ipvw

ices

CURSUSSEN STATISTIEK
PERMANENTE VORMING IN DE WETENSCHAPPEN

COURSES IN STATISTICS
CONTINUING EDUCATION IN SCIENCE

$$Var(X) = \sigma^2$$

$$n \geq \left(\frac{2 z_{1-\frac{\alpha}{2}} \sigma}{L} \right)^2$$

$$P(X \in A) = \int_A f(x)\, dx$$

$$\lambda(t) = \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} P(T < t + \Delta t \mid T \geq t)$$

$$\mu = \int_0^\infty S(t)\, dt$$

$$P(B) = P(B|A_1) P(A_1) + P(B|A_2) P(A_2) + \ldots P(B|A_k) P(A_k)$$

$$\lambda(t) = \frac{f(t)}{S(t)}$$

$$P(X \in (a, b]) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

$$P(A^c) = 1 - P(A)$$

UNIVERSITEIT GENT

CURSUSSEN
STATISTIEK
PERMANENTE VORMING IN DE WETENSCHAPPEN

COURSES IN
STATISTICS
CONTINUING EDUCATION IN SCIENCE

ipvw

ices

INDEX

COURSE LOCATIONS

Center for Statistics
Centrum voor Statistiek

Krijgslaan 281 – S9, 9000 Ghent
www.cvstat.UGent.be

Institute for Continuing
Education in Science

This series includes courses that are embedded
in the Doctoral Schools programs.
Visit your DS website for more information.

Instituut voor Permanente
in de Wetenschappen

Deze reeks omvat cursussen die zijn opgenomen
in de UGent Doctoral Schools programma's.
Ga naar de DS website voor meer informatie.

Krijgslaan 281 – S9, 9000 Ghent
T 09 264 44 26 (am)
F 09 264 85 90
ipvw.ices@UGent.be
www.ipvw-ices.UGent.be

1 Fac. of Psychology and Educational
  Sciences (PP) – H. Dunantlaan 1
2 Fac. of Sciences (WE) – Campus Sterre,
  Krijgslaan 281, building S9
3 Adviescentrum voor Studenten
4 Station Gent Sint-Pieters

The power of data and the information therein is entering the heart of almost any section of society. We are discovering that processes can be better understood and controlled, predictions made, causal effects estimated and decisions optimized. Reliable results follow when studies have been appropriately designed, data carefully gathered and analyzed. Scientists and professionals alike add tremendously to their market value when data analytic skills merge their subject matter expertise.

To prepare for 'big data' in tomorrow's world and to face the facts with statistical skills, the Ghent Center for Statistics joins the Institute for Continuing Education of the Faculty of Sciences (ICES) to organize their yearly series of targeted training modules. We aim to provide insight in the basics of statistical research while developing the technical skills to come to results with statistical software. Blended learning with hands-on sessions on PC's or laptops allows participants to gain firsthand experience in applying the knowledge. Our courses target professionals and the academically trained, who wish to refresh their knowledge or discover new areas of research. The program's modular architecture facilitates flexible entry and adaptive training trajectories.

The Flemish Community recognizes the value of lifelong learning for the region's economic development. Employers are granted financial support through the government's introduction of the KMO-portfolio. More about

'MERGE DATA ANALYTIC SKILLS WITH YOUR SUBJECT MATTER EXPERTISE'

this stimulating initiative can be found on the website: *www.kmo-portefeuille.be* (in Dutch). Doctoral Schools support young researchers in acquiring the necessary skills. Several modules can, under certain conditions, be incorporated in the program of the Doctoral Schools.

For the 2014-2015 academic year, the program offers a classic series of basic modules which gradually build up statistical knowledge and techniques, preceded by three courses introducing statistical software SPSS, SAS and R, two commercial and a freeware package, respectively. This is traditionally followed by more specialized courses: 'Bayesian Statistics' starts in January 2015, 'Survival Analysis and Competing Risks' starts in March 2015, 'Multilevel Analysis for Grouped and Longitudinal Data' and 'Data Mining' kick off in April 2015. With exception of Module 2, all modules are taught in English. Classes take place in a pleasant atmosphere with ample opportunity to interact with lecturers. Information about additional short courses is posted on our website *www.ipvw-ices.UGent.be* or you can subscribe to our mailing list. Complementary statistical training and consulting services at Ghent University are found through its Center for Statistics, *www.cvstat.ugent.be* . Wishing you an enjoyable and rewarding learning experience, we look forward to meeting you,

Professor Els Goetghebeur
Director ICES

ADDITIONAL TRAINING AND CONSULTING SERVICES AT GHENT UNIVERSITY

Training

FLAMES

Flames, Flanders training network in Methodology and Statistics (*www.flames-statistics.eu*) is an interuniversity initiative providing further training to young researchers in Ghent and beyond.

One year Master in Statistical Data Analysis

The UGent advanced Master's program in Statistical Data Analysis leads to a dedicated degree following more intensive training in the methods of practical statistics offered to scientists in diverse areas (see page 32).

Consulting

FIRE

The FIRE (Fostering Innovative Research based on Evidence) statistical consulting service offers statistical and methodological support for UGent doctoral students and post-docs during their scientific research. Personalized advice and help is provided during all phases of the research including study design, data collection, statistical analysis, and interpretation and reporting of results.

Book your FIRE consulting slot by filling out the application form at *www.cvstat.ugent.be/FIRE* or, contact us at *fire@ugent.be* for more information.

