenkilö: Eero Korpelainen

**Sisältö:** Todennäköisyyslaskennan soveltaminen tilastotieteeseen, jakaumien tunnusluvut, ehdolliset jakaumat, rajajakaumat, generoivat funktiot, todennäköisyyslaskentaa käytännön esimerkein, simulointi. Kurssilla opetellaan todennäköisyyteen liittyviä matemaattisia valmiuksia runsaalla laskemisella.

**Contents:** Probability and statistical theory, parameters of distributions, conditional distributions, convergence in distributions, moment generating functions, examples and simulation.

**Tavoitteet:** Opintojakson tavoitteena on oppia tilastollisissa analyyseissä tarvittavaa todennäköisyyslaskentaa ja todennäköisyysjakaumia.

**Objectives:** Goal of the course is to learn to use calculation of probability for statistical analysis and some distributions of probability.

**Oppimateriaali:** Casella-Berger, Statistical inference, 1990. Luvut 1-5.

DeGroot, Probability and Statistics, 1986. Luvut 1-5.

**Study materials:** Casella-Berger, Statistical inference, 1990. Chapters 1-5.

DeGroot, Probability and Statistics, 1986. Chapters 1-5.

**Toteutustavat:** Luentoja 52 t ja harjoituksia 24 t.

**Teaching methods:** Lectures 52 h and exercises 24 h.

**Suoritustapa:** Harjoitustehtävät ja tentti.

**Modes of study:** Exercises and exam.

**Arviointi:** 0-5

**Assessments:** 0-5

**Edellytykset:** Tilastotieteen perusopinnot

**Prerequisites:** Basic studies in statistics

Lisätiedot: Vastaa kurssia Matemaattinen tilastotiede 1.

**Johdatus tilastolliseen päättelyyn (3622312, 9 op)**

**Introduction to Statistical Inference**

Vastuuhenkilö: Eero Korpelainen

**Sisältö:** Tilastolliset estimointimenetelmät ja testisuureiden konstruointi. Kurssilla opetellaan kaavojen johtamista ja ymmärtämistä.

**Contents:** Statistical estimation methods and methods of evaluating tests.

**Tavoitteet:** Opintojakson tavoitteena on oppia tilastollisen estimoinnin ja testisuureiden tekemisen periaatteita.

**Objectives:** Goal of the course is to learn principles of statistical estimation and constructions of statistical tests.

**Oppimateriaali:** Casella-Berger, Statistical inference, 1990. Luvut 5-10.

DeGroot, Probability and Statistics, 1986. Luvut 6-9.

**Study materials:** Casella-Berger, Statistical inference, 1990. Chapters 5-10.

DeGroot, Probability and Statistics, 1986. Chapters 6-9.

**Toteutustavat:** Luentoja 52 t ja harjoituksia 24 t.

**Teaching methods:** Lectures 52 h and exercises 24 h.

**Suoritustapa:** Harjoitustehtävät ja tentti.

**Modes of study:** Exercises and exam.

**Arviointi:** 0-5

**Assessments:** 0-5

**Edellytykset:** Todennäköisyyslaskenta ja tilastotiede

**Prerequisites:** Probability and Statistics

Lisätiedot: Vastaa kurssia Matemaattinen tilastotiede 2.

**Monimuuttujamenetelmät (3622324, 6 op)**

**Multivariate Analysis**

Vastuhenkilö: Eero Korpelainen

**Sisältö:** Korrelaatioanalyysi, pääkomponenttianalyysi, faktorianalyysi, ryhmittelyanalyysi, erotteluanalyysi sekä kanoninen analyysi ja näiden suorittaminen tietokoneella.

**Contents:** Principal component analysis, factor analysis, canonical analysis, cluster analysis and discriminant analysis.

**Tavoitteet:** Opintojakson tavoitteena on oppia käsittelemään useampiulotteista havaintoaineistoa joillakin perusmonimuuttujamenetelmillä.

