Preface

It is my pleasure to welcoming you to the Leuven STATistics STATe of the Art Training Initiative, a scientific and educational project of the Leuven Statistics Research Centre (LStat), offering a wide range of short courses.

Statistics in Leuven is varied with roots in various disciplines. Statisticians are active throughout the university, in mathematics, computer science, economy, psychology, education, bio-engineering, engineering, biology, chemistry, medicine, pharmacy, physical education, psychology, social science, linguistics, etc. Many colleagues combine an excellent international scientific reputation with highly effective teaching skills. At the same time, statistical consulting for internal and external clients is a substantial component of LStat’s mission as well.

It is therefore not surprising that the short course programme has been highly successful and in great demand. Building on the success of previous years, we have further expanded the programme with new courses on highly relevant topics, many located at the heart of our faculty’s expertise. Due to increasing demand, some courses are offered more than once per academic year.

A selected set of courses is offered in an open educational concept, in the sense that, for example, also contingents of students of our highly successful MSc in Statistics partake in them. This ensures stimulating interaction.

Courses take place in one of the university’s campuses, dotted around the beautiful college town of Leuven.

Should your company or institute be looking for a tailor-made training initiative, perhaps on-site, then we will be delighted to explore options and work towards an individualized proposal.

Professor Dirk Cattrysse
2017-2019 Chair of LStat

Leuven Statistics Research Centre
Celestijnenlaan 200 B bus 5307
3001 Heverlee
tel. + 32 16 32 88 75
www.lstat.kuleuven.be
Presenters

Kris Bogaerts is project manager external consultancy at the Leuven Biostatistics and Statistical Bioinformatics Centre (L-BioStat) of the KU Leuven in Belgium. He studied mathematics (1994) at the KU Leuven, received his Master in Biostatistics (1995) at Universiteit Hasselt and his PhD in Science (statistics) at KU Leuven in 2007. He has worked as a statistical consultant in a great variety of medical applications. He is co-author of more than 65 peer reviewed research articles. Dr. Bogaerts is a member of the Belgian Statistical Society.

An Carbonez is professor at the Leuven Statistics Research Centre (LStat) of the KU Leuven. She received her PhD in Mathematics at the same university in 1992. She is the coordinator of the Master in Statistics program of the KU Leuven. She is also involved in statistical consulting projects and gives tailor made statistical courses in internal and external divisions and companies.

Anne-Marie De Meyer is Professor emeritus at the KU Leuven in the Faculty of Science, Department of Mathematics. She received her PhD in Mathematics (Applied Probability) in 1979. For more than 25 years she has been involved in the short course program for a variety of courses in applied statistics and she continues to be active in the statistical consulting of LStat.

After a 10 year stint as an officer in the army, Jan De Spiegeleer accumulated over 20 years of experience in the financial markets. Before founding RiskConcile in 2015, he was based both in London and Brussels, where he was a managing director at KBC Financial Products. He was responsible for the equity derivatives desk (trading and structuring). After that he took an assignment as a risk manager and became in 2007 head of risk management at Jabre Capital Partners, a Geneva-based hedge fund management company. Jan De Spiegeleer is also part of the Statistics Department within Mathematics (KU Leuven).

Peter Goos is a full professor at the Department of Biosystems of the Faculty of Bioscience Engineering of the University of Leuven and the Department of Environment, Technology and Technology Management of the Faculty of Applied Economics of the University of Antwerp. He has been a guest professor at the Econometric Institute of the Erasmus School of Economics (Erasmus University of Rotterdam), the Faculty of Business and Economics of the University of Leuven, the Antwerp Management School and the International School of Management in Saint-Petersburg. Peter Goos has received the Shewell Award and the Lloyd S. Nelson Award of the American Society for Quality, the Ziegel Award and the Statistics in Chemistry Award of the American Statistical Association, and the Young Statistician Award of the European Network for Business and Industrial Statistics. In 2013, Peter Goos was ranked 7th in the top 40 of economists in the Netherlands.

Kristel Hoydonckx works within the department “Facilities for Research” of the KU Leuven and is among others responsible for the survey service.