Stat-Gent CRESCENDO

Stat-Gent CRESCENDO unites statistical expert knowledge of the Ghent University Center for Statistics aiming to support applied research. It provides consulting services in collaboration with the private sector, the public sector, and other research groups.

We offer an operational framework for statistics and data analysis contract work, in flexible formats including data analysis projects, customized training and software solutions. Data analysis projects add maximum value when the statistical method is integrated in the complete trajectory from objective setting to report writing. We therefore strive for stable and sustained relationships with our partners in mutually rewarding research collaborations.

High-quality work is delivered by well-trained and dedicated statistical consultants, under guidance and supervision of UGent statistics professors. Stat-Gent has expertise in a broad range of applications, such as (but not limited to):
• design and analysis of clinical trials
• health economics, epidemiology, electronic health records, quality of care, drug compliance, and evidence-based medicine
• business analytics
• biotech and agriculture
• statistical genetics/genomics: biomarkers,

micro-arrays, qPCR, next-generation sequencing

We use a wide variety of statistical methods going from basic regression, analysis of variance, mixed models and multivariate techniques, to more specific methods in causal analysis, data mining, functional data analysis, experimental design, longitudinal analysis, missing data, multiple testing, robust and non-parametric statistics and survival analysis.
Do not hesitate to contact us with questions or for more information at *statgent@ugent.be*.

INTRODUCTION TO
SAS

Target audience

- This course will benefit professionals and investigators from diverse areas, ranging from investigators working with clinical trial data to data analysts working on high-dimensional data stored in data warehouses.
- It is suited for those who have not used SAS or any other statistical software before, but equally for those who have already worked with SAS but want a better understanding of the capabilities of the SAS programming language.

Teacher

- Ing. Veronique Storme studied biochemistry and obtained a Master in Statistical Data Analysis from Ghent University.
- She is consultant in statistical data analysis at the VIB department of Plant Systems Biology. She has many years of experience programming in SAS.

Course prerequisites

- The course is open to all interested persons. Knowledge of basic statistical concepts is considered an advantage, but not required for learning the SAS programming language.

The amount of data collected and stored in industry as well as society grows exponentially. Data collections range in size from small datasets as seen in preclinical trials to high-dimensional data gathered in data warehouses. Raw data are typically not suitable for immediate analysis. They may be noisy, stored across different tables or require transformation before richer information is readily extracted. Hence, the first task typically involves manipulation of the data: combining tables, selecting relevant data, addressing missing values, modifying and creating variables. Once an adequate analysis dataset is obtained, it becomes possible to apply subsequent statistical analysis techniques (to implement for instance linear regression or survival analysis).

This course aims to empower participants to manipulate data by learning the SAS programming language. SAS offers a unique combination of a complete set of database capabilities, advanced statistical techniques and a powerful programming language. Compared to the interactive SAS Enterprise Guide, using the SAS syntax directly allows total control of data manipulation, more to define model parameters and easier replication of analyses. In addition to data manipulation, we will illustrate by hands-on practice how you can use SAS statistical procedures to subsequently analyze/model your data. Upon finishing this course, you should have a solid foundation for performing data manipulation and basic statistical analyses using the SAS programming language.

Exam

- There is no exam connected to this module. Participants receive a certificate of attendance at the end of the course.
- To qualify for reimbursement from the UGent Doctoral Schools one must attend all classes. (Visit your DS-website for more information).

Course material

- Copies of lecture notes.
- The course is based on the programming manual: "The Little SAS Book: A Primer", L. D. Delwiche and S. J. Slaughter (2012), 5th ed., ISBN 978–1612903439. If you don't already own this book it is highly recommended to order a copy.

Fees

- The registration fee amounts to 325 EUR. Reduced prices apply to students and participants of nonprofit and public services. These prices are available at the ICES website.
- The book can be bought at the additional cost of 55 EUR. Please indicate this clearly on the registration form.

Dates and venue

September 30, October 2, 14 and 16, 2014 from 5.30 pm to 9 pm. Faculty of Science, Building S9, Campus Sterre, Krijgslaan 281, Ghent.

INLEIDING TOT

SPSS

Doelpubliek

- Deze practica zijn gericht op alle personen die gegevens inzamelen en/of opslaan, met de bedoeling deze statistisch te analyseren en interpreteren.

Lesgever

- Kris Erauw (UGent) is staf-medewerker bij de dienst onderwijsondersteuning van de Faculteit Psycho-logie en Pedagogische Wetenschappen aan de Universiteit Gent. Hij stond jarenlang mee in voor de begeleiding van studenten bij de vakken statistiek en methodologie, en bij het schrijven van hun scriptie.
- Daarnaast ondersteunt hij onderzoekers bij het ontwikkelen van een gepast onderzoeksopzet en bij het verwerken van hun data.

Toelatingsvoorwaarden

- Geen.

We leven in een kennismaatschappij. Heel veel mensen verzamelen gegevens of willen bepaalde ideeën met onderzoeksbevindingen ondersteunen. Denk aan de jongeren in een stedelijke jeugdraad die in hun jeugdwerkbeleidsplan de geformuleerde beleidslijnen moeten ondersteunen met onderzoeksbevindingen; of aan de voorzitter van een oudercomité die de standpunten van ouders op een wetenschappelijk verantwoorde manier wil bevragen. Het inzamelen en opslaan van al die gegevens is vaak niet zo evident als het lijkt. Zeker niet als het de bedoeling is de gegevens later op een professionele manier te verwerken.