**Objectives:** Goal of the course is to learn to use data with some basic multivariate analysis methods.

**Oppimateriaali:** Dillon-Goldstein, Multivariate analysis, 1984. Luvut 1-3, 5, 9, 10, 11.

**Study materials:** Goldstein, Multivariate analysis, 1984. Chapters 1-3, 5, 9, 10, 11.

**Toteutustavat:** Kontaktiopetusta 40 t.

**Teaching methods:** Lectures 40 h.

**Suoritustapa:** Harjoitustehtävät ja tentti.

**Modes of study:** Exercises and exam

**Arviointi:** 0-5

**Assessments:** 0-5

**Edellytykset:** Tilastotieteen matriisilaskenta ja Todennäköisyyslaskenta ja tilastotiede

**Prerequisites:** Matrix Algebra for Statistics and Probability and Statistics

**Koesuunnittelu (3622320, 6 op)**  
**Design and Analysis of Experiments**  
Vastuuhenkilö: Eero Korpelainen

**Sisältö:** Kurssilla opiskellaan mm. koesuunnittelun merkitystä, varianssianalyysiä, erilaisia koekaavioita, faktorikokeita, hierarkisia kokeita, latinalaisia neliöitä, Taguchin menetelmää ja kokeellista optiminhakua.

**Contents:** Basic knowledge of design of experiments and variance analysis.

**Tavoitteet:** Tavoitteena on, että opiskelija tuntee tilastollisen koesuunnittelun ja kokeellisen tutkimuksen lähestymistavat ja tavallisimmat kokeellisella tutkimuksella kootun aineiston analysointimenetelmät.

**Objectives:** The goal of the course is to provide concepts and methods of basic design of experiments.

**Oppimateriaali:** Montgomery D.C. Design and Analysis of Experiments, 6th edition 2004

**Study materials:** Montgomery D.C. Design and Analysis of Experiments, 6th edition 2004

**Toteutustavat:** Kontaktiopetusta 40 t., ohjattuja harjoituksia ja harjoitustöitä.

**Teaching methods:** Lectures 40 h.

**Suoritustapa:** Harjoitustehtävät

**Modes of study:** Exercises

**Arviointi:** 0-5

**Assessments:** 0-5

**Edellytykset:** Tilastotieteen perusopinnot

**Prerequisites:** Basic studies in statistics

**Tiedon louhinta (3622332, 6 op)****Data Mining**

Vastuhenkilö: Eero Korpelainen

**Sisältö:** Assosiaatiosääntöjä, luokittelumenetelmiä, ryhmittely- ja erotteluanalyysiä sekä joitakin muita tilastollisia menetelmiä.

**Contents:** Association rules, classification methods, cluster and discrimination analysis and some other statistical methods

**Tavoitteet:** Tavoitteena oppia tiedon louhinnan ja sen prosessin käsitteitä ja menetelmiä, aineiston kuvailua, assosiaatiosääntöjä, luokittelua sekä tiedon louhinnan tilastollisia menetelmiä kuten klusterointia, ja erotteluanalyysiä.

**Objectives:** The goal of the course is to learn to some concepts of data mining as association rules, classification methods and some statistical methods.

**Oppimateriaali:** Hand-Mannila-Smyth, Principles of Data Mining, 2001. Han-Kambler, Data Mining, Concepts and Techniques, 2001. Dillon-Goldstein, Multivariate analysis, 1984.

**Study materials:** Hand-Mannila-Smyth, Principles of Data Mining, 2001. Han-Kambler, Data Mining, Concepts and Techniques, 2001. Dillon-Goldstein, Multivariate analysis, 1984.

**Toteutustavat:** Kontaktiopetusta 40 t., ohjattuja harjoituksia ja harjoitustöitä.

**Teaching methods:** Lectures 40 h.

**Suoritustapa:** Harjoitustehtävät

**Modes of study:** Exercises

**Arviointi:** 0-5

**Assessments:** 0-5

**Edellytykset:** Tilastotieteen perusopinnot

**Prerequisites:** Basic studies in statistics

**Aikasarja-analyysi (3622316, 6 op)**

**Time Series Analysis**

Vastuuhenkilö: Eero Korpelainen

**Sisältö:** Kuvailevat tekniikat aikasarja-analyysissä. Aikasarjojen todennäköisyysmallit, erityisesti ARIMA- ja rakennemallit. Ennustusmenetelmät.

**Contents:** Graphical techniques and components of time series. Statistical models of time series, f.ex. ARIMA-models and forecasting methods.