Marlies Lacante is since 1974 part of the Psychology Department of the KU Leuven. For more than 20 years she was and remains involved in the statistical education within Psychology. She is currently teaching as an associate professor at the academic teacher training cell, the Leuven Statistics Research Centre (LStat) and the Master in Psychology. She is also active in education-research, focusing on survey research and with special attention to the research methodology.
Emmanuel Lesaffre is Professor of Biostatistics at the KU Leuven and U Hasselt in Belgium. He also holds a honorary professorship at the university of Erasmus, Rotterdam, the Netherlands. He studied mathematics at the Universiteit Antwerpen, and received his Doctorate of Science at KU Leuven in 1986. The statistical research of Dr. Lesaffre deals with hierarchical and clustered data with a focus on longitudinal studies, interval censoring, missing data problems, Bayesian methods and in general statistical methods in clinical trials. He has worked in a great variety of medical applications, with focus on research in oral health, cardiology and nursing studies. Dr. Lesaffre is the founding chair of the Statistical Modeling Society, and is a past-president of the International Society of Clinical Biostatistics. In addition he started up a bi-annual international conference on statistical methods in oral health. He is one of the three founding editors of Statistical Modeling and has been Associate Editor of Biometrics, and is currently Associate Editor of Biostatistics. He has worked in a great variety of medical applications, with focus on research in oral health, cardiology and nursing studies. Dr. Lesaffre is the founding chair of the Statistical Modeling Society, and is a past-president of the International Society of Clinical Biostatistics. In addition he started up a bi-annual international conference on statistical methods in oral health. He is one of the three founding editors of Statistical Modeling and has been Associate Editor of Biometrics, and is currently Associate Editor of Biostatistics.

José Nunez Ares is since 2014 a PhD candidate at KU Leuven under the supervision of prof. Peter Goos. He holds a bachelor and master degrees in civil engineering (La Coruña, Spain, 2001). After working for several years in the private sector in Spain and in The Netherlands, in 2014 he completed a master in Operations Research at Erasmus University. In his research he combines techniques from mathematical optimization and experimental design. The main focus of his research lies on trend-robust response surface designs and the discovery of new efficient response surface designs. His first paper in an international peer reviewed journal was accepted for publication in April 2016.

Geert Molenberghs is Professor of Biostatistics at the Universiteit Hasselt and KU Leuven in Belgium. He received the B.S. degree in mathematics (1988) and a PhD in biostatistics (1993) from the Universiteit Antwerpen. Dr Molenberghs published methodological work on surrogate markers in clinical trials, categorical data, longitudinal data analysis, and on the analysis of non-response in clinical and epidemiological studies. He served as Joint Editor for Applied Statistics (2001-2004), Co-editor for Biometrics (2007-2009) and as President of the International Biometric Society (2004-2005). He currently is Co-editor for Biostatistics (2010-2013). He was elected Fellow of the American Statistical Association and is an honorary member of the Society of Clinical Biostatistics. Dr. Lesaffre started up the Biostatistical Centre, a predecessor of the Leuven Institute for Biostatistics and statistical Bioinformatics. He was also chair of the department of Biostatistics at Erasmus MC from 2007 to 2014.

Peter Rousseeuw is a statistician known mainly for his work on robust statistics and cluster analysis. He was professor at the Delft University of Technology, The Netherlands, at the University of Fribourg, Switzerland, and at the University of Antwerp, and worked in the financial industry in New York. Currently he is professor at the KU Leuven. He has authored three books and over 170 papers.

Francis Tuerlinckx is Professor of Psychology at the KU Leuven in Belgium. He received the Master degree in psychology (1996) and a PhD in psychology (2000) from the KU Leuven. He was a postdoc at the Department of Statistics of Columbia University (New York). In general, Francis Tuerlinckx' research deals with the mathematical modeling of various aspects of human behavior. More specifically, he works on item response theory, reaction time modeling, and dynamical systems data analysis for human emotions.
**Presenters**

**Seppe vanden Broecke** received a PhD in Applied Economics at KU Leuven, Belgium in 2014. Currently, Seppe is working as an assistant professor in data and process science at the department of Decision Sciences and Information Management at KU Leuven. Seppe’s research interests include business data mining and analytics, machine learning, process management, process mining. His work has been published in well-known international journals and presented at top conferences.