Deze lessenreeks is er op gericht data in een bruikbare vorm te verzamelen, de ingezamelde data in SPSS op te slaan en met SPSS de eerste beschrijvende statistieken te produceren. De verschillende lessen in de reeks zijn ervaringsgericht opgevat. De deelnemers worden met een aantal problemen geconfronteerd waarna mogelijke oplossingen besproken en gedemonstreerd worden.

Les 1
Data en dataverzameling: data in SPSS invoeren en definiëren, data uit andere programma's importeren en gebruiken.

Les 2
Elementaire bewerkingen: samenvattende statistieken en voorstellingen genereren, variabelen herschrijven en combineren, databestanden bewerken en combineren.

Les 3
Gemiddelden vergelijken: grafische voorstelling van gemiddelden, t-toetsen en one-way variantie-analyse.

Les 4
Meer uit je databestand halen: de eerste stappen in lineaire regressie.

Examen

- Er is geen examen verbonden aan deze module. Deelnemers ontvangen een aanwezig- heidscertificaat aan het einde van de cursus.
- Om in aanmerking te komen voor een terug- betaling door de UGent Doctoral Schools moet de deelnemer aanwezig zijn bij alle lessen. (Ga naar de DS-website voor meer informatie).

Lesmateriaal

- Oefeningenbundel.

Prijs

- De deelnameprijs bedraagt 325 EUR. Gereduceerde prijzen gelden voor studenten en personeel in de non-profit- en overheidssector. Deze vindt u op de IPVW-website.

Data en locatie

- 6, 7, 9 en 10 oktober 2014, telkens van 17u tot 20u.
- pc-klas 1 van de Faculteit Psychologische en Pedagogische Wetenschappen, Henri Dunantlaan 1, Gent.

INTRODUCTION TO **R**

Target audience

- This course targets professionals and investigators from diverse areas with little to no R-programming experience who wish to start using R for their data manipulation, data exploration or statistical analysis.

Teacher

- Joris Meys is a statistical consultant in the Biostatistics Department at the University of Ghent. He is co-author (with Andrie de Vries) of 'R for Dummies' (Wiley, 2012).
- He is an accomplished R programmer, and produces R packages both for specific research projects and, via R-Forge, for more general application. His statistical expertise is in the areas of ecotoxicology, analysis of environmental data, clinical research and meta analysis.

Course prerequisites

- The course is open to all interested persons.
- Knowledge of basic statistical concepts and experience with other programming languages are considered advantages, but not required for learning the R language.

R is an environment for statistical computing and graphics, which is becoming increasingly popular as a tool to get insight in often complex data. While in some ways similar to other programming languages (such as C, Java and Perl), R is particularly suited for data analysis because ready-made functions are available for a wide variety of statistical (classical statistical tests, linear and nonlinear modeling, time-series analysis, classification, clustering, …) and graphical techniques.

The base R program can be extended with user-submitted packages, which means new techniques are often implemented in R prior to being available in other software. This is one of the reasons why R is becoming the de facto standard in certain fields such as bioinformatics (Bioconductor) and financial services.

This course introduces the use of the R environment for the implementation of data management, data exploration, basic statistical analysis and automation of procedures.

The course starts with a description of the R GUI, the use of the command line and an overview of basic data structures. The application of standard procedures to import data or to export results to external files will be illustrated. Creation of new variables, subsetting, merging and stacking of data sets will be covered in the data management section. Exploration of the data by histograms, box plots, scatter plots, summary numbers, correlation coefficients and cross-tabulations will be performed.

Simple statistical procedures that will be covered are:
- comparison of observed group means (t-test, ANOVA and their nonparametric versions) and proportions
- test for independence in 2-way cross tables and linear regression (focusing on the implementation in R of the statistical methods that form the subject of other modules of the statistics course)

Finally, installing new packages and automation of analysis procedures will also be discussed.

Practical sessions and specific exercises will be provided to allow participants to practice their R skills in interaction with the teacher.

Exam
- There is no exam connected to this module. Participants receive a certificate of attendance at the end of the course.
- To qualify for reimbursement from the UGent Doctoral Schools one must attend all classes. (Visit your DS-website for more information).

Course material
- Copies of slides.
- Recommended but optional handbook: "R for Dummies", J. Meys & A. de Vries, 1st ed. (2012), Wiley, ISBN 978–1119962847.

Fees
- The registration fee amounts to 325 EUR. Reduced prices apply to students and participants of nonprofit and public services. These prices are available at the ICES website.
- The book is optional and can be bought at the additional cost of 55 EUR. Please indicate this clearly on the registration form.

Dates and venue
- October 22, 23, 29 and 30, 2014 from 5.30 pm to 9 pm.
- Faculty of Science, Building S9, Campus Sterre, Krijgslaan 281, Ghent.

INTRODUCTORY STATISTICS
BASICS OF STATISTICAL INFERENCE

Target audience

- This course will benefit professionals and investigators from diverse areas, including research scientists and clinical research associates, investing in data handling and wishing to acquire insight into basic statistical methods or to refresh their knowledge and practice of statistics.

Teacher

- Dr. Els Adriaens (Adriaens Consulting bvba) studied biology, obtained a PhD in pharmaceutical sciences and a Master in Statistical Data Analysis at Ghent University.
- She is consultant in statistical data analysis specialized in the field of the development and validation of alternatives to laboratory animals.