**Tavoitteet:** Tavoitteena oppia aikasarjojen käsittelyn käsitteitä ja perustekniikoita.

**Objectives:** The goal of the course is to learn to some concepts and methods of time series.

**Oppimateriaali:** Chatfield, The Analysis of Time Series, 6th Edition, 2004.

Brockwell & Davis, Introduction to Time Series and Forecasting, 2nd Edition, 2002

**Study materials:** Chatfield, The Analysis of Time Series, 6th Edition, 2004.

Brockwell & Davis, Introduction to Time Series and Forecasting, 2nd Edition, 2002

**Toteutustavat:** Kontaktiopetusta 40 t.

**Teaching methods:** Lectures 40 h

**Suoritustavat:** Harjoitustehtävät ja tentti.

**Modes of study:** Exercises and exam

**Arviointi:** 0-5

**Assessment:** 0-5

**Edellytykset:** Tilastotieteen perusopinnot ja Todennäköisyyslaskenta ja tilastotiede

**Prerequisites:** Basic studies and Probability and Statistics

## **Regressioanalyysi (3622313, 8 op)**

### **Regression Analysis**

Vastuuhenkilö: Osmo Kolehmainen

**Sisältö:** Lineaarinen regressioanalyysi on tilastotieteen perusmetodi, jonka yleistyksiä ovat esimerkiksi lineaariset mallit ja yleistetyt lineaariset mallit. Kurssilla käsitellään estimoinnin, ennustamisen ja hypoteesintestauksen perusteorioita ja niihin liittyvää R-koodia.

**Content:** Linear regression is a basic method of statistics, which can be generalized in many ways, for example the general linear models and the generalized linear models. The course consists of the basic theories of estimation, prediction and hypothesis testing. The relevant R code is also included.

**Tavoitteet:** Opitaan esimerkiksi regressiokertoimien estimointia, testausta, tulkintaa ja estimoidulla mallilla ennustamista sekä diagnostiikkaa, muunnosten käyttöä, selittäjien valintaa ja robustia regressiota.

**Objectives:** The goal is to learn how to estimate, test and interpret regression coefficients; prediction, diagnostics, use of transformations, variable selection and robust regression.

**Oppimateriaali:** Weisberg, Applied linear regression, 1985 (tai uudempi). Luvut 1-9. Faraway, Linear Models with R

**Study materials:** Weisberg. Applied linear regression, 1985 (or newer). Chapters 1-9, Faraway, Linear Models with R

**Toteutustavat:** Luentoja 40 h ja harjoituksia 20 h

**Teaching Methods:** Lectures 40 h and exercises 20 h

**Suoritustavat:** Harjoitustehtävät ja kirjallinen kuulustelu

**Modes of study:** Exercises and written exam

**Arvostelu:** 0 - 5

**Assessment:** 0 – 5

**Edellytykset :** Tilastotieteen matriisilaskenta

**Prerequisites:** Matrix Algebra for Statistics

## **Laskennallinen tilastotiede (3622322, 6 op)**

### **Computational Statistics**

Vastuuhenkilö: Osmo Kolehmainen

**Sisältö:** Laskennallinen tilastotiede on nopeasti kehittyvä alue tilastollisessa tutkimuksessa ja sovelluksissa. Kurssiin sisältyy satunnaislukujen generointi, moniulotteisen datan visualisointi, Monte Carlo –metodit, bootstrap, MCMC ja tiheysfunktion estimointi.

**Content:** Computational statistics is a rapidly expanding area in statistical research and applications. generation of random numbers, visualization of multivariate data, Monte Carlo methods, bootstrap, MCMC and the probability density estimation.

**Tavoite:** Tavoitteena on oppia R-ympäristössä alan perusmetodeja.

**Objectives:** The goal is to learn in the R environment the basic methods of computational statistics.

**Oppimateriaali:** Rizzo, Statistical Computing with R, 2008

**Study materials:** Rizzo, Statistical Computing with R, 2008

**Toteutustavat:** Luentoja 40 h

**Teaching Methods:** Lectures 40 h

**Suoritustavat:** Harjoitustyöt

**Modes of study:** Exercises

**Arvostelu:** 0 - 5

**Assesment:** 0 – 5

**Edellytykset:** Todennäköisyyslaskenta ja tilastotiede (3622311)

**Prerequisites:** Probability and Statistics (3622311)

## **Otantateoria 1 ( 3622326, 6 op)**

### **First Course in Sampling Theory**

Vastuuhenkilö: Osmo Kolehmainen

**Sisältö:** Ositettu otanta, ryväs- ja systemaattinen otanta, moniasteiset otanta-asetelmat sekä suhde-, regressio- ja Horwitz-Thompson estimaattorit.

**Content:** Cluster- and systematic sampling, stratified sampling and multistage designs, as well as the ratio, regression and Horwitz-Thompson estimators.