**Katrijn Van Deun** is assistant professor in Methodology and Statistics at Tilburg University and a research fellow of the KU Leuven. She obtained a Master in psychology, a Master’s degree in statistics and a PhD in psychology. Her main area of expertise is scaling, clustering and component analysis techniques, which she applies in the fields of psychology, chemometrics and bioinformatics. She has various publications in both methodological and substantive journals in psychometrics, chemometrics and bioinformatics. Katrijn is secretary of the Dutch/Flemish Classification society.

**Geert Verbeke** is Professor in Biostatistics at KU Leuven and Universiteit Hasselt. He received the B.S. degree in mathematics (1989) from the KU Leuven, the M.S. in biostatistics (1992) from Universiteit Hasselt, and earned a Ph.D. in biostatistics (1995) from the KU Leuven. Geert Verbeke has published extensively on longitudinal data analyses. He has held visiting positions at the Gerontology Research Center and the Johns Hopkins University (Baltimore, MD). Geert Verbeke is Past President of the Belgian Region of the International Biometric Society, International Program Chair for the International Biometric Conference in Montreal (2006). He was Board Member of the American Statistical Association (2008-2010) and currently is member of the Council of the Royal Statistical Society (2013-2016). He is past Joint Editor of the Journal of the Royal Statistical Society, Series A (2005–2008) and editor of Biometrics (2010–2013). He is the director of the Leuven Center for Bio-statistics and statistical Bioinformatics (L-BioStat), and vice-director of the Interuniversity Institute for Biostatistics and statistical Bioinformatics (I-BioStat), a joint initiative of the Hasselt and Leuven universities in Belgium.

**Jan Wijffels** is the founder of www.bnosac.be - a consultancy company specialised in statistical analysis and data mining. He holds a Master in Commercial Engineering, a MSc in Statistics and a Master in Artificial Intelligence and has been using R for 10 years, developing and deploying R-based solutions for clients in the private sector. He has developed and co-developed the R packages fbase, ETLtiles, RMOA and RMyrrix.

**Tom Wilderjans** is a post-doctoral researcher at the Fund for Scientific Research (FWO-Flanders). He obtained a Master’s degree (2005) and a PhD (2009) in Mathematical Psychology from the KU Leuven. Tom’s research deals with multivariate data analysis (component analysis, clustering, and combinations thereof) and model selection.
Essential tools for R

Course outline
This course gives an introduction to the use of the statistical software language R. R is a language for data analysis and graphics. This introduction course to R is aimed at beginners. The course covers data handling, graphics, mathematical functions and some statistical techniques. R is for free and for more information you can visit the site http://cran.r-project.org/

Target audience
Everybody who is interested in using the R programming language. You will learn how to write and manage your R scripts.

Prerequisites
There are no prerequisites.

Presenter
Dominik Sznajder

Course Material
The course material will be made available before the start of the course.

Dates
- 26, 27, 28 September and 2, 3 October 2017, 18.00 – 21.00 (max. 20 participants)
- 8, 9, 10, 13 and 14 November 2017, 18.00 – 21.00 (max. 20 participants)
- 19, 20, 21, 26 and 27 February 2018, 18.00 – 21.00 (max. 20 participants)

Language
- English

Price
- (PhD) students KU Leuven and Association KU Leuven € 125
- Staff KU Leuven and Association KU Leuven, other (PhD) students € 200
- Non profit/social sector € 312,50
- Private sector € 750
Fundamentele statistische methoden

Beschrijving


INHOUD VAN DE CURSUS

- Beschrijvende grootheden: grafische en numerische samenvatting van de data
- Verdelingen: Binomiale, Poisson, Normale, T-verdeling
- Steekproefverdeling van het gemiddelde
- Betrouwbaarheidsintervallen
- Hypothese testen omtrent een gemiddelde (één en twee steekproeven)
- Gepaarde t-test
- Schatten en testen van proporties

Doelgroep

Iedereen die een opfrissing van fundamentele statistische technieken wenst.

Voorkennis

Er wordt geen voorkennis verondersteld.

Lesgever

Prof. Marlies Lacante

Cursusmateriaal

Cursusmateriaal wordt ter beschikking gesteld.
Fundamental statistical methods, applications with R

Course outline
By using cases, one explores data by using R. Attention is paid to the interpretation of the output. Topics as exploring data, construction of confidence intervals and hypothesis testing is covered. This is a hands-on session.