Course prerequisites

- The course is open to all interested persons. It is necessary to have an understanding of basic algebra (basic rules, solving equations, ...), exponents and square roots.

Exam

- Participants can, if they wish, take part in an exam. Upon succeeding in this test a certificate from Ghent University will be issued to participants with a university degree at the bachelor level or an equivalent degree.
- To qualify for reimbursement from the UGent Doctoral Schools one must pass the exam. (Visit your DS-website for more information).

This course aims to provide insight into basic statistical concepts with emphasis on practical applications. Mathematical formulae are kept to a minimum. The theory and the methods of analysis will be extensively illustrated with examples relating to a wide variety of different fields.

We start with concise graphical and numerical descriptions of data obtained from observational or experimental studies. The most common and frequently used probability distributions of discrete and continuous variables will be presented. Statistical inference draws conclusions about a population based on sampled data. Chance variations are taken into account such that a level of confidence is attached to these conclusions. We present the reasoning behind significance tests for the comparison of observed data with a hypothesis. We apply this procedure to data obtained either from one or from two populations. The correct use of the t-test will be discussed. Nonparametric methods are considered as a possible alternative in case the requirements of the t-test are not met.

We cover the basic concepts of hypothesis testing for categorical data, including the chi-square test.

Quite often the relationship between two variables, where the outcome of one variable is seen as depending on the value of the other, is the focus of scientific interest. We will give an introduction to linear regression analysis, where a regression line based on observations obtained in a sample describes this relation.

Course material

- Copies of slides.
- Recommended handbooks are:
- Medical field: Book 1: "Fundamentals of Biostatistics", B. Rosner, 7th ed. (2010), Thomson Brooks/Cole, ISBN 978–0538733496. The examples used in this book are restricted to the field of bioscience. The book is therefore recommended if you have a background in a related research area, such as (veterinary) medicine, biotechnology, biology, pharmacy, a.s.o.
- All fields: Book 2: "Introduction to the Practice of Statistics", D.S. Moore, G.P. McCabe and B. Craig, 8th revised ed. (2014), W.H. Freeman, ISBN 978-146415833. This book uses examples from a wide range of research areas and is therefore recommended if you have no background in the research areas mentioned for book 1.

Fees

- The registration fee amounts to 800 EUR. Reduced prices apply to students and participants of nonprofit and public services. These prices are available at the ICES website.
- Both books are optional and can be bought at the additional cost of 95 EUR for book 1 and 90 EUR for book 2. Please indicate this clearly on the registration form.
- The examination fee is 30 EUR.

Dates and venue

- November 4, 18 and 25, December 2, 9, 16 and 23, 2014 from 5.30 pm till 9.30 pm (each lecture is followed by a hands-on practical session in SPSS).
- Faculty of Science, Building S9, Campus Sterre, Krijgslaan 281, Ghent.

BAYESIAN
STATISTICS

Target audience

- This course targets professionals and investigators from diverse areas who wish to get acquainted with Bayesian techniques to be able to apply them to their practical applications.
- This module presents a practical short course on Bayesian Statistics. Participants needing a more elaborate and in depth understanding can follow the second semester course presented by the same instructor in the context of the Master in Statistical Data-Analysis: *www.mastat.ugent.be*

Teacher

- Prof. dr. Dries Benoit is professor of Business Analytics at the faculty of Economics and Business Administration of Ghent University. He teaches Bayesian statistics to the Business Engineering students and in the Master of Statistical Data-Analysis.
- His research is focused on the use of Bayesian statistics for business analytics. He worked on problems ranging from customer relationship management, over credit scoring to pricing and revenue management.
- Methodologically, he contributed to the field of Bayesian quantile regression and Bayesian robustness.

Course prerequisites

- Participants are expected to have an active knowledge of the basic principles underlying statistical strategies, at a level equivalent to the "Introductory Statistics" course of this program.
- Basic knowledge of the statistical programming language R is required.
- In the first session of Module 6 'Analysis of Variance', on January 13, 2015, these principles will be briefly reviewed. This review session is open to interested participants of subsequent modules.

Recent years have seen a tremendous increase in the use and development of Bayesian methods for academic research. In its wake, more and more companies employing statisticians are valuing the knowledge brought by these approaches. The goal of this course is to give participants a brief and intensive introduction to Bayesian statistics.

Participants will learn how Bayesian inference differs from classical inference and how to interpret its results in a meaningful way. They will acquire the skills to use Bayesian techniques correctly in a range of practical applications.

Topics that will be discussed include the difference between Bayesian and frequentist/classical probability, the likelihood function, choice of prior distributions, conjugate priors, the posterior distribution and methods for summarizing the posterior. In addition, an overview will be given about the most important Markov Chain Monte Carlo Methods that are often used to simulate the posterior distribution. These methods include the Gibbs sampler, Importance sampling, Metropolis-Hastings and the Slice sampler.

Depending on the interest and background of the participants, the Bayesian estimation of one (or more) of the following approaches will be explained and discussed: linear regression, choice models (logit, probit, multinomial), Bayesian hypothesis testing, quantile regression, mixed models, Bayesian variable selection, …

All exercises in this course will use R together with the rjags R-package and the JAGS software. Note that JAGS is very similar (if not identical) to the popular BUGS/winBUGS language for Bayesian modeling.

Exam

- Participants can, if they wish, take part in an exam. Upon succeeding in this test a certificate from Ghent University will be issued to participants with a university degree at the bachelor level or an equivalent degree.
- To qualify for reimbursement from the UGent Doctoral Schools one must pass the exam. (Visit your DS-website for more information)..