**Tavoitteet:** Tavoite on oppia tärkeimmät otanta-asetelmat ja estimointimenetelmät. Sovelluksena esitellään esimerkiksi eläinpopulaation koon estimointia.

**Objectives:** The goal is to learn the most important sampling designs and estimation methods. Applications such as the estimation of animal population size are also considered.

**Oppimateriaali:** Thompson, Sampling, 1992

**Study materials:** Thompson, Sampling, 1992

**Toteutustavat:** Luentoja 40 h

**Teaching Methods:** Lectures 40 h

**Suoritustavat:** Harjoitustehtävät ja kirjallinen kuulustelu

**Modes of study:** Exercises and written exam

**Arvostelu:** 0 - 5

**Assessment:** 0 – 5

**Edellytykset:** Tilastotieteen peruskurssi 1 ja 2 tai vastaavat opinnot

**Prerequisites:** Basic Course in Statistics 1 and 2 or equivalent

**Spatiaalinen tilastotiede (3622328, 6 op)**

**Spatial Statistics**

Vastuuhenkilö: Osmo Kolehmainen

**Sisältö:** Koostuu osasuorituksista Johdatus spatiaaliseen tilastotieteeseen (3622215) ja Spatiaalisen tilastotieteen harjoitustyöt (3622216).

**Content:** The courses Introduction to Spatial Statistics and Applied Project in Spatial Statistics together make up this course.

**Suoritustavat:** Harjoitustyöt

**Modes of study:** Exercises

**Arvostelu:** Hyväksytty/Hylätty

**Assessment:** Pass/Fail

## **Biometria (3622318, 6 op)**

### **Biometrics**

Vastuhenkilö: Juha Alho

**Sisältö:** Kurssilla perehdytään epidemiologian tutkimusasetelmiin sekä niissä kerätyn lukumäärätyyppisen aineiston analyysiin suhteellisen riskin malleilla (logistinen, Poisson- ja Cox-regressio). Yleistettyjen lineaaristen mallien estimointiteoriaa. R-ohjelman käyttö.

**Contents:** Epidemiological study designs, and the count type of data that arises from them, is studied from the perspective of relative risk regression techniques (logistic, Poisson and Cox regressions). Theory of generalized linear models is surveyed. R serves as a computational platform.

**Tavoitteet:** Tutustutaan epidemiologian tutkimustyypeihin sekä niissä sovellettaviin tilastollisiin menetelmiin.

**Objectives:** Familiarity with basic epidemiologic study types and statistical models used to analyze data from them.

**Oppimateriaali:** Alho: Lecture notes on Biometrics.

Alho-Spencer, Statistical Demography and Forecasting, luvut 2, 4 ja 5.

Breslow-Day, Statistical Methods in Cancer Research, vols. I and II, 1980, 1987.

Hosmer-Lemeshow, Applied Logistic Regression, 1989.

**Study Materials:** Alho: Lecture notes on Biometrics.

Alho-Spencer, Statistical Demography and Forecasting, Chapters 2, 4 and 5.

Breslow-Day, Statistical Methods in Cancer Research, vols. I and II, 1980, 1987.

Hosmer-Lemeshow, Applied Logistic Regression, 1989.

**Toteutus:** Kontaktiopetusta 40 t. Kurssi luennoidaan englanniksi.

**Teaching methods:** Lectures 40 h. Course is given in English.

**Suoritustavat:** Harjoitustehtävät, välikokeet ja tentti.

**Modes of Study:** Exercises, period exams, final exam.

**Arviointi:** 0-5

**Assessment:** 0-5

**Edellytykset:** Perusopinnojen lisäksi Regressioanalyysi ja Todennäköisyyslaskenta ja tilastotiede.

**Prerequisites:** Basic studies, Regression Analysis and Probability and Statistics.

## **Stokastiset prosessit (3622330, 6 op)**

### **Stochastic Processes**

Vastuuhenkilö: Juha Alho

**Sisältö:** Kurssilla käsitellään tärkeimpiä Markov-prosesseja: Poisson-prosessi, Markovin ketju, Brownin liike. Jatkuvatilaisten prosessien teoriaa esitetään heuristisesti ns. Markov Chain Monte Carlo -tekniikan alkeiden (Gibbs- ja Metropolis-Hastings -otanta) sekä optiokaupassa sovellettavan ns. Black-Scholes -kaavan ymmärtämiseksi.