Target audience
Everybody who wants to explore data by using R

Prerequisites
Fundamental Statistical Methods (distributions, confidence intervals, hypothesis testing) and Introduction to R.

Presenter
N.

Course Materials
The course material will be made available.

Date
- 9 October 2017, 09.00 – 12.00 and 13.00 – 16.00

Language
- English

Price
- (PhD) students KU Leuven and Association KU Leuven € 50
- Staff KU Leuven and Association KU Leuven, other (PhD) students € 80
- Non profit/social sector € 125
- Private sector € 300
Models for Longitudinal and Incomplete Data

CONCEPTS, MODELS AND HANDS-ON APPLICATION WITH THE OPTION TO ANALYSE ONE’S OWN DATA

Course outline
This is a web lecture with introductory and Q & A sessions. We first present linear mixed models for continuous hierarchical data. The focus lies on the modeler’s perspective and on applications. Emphasis will be on model formulation, parameter estimation, and hypothesis testing, as well as on the distinction between the random-effects (hierarchical) model and the implied marginal model. Apart from classical model building strategies, many of which have been implemented in standard statistical software, a number of flexible extensions and additional tools for model diagnosis will be indicated. Second, models for non-Gaussian data will be discussed, with a strong emphasis on generalized estimating equations (GEE) and the generalized linear mixed model (GLMM). To usefully introduce this theme, a brief review of the classical generalized linear modeling framework will be presented. Similarities and differences with the continuous case will be discussed. The differences between marginal models, such as GEE, and random-effects models, such as the GLMM, will be explained in detail. Third, it is oftentimes necessary to consider fully non-linear models for longitudinal data. We will discuss such situations, and place some emphasis on the non-linear mixed-effects model. Fourth, non-linear mixed models will be discussed. Applications in the PK/PD world will be brought to the front. Fifth, when analyzing hierarchical and longitudinal data, one is often confronted with missing observations, i.e., scheduled measurements have not been made, due to a variety of (known or unknown) reasons. It will be shown that, if no appropriate measures are taken, missing data can cause seriously jeopardize results, and interpretation difficulties are bound to occur. Methods to properly analyze incomplete data, under flexible assumptions, are presented. Key concepts of sensitivity analysis are introduced. All developments will be illustrated with worked examples using the SAS System. However, the course is conceived such that it will be of benefit to both SAS users and users of other platforms.

Prerequisites
Throughout the course, it will be assumed that the participants are familiar with basic statistical modeling concepts, including linear models (regression and analysis of variance), as well as generalized linear models (logistic and Poisson regression) and basic knowledge of mixed and multilevel models. Moreover, pre-requisite knowledge should also include general estimation and testing theory (maximum likelihood, likelihood ratio). When registering for this course, you have to mention the topics you have followed before and/or indicate where you became acquainted with the requested material.

Presenters
Geert Verbeke and Geert Molenberghs

Course Materials
The course material will be made available.

Background reading

Target audience
The targeted audience includes methodological and applied statisticians and researchers in industry, public health organizations, contract research organizations, and academia.

Important: The course will also serve for the MSc in Statistics students.
LEUVEN STATISTICS STATE OF THE ART TRAINING INITIATIVE

Dates

- 9 October 2017: 13.00 – 14.00 (Introductory)
- 17 October 2017: 09.00 – 11.00 (Q & A)
- 8 November 2017: 09.00 – 11.00 (Q & A non-Gaussian LDA)
- 1 December 2016: 09.00 – 11.00 (Q & A missing data)

Web lectures have to be watched prior to the Q & A session(s)!

Language

- English

Price

- (PhD) students KU Leuven and Association KU Leuven € 250
- Staff KU Leuven and Association KU Leuven, other (PhD) students € 400
- Non profit/social sector € 625
- Private sector € 1500
Statistical Machine Learning with R

Course outline

This course is a hands-on course covering the use of statistical machine learning methods available in R. The following basic learning methods will be covered and used on common datasets.