Course material

- Handouts of slides.
- Recommended books (optional):

- Book 1: Albert, J. (2007). Bayesian Computation with R, Springer, New York (USA), ISBN 978–0387922973. The present course is largely based on this book.
- Book 2: Kruschke, J.K. (2011). Doing Bayesian Data Analysis, Elsevier, Oxford (UK), ISBN 978–0123814852. This book discusses R in more depth but stays very accessible.
- Book 3 – Bernardo J.M. And Smith, A.F.M. (2002). Bayesian Theory, Wiley, New York (USA), ISBN 978–0471494645.

Fees

- The registration fee amounts to 600 EUR. Reduced prices apply to students and participants of nonprofit and public services. These prices are available at the ICES website.
- All three books are optional and can be bought at the additional cost of 50 EUR for book 1, 80 EUR for book 2 and 90 EUR for book 3. Please indicate this clearly on the registration form.
- The exam fee is 30 EUR.

Dates and venue

- January 8, 15, 22 and 29, 2015 from 5.30 pm to 9.30 pm (each lecture is followed by a hands-on practical session).
- Faculty of Science, Building S9, Campus Sterre, Krijgslaan 281, Ghent.
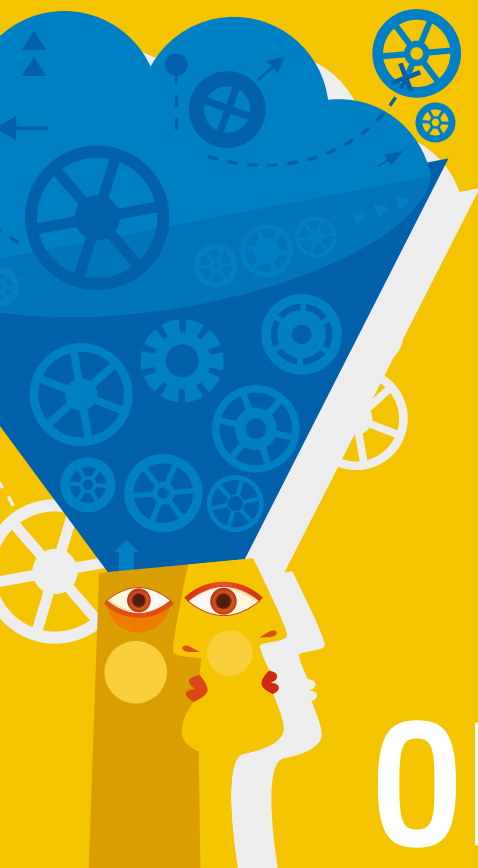
ANALYSIS
OF VARIANCE

Target audience

- This course targets professionals and investigators from diverse areas, who need to use statistical methods in the collection and handling of data in their research, in particular for assessing the effect of e.g. different treatments.

Teacher

- Dr. Els Adriaens (Adriaens Consulting bvba) studied biology, obtained a PhD in pharmaceutical sciences and a Master in Statistical Data Analysis at Ghent University.
- She is consultant in statistical data analysis mainly in the field of the development and validation of alternatives to laboratory animals.

Course prerequisites

- Participants are expected to have an active knowledge of the basic principles underlying statistical strategies, at a level equivalent to the "Introductory Statistics" course of this program.
- In the first session on January 13, 2015, these principles will be briefly reviewed. This review session is open to interested participants of subsequent modules. Participants who have recently followed the introductory course are exempt from that first session.

Analysis of variance (ANOVA) is a statistical tool used in the comparison of means of a random variable in populations that differ in one or more characteristics (factors), e.g. treatment, age, sex, subject, etc.

First, we cover one-way ANOVA, where only a single factor is of concern. Depending on the type of the factor, the conclusions pertain to just those factor levels included in the study (fixed factor model), or to a population of factor levels of which we observed a sample (random effects model).

In two-way and multi-way ANOVA where populations differ in more than one characteristic, the effects of factors are studied simultaneously. This yields information about the main effects of each of the factors as well as about any special joint effects (factorial design). We also consider nested designs, where each level of a second (mostly random) factor occurs in conjunction with only one level of the first factor. One special challenge in multi-way ANOVA lies in verifying the assumptions that must be satisfied.

In this course we will focus on correct execution of data analysis and understanding its results. We pay attention to expressing these conclusions in a correct and understandable way. The different methods will be extensively illustrated with examples from scientific studies in a variety of fields.

Exam

- Participants can, if they wish, take part in an exam. Upon succeeding in this test a certificate from Ghent University will be issued to participants with a university degree at the bachelor level or an equivalent degree.
- To qualify for reimbursement from the UGent Doctoral Schools one must pass the exam. (Visit your DS-website for more information).

Course material

- Handouts of slides.
- Recommended handbook: "Applied Linear Statistical Models", M.H. Kutner, C.J. Nachtsheim, J. Neter and W. Li, 5th ed. (2004), McGraw-Hill, ISBN 978–0071122214.

Fees

- The registration fee amounts to 800 EUR. Reduced prices apply to students and participants of nonprofit and public services. These prices are available at the ICES website.
- The book is optional and can be bought at the additional cost of 75 EUR. Please indicate this clearly on the registration form.
- The examination fee is 30 EUR.

Dates and venue

- January 13, 20 and 27, February 3, 10, 17 and 24, 2015, from 5.30 pm to 9.30 pm (each lecture, except on January 13, is followed by a hands-on practical session in SPSS).
- Faculty of Science, Building S9, Campus Sterre, Krijgslaan 281, Ghent.