**Contents:** The most important Markov processes: Poisson process, Markov chain and Brownian motion are discussed. Continuous state processes are covered heuristically with the aim of understanding the basic properties of Markov Chain Monte Carlo techniques (Gibbs and Metropolis-Hastings sampling), and to understand the so-called Black-Scholes formula of option trading.

**Tavoitteet:** Päämääränä on ymmärtää prosessien taustaoletusten merkitystä tilastollisten sovellutusten kannalta sekä oppia prosessien ominaisuuksien todistamisessa tarvittavaa matemaattista ajattelua.

**Objectives:** The goal is to understand the underlying assumptions of the processes from the perspective of statistical applications, and to learn mathematical thinking that is used in proving their properties.

**Oppimateriaali:** Alho: Lecture notes on Stochastic Processes.

Black-Scholes, The Pricing of Options and Corporate Liabilities, 1973.

Çınlar, Introduction to Stochastic Processes, 1975. Luvut 4 ja 5 sekä osia luvuista 1-3 ja 6.

Rozanov, Probability Theory, Random Processes and Mathematical Statistics, 1995. Osia luvuista 2 ja 5.

Williams, Weighing the odds, 2001. Osia luvusta 7.

**Study Materials:** Alho: Lecture notes on Stochastic Processes.

Black-Scholes, The Pricing of Options and Corporate Liabilities, 1973.

Çınlar, Introduction to Stochastic Processes, 1975. Chapters 4 and 5, parts of chapters 1-3 ja 6.

Rozanov, Probability Theory, Random Processes and Mathematical Statistics, 1995. Parts of chapters 2 ja 5.

Williams, Weighing the odds, 2001. Parts of chapter 7.

**Toteutustavat:** Kontaktiopetusta 48 t. Kurssi luennoidaan englanniksi.

**Teaching Methods:** Lectures 48 h. Course is given in English.

**Suoritustavat:** Harjoitustehtävät, välikokeet ja tentti.

**Modes of Study:** Exercises, period exams, final exam.

**Arviointi:** 0-5

**Assessment:** 0-5

**Edellytykset:** Perusopinnot lisäksi Regressioanalyysi ja Todennäköisyyslaskenta ja tilastotiede.

**Prerequisites:** Basic studies, Regression Analysis and Probability and Statistics.

### **Proseminaari (3622315, 8 tai 10 op)**

#### **Proseminar**

Vastuuhenkilö: Juha Alho

#### **Sisältö:**

Proseminaarin ensimmäisellä lukukaudella tutustutaan internetin (mm. Wikipedian) käyttökelpoisuuteen tilastollisen tiedon lähteenä. Toisella kaudella tehdään empiiriseen aineistoon perustuva esitys jostain opiskelijan valitsemasta ongelmasta. Proseminariesitelmä muokataan kandidaatintutkielmaksi (10 op).

#### **Contents:**

During the first semester, students familiarize themselves with the internet as a source of statistical information (e.g. Wikipedia). During the second semester students carry out an empirical analysis on a problem of their choice. The seminar paper serves as a basis for the candidate thesis.

**Tavoitteet:** Proseminaarissa harjoitellaan tilastollisen teorian ja empiiristen tutkimustulosten suullista ja kirjallista esittämistä.

**Objectives:** Oral and written presentation of statistical theory and empirical results is practiced.

**Oppimateriaali:** internet

**Study materials:** internet

**Toteutustavat:** Kummallakin lukukaudella opiskelijat seuraavat kvantitatiivisten menetelmien seminaarin esitelmää, 56 t.