- classification trees (caret)
- feed-forward neural networks and multinomial regression
- random forests
- boosting for classification and regression
- bagging for classification and regression
- penalized regression modelling (lasso/ridge regularized generalized linear models)
- model based recursive partitioning (trees with statistical models at the nodes)
- training and evaluation will be done through the use of the caret and ROCR packages

The course will cover the techniques from a high-level viewpoint, useful for day-to-day R users.

Target audience

The course is for R users in industry/academics who are interested in building predictive models in R which have some experience with regressions but have less knowledge of machine learning and techniques of artificial intelligence.

Also persons interested in the statistical learning techniques itself will find this course useful. Or people with a data science background with less knowledge of R and which are interested in machine learning in general.

Prerequisites

Initial experience in R ranging from a few weeks to several years (at least understanding of the course ‘Essential Tools for R’ is needed). Some practical experience in regression modelling.

Course materials

The course material will be made available.

Presenter

Jan Wijffels
Fundamentele statistische methoden, toegepast met SPSS

Beschrijving

Dit is een inleidende cursus tot het gebruik van SPSS. Aan de hand van cases wordt geïllustreerd hoe men met SPSS tot exploratie van gegevens komt. Hierbij wordt de nodige aandacht besteed aan het interpreteren van de verkregen output. Hypothesetesten voor onafhankelijke en gepaarde groepen worden uitgevoerd en besproken. Er is tijd om zelf te werken met deze software.

Doelgroep

Iedereen die gegevens wenst te exploreren met SPSS.

Lesgever

Prof. Marlies Lacante

Voorkennis

De technieken die aangeleerd werden bij Fundamentele Statistische Methoden.

Cursusmateriaal

Cursusmateriaal wordt ter beschikking gesteld.

Datum

• 19 oktober 2017, van 9 tot 12 u. en van 13 tot 16 u.

Taal

• Nederlands

Prijs

• (PhD) studenten van KU Leuven en Associatie KU Leuven € 75
• Personeel KU Leuven en Associatie KU Leuven, andere (PhD) studenten € 80
• Non profit/sociale sector € 125
• Private sector € 300
Regression and Analysis of Variance

Course outline

Linear statistical models are widely used today in many applications. Successfully applying these techniques require a good understanding of the underlying theory and the practical problems that you may encounter in real-life situations.

DAY 1: REGRESSION ANALYSIS

• Correlation
• Simple linear regression
  Ordinary least squares: estimating parameters, confidence intervals and tests, diagnostics, prediction.
• Multiple regression
  Ordinary least squares: estimating parameters, confidence intervals and tests, diagnostics, prediction.
  Variable selection techniques.

DAY 2: ANALYSIS OF VARIANCE

• One-way Anova
  Comparing means
  The Anova model: estimating parameters, hypothesis tests, Anova tabel, F test
  Multiple comparisons, Contrasts
• Two-way Anova
  The two-way Anova Model
  Main effects, interaction effects
  Multiple comparisons

DAY 3

• Detection of outliers and influential observations
• Multicollinearity
• Analysis of covariance

Target audience

This course is important for persons involved with modeling data.

Prerequisites

Participants are familiar with basic statistical modeling concepts (see topics described in Fundamental Statistical Methods).

Presenters

TBA

Course Materials

The course material will be made available.
Text Mining with R

Course outline

This course is a hands-on course covering the use of text mining tools for the purpose of data analysis. It covers basic text handling, natural language engineering and statistical modelling on top of textual data.

The following items are covered:
- Cleaning of text data, regular expressions
- String distances
- Graphical displays of text data
- Natural language processing: stemming, parts-of-speech tagging, tokenization, lemmatisation
- Sentiment analysis
- Statistical topic detection modelling (latent dirichlet allocation)
- Automatic classification using predictive modelling based on text data
- Visualisation of correlations & topics
- Word embeddings
- Document similarities & Text alignment

Target audience

The course is for R users in industry/academics who are interested in practical natural language processing and statistical learning on text data. People with a data science background with less knowledge of R and which are interested in machine learning & text mining in general will find this course, alongside the course on Statistical Machine Learning with R very useful.

Prerequisites

Initial experience in R ranging from a few weeks to several years (at least understanding of the course ‘Essential Tools for R’ is needed, preferably also ‘Advanced R programming topics’). Some practical experience in regression modelling. In order to follow the parts on topic modelling and automatic classification of text data, you should have knowledge of statistical modelling.