SURVIVAL ANALYSIS
AND COMPETING RISKS

Target audience

- Professional statisticians working in applied environments where survival analysis plays a role. This includes biostatisticians working in the pharmaceutical industry, regulatory agencies, or academic centers.
- Researchers and students from epidemiology, biostatistics, statistics, health services research, and related fields wishing to learn about survival analysis and competing risks methodology, to solve applied problems with readily available software.

Teacher

- Dr. Mioara Alina Nicolaie obtained a PhD in Mathematics on the topic of finite Markov chains from Transilvania University of Brașov, Romania, in 2012 and a PhD in Medical Statistics on competing risks from Leiden University Medical Center, the Netherlands, in 2014.
- She worked on medical applications as a post-doctoral researcher at the Catholic University of Louvain since February 2013, developing novel statistical methodology and software for the analysis of time-to-event data, including multi-state data and competing risks with a cure fraction.

Course prerequisites

- Basic knowledge of survival analysis.
- Some experience with the R software.

Analysis of time to event data, often termed survival analysis, is a broad topic with applications across medical sciences, economics, engineering, demography and social sciences.

The purpose of this short course is to present modern survival analysis, with a focus on medical applications. The course starts with an introduction to key principles and statistical methods for survival data studying an inevitable event such as death or eventual failure from any cause. The concepts of non-informative censoring, the hazard and survival functions will be defined and discussed in practical settings along with corresponding estimators. We then build on this to study multiple event types, where some (like death) may preclude others (like disease recurrence) from happening in the future. Concepts then extend to cause-specific hazards and cumulative incidence functions. We develop hypothesis testing and regression modeling for explanatory and prediction purposes, emphasizing the Cox proportional hazards model. In the second part of the course, we turn to some recent developments in competing risks, such as:

• The problem of dynamic prediction of survival outcomes in the presence of time-dependent covariates
• The issue of inference when missing causes of failure show up in competing risks data
• Presence of a cure fraction in the population

Application to real data analyses and their R code implementation will be provided during the course.

Upon completing this course, participants will be able to identify settings in which competing risks modeling is required, to construct and fit an appropriate survival model and correctly interpret the obtained results.

Exam

• Participants can, if they wish, take part in an exam. Upon succeeding in this test a certificate from Ghent University will be issued to participants with a university degree at the bachelor level or an equivalent degree. (Visit your DS-website for more information).
• To qualify for reimbursement from the UGent Doctoral Schools one must pass the exam. (Visit your DS-website for more information).

Course material

• Copies of lecture notes. Recommended (but not required) book: "Survival Analysis Techniques for Censored and Truncated Data", J. Klein and M. Moeschberger (2003), New York: Springer-Verlag. ISBN 978–1441929853.

Fees

• The registration fee amounts to 600 EUR. Reduced prices apply to students and participants of nonprofit and public services. These prices are available at the ICES website.
• The book is optional and can be bought at the additional cost of 100 EUR. Please indicate this clearly on the registration form.
• The examination fee is 30 EUR.

Dates and venue

• March 3, 10, 17 and 24, 2015, from 5.30 pm to 9 pm.
• Faculty of Science, Building S9, Campus Sterre, Krijgslaan 281, Ghent.

APPLIED
LINEAR REGRESSION

Target audience

- This course targets professionals and investigators from all areas who are involved in prediction problems and need to model the relationship between a dependent variable and one or more explanatory variables.

Teacher

- Lizzy De Lobel (Stat-Gent Crescendo, Consulting, UGent) studied Mathematics, Statistical Data Analysis and statistical genomics at Ghent University, where she also worked as a teaching assistant.
- She currently works as consultant for the Stat-Gent consortium. Her experience in teaching and consulting on model building will contribute greatly to this course.

Course prerequisites

- Participants are expected to have an active knowledge of the basic principles underlying statistical strategies, at a level equivalent to the "Introductory Statistics" course of this program.
- In the first session of Module 6 'Analysis of Variance', on January 13, 2015, these principles will be briefly reviewed. This session is free and open to interested participants of this year's program.

Linear regression addresses how a continuous dependent variable is affected by one or more predictors. The fact that many practical problems deal with continuous outcomes (e.g. income, blood pressure, temperature, affect) makes linear regression a popular tool, and most of us will be familiar with the concept of drawing a line through a cloud of data points.

The first two sessions of this module introduce the conceptual framework of this method using the simple case of a single predictor. Formulas and technicalities are kept to a minimum and the main focus will be on interpretation of results and assessing model validity. This includes confidence statements on the predictor effect (hypothesis tests and confidence intervals), using the regression model to predict future results, and verification of model assumptions.

In session 3 and 4 we allow for more than one predictor leading to the multiple linear regression model. We focus on either explanation or prediction. How to come to a parsimonious model starting from a large number of predictors will be discussed in detail. In these complex linear models special attention will be given to interpreting individual predictor effects, as they critically depend on other terms in the model and underlying relations between predictors (confounding).

In the last session a more elaborate data analysis is discussed. We touch on problems where linear regression is not appropriate and replaced by related approaches such as generalized linear models and mixed models.

Different features will be illustrated with case examples from the instructors practical experience, and participants are encouraged to bring examples from their work. Exercises on PC are made with the SPSS software. If preferred, participants can use SAS or R.