**Teaching Methods:** Students attend the quantitative methods seminar during both semesters, 56 h.

**Suoritustavat:** Opiskelijat kirjoittavat kaksi seminaariesitelmää ja esittävät ne suullisesti.

**Modes of Study:** Students write two papers and present them in the seminar.

**Arvostelu:** Hyväksytty/hylätty.

**Assessment:** Pass/Fail

**Edellytykset:** Perusopintojen lisäksi Regressioanalyysi ja Todennäköisyyslaskenta ja tilastotiede.

**Prerequisites:** Basic studies and Regression Analysis and Probability and Statistics.

### **Tilastotieteen kandidaatintutkielma (3622390, 0 op)**

#### **Bachelor's Thesis in Statistics**

Vastuhenkilö: Juha Alho

**Sisältö:** Proseminariesitelmän pohjalta kirjoitettu tilastotieteellisiä menetelmiä käyttävä tutkielma.

**Contents:** A thesis written on the basis of the proseminar paper that uses statistical methods.

**Tavoite:** Opetellaan tilastollisen teorian ja empiiristen tulosten kirjallista esittämistä.

**Objectives:** Students learn to present statistical theory and empirical results both in writing and orally.

**Suoritustavat:** Pääaineopiskelijan proseminarityönsä pohjalta kirjoittama tutkielma.

**Modes of study:** A thesis written by a Statistics major based on the proseminar paper.

**Arviointi:** 0-5

**Assessment:** 0-5

**Edellytykset:** Proseminaarin suorittaminen

**Prerequisites:** Proseminar completed.

### **Tilastotieteen kandidaatin tutkinnon kypsyysnäyte (3622391, 0 op)**

#### **Bachelor's Essay**

Vastuhenkilö: Juha Alho

Ennen kandidaatintutkielman arvostelua suoritettava kirjallinen kypsyysnäyte työn aihepiiristä.

A short essay on the general area of Bachelor's thesis.

## TILASTOTIETEEN SYVENTÄVIÄ OPINTOJA

Advanced Studies in Statistics

### **Tilastollinen päättely (3622510, 5-9 op)**

#### **Statistical Inference**

Vastuuhenkilö: Juha Alho

**Sisältö:** Kurssi tarjoaa johdatuksen todennäköisyysmalleihin, joita käytetään tilastollisessa päättelyssä. Pääpaino on käsitteissä eikä niinkään matemaattisessa eksaktisuudessa. Laskennallisia tekniikoilla on tärkeä sija. Edellytyksenä kurssille osallistumiselle on, että opiskelijalla on aineopintotason tiedot tilastotieteen teoriasta.

**Contents:** The course provides an introduction to probability models used in statistical inference. The emphasis is on conceptual issues rather than mathematical rigor. Computational techniques are presented. A prerequisite is knowledge of statistical theory at the level of Subject Studies.

**Tavoitteet:** Päämääränä on antaa opiskelijalle riittävä pohja itsenäiseen tilastotieteellisen kirjallisuuden käyttöön.

**Objectives:** The goal is to provide a sufficient basis for independent study of statistical literature.

**Oppimateriaali:** Alho: Lecture notes on Statistical Inference

Cox-Hinkley, Theoretical Statistics, 1974. Valitut kohdat.

Nyblom-Alho, Lecture Notes in Statistical Inference.

**Study materials:** Alho: Lecture notes on Statistical Inference.

Cox-Hinkley, Theoretical Statistics, 1974. Selected topics.

Nyblom-Alho, Lecture Notes in Statistical Inference.

**Toteutustavat:** Kontaktiopetusta 65-70 tuntia sisältäen luentoja ja harjoituksia. Kurssi luennoidaan englanniksi.

**Teaching Methods:** Lectures 65-70 h, including exercise sessions. Course is given in English.

**Suoritustavat:** Suoritustapana harjoitustehtävät ja tentti. Englanninkielisessä maisteriohjelmassa opiskeleva voi suorittaa kurssin suppeampana.

**Modes of study:** Exercises and final exam. Students in an English language Master's Program may complete a shorter version of the course.

**Arviointi:** 0-5

**Assessment:** 0-5

**Edellytykset:** Perusopintojen ja Todennäköisyyslaskenta ja tilastotiede sekä Johdatus tilastolliseen päättelyyn.

**Prerequisites:** Basic studies, Probability and statistics and Introduction to statistical inference.

## **Seminaari (3622511, 10 op)**

### **Undergraduate Seminar**

Vastuhenkilö: Juha Alho

**Sisältö:** Seminaariesitelmä tehdään opiskelijan valitsemasta aiheesta ja sen on tarkoitus muodostaa pohja pro gradu –tutkielmalle.

**Contents:** Seminar paper is written on the topic of the student's choice and it is intended to form a basis for the pro gradu thesis.

**Tavoitteet:** Jatkaen proseminaarin pohjalta, seminaarissa perehdytään yksityiskohtaisemmin tilastollisen teorian ja empiiristen tulosten suullisen ja kirjallisen esittämisen ongelmakohtiin.

**Objectives:** Continuing where the proseminar left off, a deeper discussion of the difficulties of oral and written presentation of statistical theory and empirical results is carried out.

**Toteutustavat:** Kummallakin lukukaudella opiskelijat seuraavat kvantitatiivisten menetelmien seminaarin esitelmiä, 56 t.