Exam

- Participants can, if they wish, take part in an exam. Upon succeeding in this test a certificate from Ghent University will be issued to participants with a university degree at the bachelor level or an equivalent degree.
- To qualify for reimbursement from the UGent Doctoral Schools one must pass the exam. (Visit your DS-website for more information).

Course material

- Copies of lecture notes.
- Recommended handbook: "Applied Linear Statistical Models", M.H. Kutner, C.J. Nachtsheim, J. Neter and W. Li, 5th ed. (2004), McGraw-Hill, ISBN 978–0071122214. Please note that this is the same book as recommended for Module 6 "Analysis of Variance".

Fees

- The registration fee amounts to 800 EUR. Reduced prices apply to students and participants of nonprofit and public services. These prices are available at the ICES website.
- The book is optional and can be bought at the additional cost of 75 EUR. Please indicate this clearly on the registration form.
- The examination fee is 30 EUR.

Dates and venue

- March 5, 12, 19 and 26, April 2, 2015 from 5.30 pm to 9.30 pm (each lecture is followed by a hands-on practical session).
- Faculty of Science, Building S9, Campus Sterre, Krijgslaan 281, Ghent.
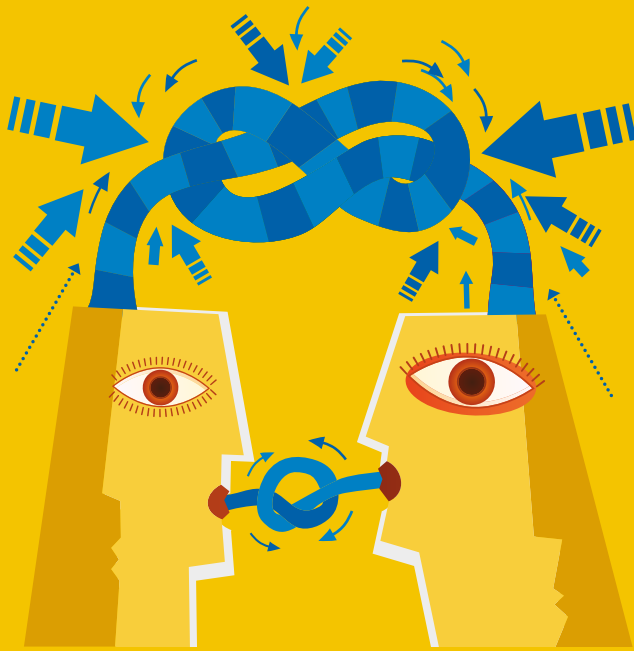
MULTILEVEL ANALYSIS
FOR GROUPED AND LONGITUDINAL DATA

Target audience

- This course targets professionals and investigators from diverse areas ranging from researchers in the behavioral and social sciences to whoever deals with data with a hierarchical or multilevel structure.

Teacher

- Prof. dr. Leoniek Wijngaards-de Meij is Professor in Applied Statistics at the Department of Methodology and Statistics of the Faculty of Social Sciences at Utrecht University, the Netherlands. She received a Master in Clinical Psychology and a PhD in Dyadic processes of parents grieving their child in Clinical Psychology at Utrecht University.
- She has worked as statistical consultant on multilevel analysis for numerous social science research projects.
- She has been teaching courses on Multilevel Analysis at both the Graduate and Undergraduate level. These courses include Multilevel for Research Masters, ML Minor courses and a Multilevel Summer School at Utrecht University, and PhD courses for several institutes including the KLI (Research Institute for Social Psychology), the EPP (Research Institute for PsychoPathology) and the Erasmus University.

Course prerequisites

- The course assumes reasonable familiarity with analysis of variance and multiple regression analysis, but prior knowledge of multilevel modeling is not assumed.

Social research often concerns relationships between individuals and the social contexts to which they belong. This can be conceptualized as a hierarchical structure, with individuals nested within groups. Classical examples are educational research with pupils nested within schools, and cross-national research with individuals nested within their national units. They involve two level data: group level and individual level variables. We need multilevel modeling to study the relationships between variables observed at different levels in the hierarchical structure. This can also cover longitudinal research, by viewing measurement occasions as nested within respondents, and extends to situations where data have a more complex multilevel structure, such as cross-classified data or multiple-membership models.

This short course is intended as a basic and nontechnical introduction to multilevel analysis. It starts with a description of some examples, and shows why multi-level models are necessary if the data have a hierarchical structure. It then covers the basic theory of two- and three-level models.

Next it explains how multilevel models can be applied to analyze longitudinal data, and why and when this may be an attractive analysis approach, as compared to more classical analysis methods such as multivariate analysis of variance (Manova). The course includes three computer labs, where multi-group and longitudinal data are analyzed. The computer labs in the course use the multilevel program HLM and the SPSS Mixed procedure, which is available in SPSS starting with version 11.5.

Exam

- Participants can, if they wish, take part in an exam. Upon succeeding in this test a certificate from Ghent University will be issued to participants with a university degree at the bachelor level or an equivalent degree.
- To qualify for reimbursement from the UGent Doctoral Schools one must pass the exam. (Visit your DS-website for more information).

Course material

- Copies of course notes.
- The course is based on: "Multilevel Analysis. Techniques and Applications", J.J. Hox (2010), 2nd ed., New York: Routledge, ISBN 978–1848728462.

Fees

- The registration fee amounts to 900 EUR. Reduced prices apply to students and partici-pants of nonprofit and public services. These prices are available at the ICES website.
- The book is optional and can be bought at the additional cost of 55 EUR. Please indicate this clearly on the registration form.
- The examination fee is 30 EUR.

Dates and venue

- April 15, 16 and 17, 2015, from 9 am till 4 pm.
- Faculty of Psychology and Educational Sciences, Dunantlaan 1, Ghent.
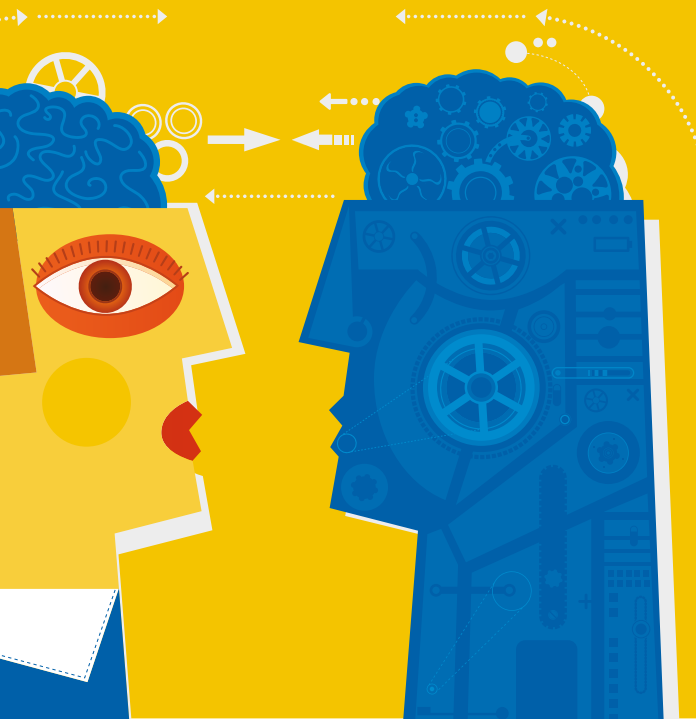
DATA
MINING

Target audience

- This course targets professionals and investigators from all areas that are involved in prediction problems based on large and/or high-dimensional databases.

Teacher

- Prof. Dr. Arnout Van Messem is Guest Professor at Ghent University, Department of Applied Mathematics, Computer Science and Statistics, Faculty of Science. He obtained his PhD at the VUB on the topic of Support Vector Machines and held postdoctoral positions at the KULeuven and the UGent.
- His research interests lie in the domains of robustness, nonparametric statistics and machine learning. He teaches statistical courses to students at the Faculty of Science and the Master in Statistical Data Analysis.

Course prerequisites

- Participants are expected to be familiar with linear regression as for instance taught in Module 8 of this program.

The increased availability of computer storage space has lead to the collection of large databases in high dimensions. Also modern techniques lead to high-dimensional data sets such as genetic data.

This course focuses on modern statistical learning techniques to construct prediction rules based on such large, high-dimensional data. The outcome can be continuous (regression) or categorical (classification). The relation between bias, variance and model complexity will be explained.

The course will discuss penalty methods, such as ridge regression and Lasso, local methods, such as nearest neighbors and kernels, splines, mixture models, classification and regression trees, and support vector machines.

Techniques such as cross-validation, bagging and boosting will be introduced as well. The curse of dimensionality will be explained and illustrated, and dimension reduction techniques such as principal component analysis will be discussed.

Exam

- Participants can, if they wish, take part in an exam. Upon succeeding in this test a certificate from Ghent University will be issued to participants with a university degree at the bachelor level or an equivalent degree.
- The exam will consist in completing an individual project with an oral defense at a date which will be specified during the course.
- To qualify for reimbursement from the UGent Doctoral Schools one must pass the exam. (Visit your DS-website for more information).

Course material

- Course notes and copies of the slides.
- Recommended handbook: The course is based on "The Elements of Statistical Learning: Data Mining , Inference and Prediction," Hastie, T., Tibshirani, R., and Friedman, J. 2nd ed. (2009), Springer-Verlag, ISBN 978–0387848570. This handbook is not required to participate in the course and can be downloaded free of charge from *www-stat.stanford. edu/~tibs/ElemStatLearn*.

Fees

- The registration fee amounts to 900 EUR. Reduced prices apply to students and participants of nonprofit and public services. These prices are available at the ICES website.
- The examination fee is 30 EUR.

Dates and venue

- April 22, 23, 29 and 30, May 6 and 7, 2015, from 5.30 pm to 9 pm.
- Faculty of Science, Building S9, Campus Sterre, Krijgslaan 281, Ghent.

PRACTICAL
INFORMATION

Please use the registration form on our website:
www.ipvw-ices.UGent.be.

- Your registration is valid from the moment you receive an e-mail confirmation from ICES. If you have not received this mail within a week, please contact ICES to double check. From the moment the confirmation e-mail is sent the payment and cancellation conditions are in effect.

- The registration fee covers tuition, some or all of the course materials, use of auditoria and PCs, drinks and sandwiches. Reduced prices apply to students and participants of nonprofit and public services. These prices are available on the ICES website.

- The examination fee for each module that has an exam connected to it is 30 EUR.

- Visit the ICES website further for information on the payment and cancellation conditions, additional reductions for participants from the non-profit and private sector, support measures from the governement, the Doctoral Schools, and how to stay informed about future courses.

2014-2015

UGent celebrates 10 years of successful learning in the one year Master in Statistical Data Analysis

http://www.mastat.ugent.be