**Teaching Methods:** Students attend the quantitative methods seminar during both semesters, 56 h.

**Suoritustavat:** Opiskelijat kirjoittavat kaksi seminaariesitelmän ja esittävät ne suullisesti.

**Modes of Study:** Students write two papers and present them in the seminar.

**Arvostelu:** Hyväksytty/hylätty.

**Assessment:** Pass/Fail

**Edellytykset:** Perusopinnot ja aineopinnot.

**Prerequisites:** Basic studies and Subject Studies

**Tutkielma (3622513, 30 op)**

**Thesis**

Vastuuhenkilö: Tutkielmaa ohjaava professori

**Sisältö:** Tutkielma on kirjallinen opinnäytetyö, jossa opiskelija perehtyy johonkin tilastotieteelliseen ongelmaan ja osoittaa hallitsevansa siinä tarvittavat käsitteet ja menetelmät.

**Contents:** Thesis is a written presentation on a statistical problem, in which the student displays knowledge of the necessary statistical concepts and methods needed for its handling.

**Tavoitteet:** Tilastotieteellisiin kysymyksiin liittyvän tieteellisen kirjoittamisen oppiminen.

**Objectives:** Learning scientific writing on statistical topics.

**Arviointi:** 0-5

**Assessment:** 0-5

## **Bayes-päätely (3622514, 5-8 op)**

### **Bayesian Inference**

Vastuhenkilö: Juha Alho

**Sisältö:** Ehdolliseen todennäköisyyteen perustuvan Bayes-päätelyn peruseriaatteen, esimerkkejä konjugaattisista jakaumaperheistä ja regressioanalyysistä. Gibbs-otannan sekä Metropolis-Hastings-alkotimin soveltaminen.

**Content:** Starting from conditional probability, basic principles of Bayesian inference are derived. Conjugate families of distributions and regression analysis are considered as examples. The use of Gibbs and Metropolis-Hastings sampling.

**Tavoitteet:** Kurssi tarjoaa johdatuksen Bayesläiseen tilastotieteeseen empiirisen tutkimuksen näkökulmasta.

**Objectives:** The aim of the course is to provide an introduction to Bayesian statistics as it is applied to empirical problems.

**Oppimateriaali:** Alho: Lecture notes on Bayesian Inference.

Marin-Robert, Bayesian Core, A Practical Approach to Computational Bayesian Statistics, 2007.

**Study materials:** Alho: Lecture notes on Bayesian Inference.

Marin-Robert, Bayesian Core, A Practical Approach to Computational Bayesian Statistics, 2007

**Toteustavat:** Kontaktiopetusta 40 t. Kurssi luennoidaan englanniksi. Englanninkielisessä maisteriohjelmassa opiskeleva voi suorittaa kurssin suppeampana.

**Teaching Methods:** Lectures 40 h. Course is given in English. Students in an English language Master's Program may complete a shorter version of the course.

**Suoritustavat:** Harjoitustehtävät ja tentti.

**Modes of study:** Exercises and final exam.

**Arviointi:** 0-5

**Assessment:** 0-5

**Edellytykset:** Perusopinnot ja Todennäköisyyslaskenta ja tilastotiede sekä Johdatus tilastolliseen päättelyyn.

**Prerequisites:** Basic studies, Probability and statistics and Introduction to statistical inference.

## **Harjoittelu (3622512, 9 op)**

### **Practical Training**

Vastuuhenkilö: Esko Valtonen

**Sisältö:** Yksityiskohtainen sisältö vaihtelee työtehtävän mukaan. Siihen voi kuulua tehtäviä aineiston hankinnasta aina tilastollisen analyysin raportointiin saakka.

**Content:** Detailed content varies depending on work tasks. It may contain anything from data acquisition to the reporting of statistical analyses.

**Tavoitteet:** Opiskelija perehtyy tilastotieteen soveltamiseen ja tilastotieteilijän toimenkuvaan.

**Objectives:** Student familiarizes him/herself to the practice of statistics and statisticians role.

**Toteutustavat:** Harjoittelu voi tapahtua yksityisen tai julkisen työnantajan palveluksessa esimerkiksi työskentelemällä tutkimusprojekteissa

**Teaching methods:** Training may occur in a private or public enterprise, in connection to a research project, for example.

**Suoritustavat:** tilastotieteen osaamista vaativa työtehtävä ja sen raportointi.

**Modes of study:** Work task requiring statistical knowledge and its reporting.

**Arvostelu:** hyväksytty / hylätty

**Assessment:** accepted / fail

**Edellytykset:** Kandidaatin tutkinto tilastotieteessä tai vastaavat opinnot

**Prerequisites:** Bachelor's degree in statistics or the equivalent.

Lisätiedot:

Ennen harjoitteluksi aiotun työn aloittamista on sen sisällöstä laaditulle suunnitelmalle saatava hyväksyntä vastaavalta opettajalta. Työn päätyttyä opettajalle on annettava selvitys harjoittelusta.

## **Tilastollinen väestötiede (3622516, 5-8 op)**

### **Statistical Demography**

Vastuhenkilö: Juha Alho

**Sisältö:** Kurssi on tarkoitettu johdatukseksi väestötieteeseen tilastotieteilijöille, matemaatikoille, tietojenkäsittelijöille, taloustieteilijöille jne. Ohjelmien PEP ja BEGIN käyttö.

**Contents:** This is an introductory course in demography, aimed at statisticians, mathematicians, computer scientists, economists etc. A brief introduction to stochastic methods of producing population forecasts is given using programs PEP and BEGIN.

**Tavoitteet:** Kurssin päämääränä on perehdyttää väestöä koskeviin tietolähteisiin, väestöilmidiiden kuvaamisessa käytettyihin tilastollisiin malleihin, stokastisiin ennusteisiin sekä niiden soveltamiseen mm. eläkejärjestelmien tutkimuksessa.

**Objectives:** After the course the students will be familiar with sources of population data, fundamental demographic measures and their statistical foundation, and models of population change.

**Oppimateriaali:** Alho: Lecture notes on statistical demography.

Alho-Spencer, Statistical Demography and Forecasting, 2005, luvut 2, 4, 6-9, 11.

**Study materials:** Alho: Lecture notes on statistical demography.

Alho-Spencer, Statistical Demography and Forecasting, 2005, chapters 2, 4, 6-9, 11.

**Toteustavat:** Kontaktiopetusta 40 t. Kurssi luennoidaan englanniksi. Englanninkielisessä maisteriohjelmassa opiskeleva voi suorittaa kurssin suppeampana.

**Teaching Methods:** Lectures 40 h. Course is given in English. Students in an English language Master's Program may complete a shorter version of the course.

**Suoritustavat:** Harjoitustehtävät ja tentti.

**Modes of study:** Exercises and final exam.

**Arviointi:** 0-5

**Assessment:** 0-5

**Edellytykset:** Esitietoja väestötieteestä ei tarvita, sen sijaan edellytetään, että opiskelijalla on tilastotieteen perusopinnojen lisäksi regressioanalyysin ja tilastollisen päättelyn perustiedot.

**Prerequisites:** No prior knowledge of demography is assumed, but it is expected that the students have taken the equivalent of Basic Courses of Statistics, know regression analysis, and understand basic principles of statistical inference.

## **Otantateoria 2 (3622515, 8 op)**

### **Second Course in Sampling Theory**

Vastuuhenkilö: Osmo Kolehmainen

**Sisältö:** Horwitzin ja Thompsonin estimointiteoria ja sen sovelluksia eri otanta-asetelmiin, lisäinformaation käyttö, osapopulaatioiden tunnuslukujen estimointi ja varianssin estimointi monimutkaisissa otanta-asetelmissä.

**Content:** Horwitz-Thompson estimation theory and its applications to various sampling designs, use of auxiliary information, estimation of subpopulation characteristics and estimation of variance in complex designs.

**Tavoitteet:** Kurssilla syvennetään ja laajennetaan monia Otantateoria 1:n käsitteitä sekä esitellään myös uutta materiaalia.

**Objectives:** The course deepens and extends many concepts of the First Course in Sampling Theory and also adds some new material.

**Oppimateriaali:** Särndal-Swensson-Wretman, Model Assisted Survey Sampling, 1997

**Study materials:** Särndal-Swensson-Wretman, Model Assisted Survey Sampling, 1997

**Toteutustavat:** Luentoja 40 h

**Teaching Methods:** Lectures 40 h

**Suoritustavat:** Harjoitustehtävät ja kirjallinen kuulustelu

**Modes of study:** Exercises and written exam

**Arvostelu:** 0 - 5

**Assessment:** 0 – 5

**Edellytykset:** Otantateoria 1 (3622326) ja Regressioanalyysi (3622313)

**Prerequisites:** First Course in Sampling Theory (3622326) and Regression Analysis (3622313)