Course materials

The course material will be made available.

Presenter

Jan Wijffels
Regression and Analysis of Variance, applications with SPSS

Course outline

This is a hands-on session where by means of cases we illustrate how to run linear models in SPSS.

The topics given in the course ‘Regression and Analysis of Variance’, are now applied on examples by using SPSS. We also stress on the correct interpretation of the output. There is time to practice with the software.

- Day 1, morning + 1 hour in the afternoon: Applications and exercises on regression analysis (linear regression model, diagnostics, multicollinearity)
- Day 1, afternoon + day 2 morning: Applications and exercises on analysis of variance (One way Anova + two-way Anova + AnCova).

Target audience

Everyone who wants to use linear models with SPSS.

Presenter

An Carbonez

Prerequisites

- Basic knowledge of SPSS.
- Having followed the course on ‘Regression and Analysis of Variance’ or being familiar with the topics covered in this course.

Course Materials

The course material will be made available.
Bayesian Data Analysis with SAS

Course outline

There is a growing acknowledgement of the value of Bayesian methods for modelling complex data structures in basically every application area. This course will introduce the essentials of Bayesian ideas, emphasizing the practical application using simulation-based SAS software, such as the procedures BCHOICE, FMM, GENMOD, MI, MIXED LIFEREG, PHREG and MCMC. Examples will include the use of Bayesian methods in hierarchical models, measurement error models, longitudinal studies, missing data problems, survival models, etc.

The course is primarily based on the Wiley book of Lesaffre and Lawson (2012): Bayesian Biostatistics.

DAY 1
- Morning: Introduction to Bayesian concepts.
- Afternoon: Sampling with SAS.

DAY 2
- Morning: Gibbs sampling to obtain the posterior summary measures.
- Afternoon: Bayesian analysis of generalized linear models using PROC GENMOD. Parametric Bayesian survival models using PROC LIFEREG and PROC PHREG.

DAY 3
- Morning: Metropolis(-Hastings) sampling to obtain the posterior summary measures.
- Afternoon: Bayesian analysis of complex data structures using PROC MCMC.

Target audience

SAS users who wish to understand Bayesian analysis more thoroughly.

Prerequisites

Basic mathematical skills such as derivatives and integration are needed. Also an intermediate level of statistical knowledge is required, such as knowledge on regression models is mandatory. This course has an intermediate to advanced level.

Presenter(s)

Emmanuel Lesaffre and Kris Bogaerts

Course materials

The course material will be made available.
Regression and Analysis of Variance, applications with R

Course outline

The linear models, provided by the course ‘Regression and Analysis of Variance’, are applied on examples. In this course, the R package is used. By means of cases, we illustrate how to model your data in R and how to interpret the corresponding output. There is a hands-on session to train you with the functionality of R.

Target audience

Everybody who wants to model data with R.

Presenter

N.

Prerequisites

Everybody should be familiar with the techniques covered in ‘Regression and Analysis of Variance’ and have a basic knowledge of working with R.

Course Materials

The course material will be made available.

Dates

- 16 and 17 November 2017,
  09.00 – 12.00 and 13.00 – 16.00

Language

- English

Price

- (PhD) students KU Leuven and Association KU Leuven € 100
- Staff KU Leuven and Association KU Leuven, other (PhD) students € 160
- Non profit/social sector € 250
- Private sector € 600
Optimization & Numerical Methods in Statistics

Course outline

THIS IS A WEB LECTURE WITH Q & A SESSIONS

Numerical problems are frequently encountered by statisticians. Prominently, the estimation of the parameters of a statistical model requires the solution of an optimization problem. In a few simple cases, closed-form solutions exist but for many probability models the optimal parameter estimates have to be determined by means of an iterative algorithm. The goal of this course is threefold. First, we want to offer the readers an overview of some frequently used optimization algorithms in (applied) statistics. Second, we want to provide a framework for understanding the connections among several optimization algorithms as well as between optimization and aspects of statistical inference. Third, although very common, optimization is not the only numerical problem and therefore some important related topics such as numerical differentiation and integration will be covered.

Target audience

The intended target audience includes PhD students and researchers in a variety of fields, including biostatistics, psychometrics, educational measurement, public health, sociology. We aim at readers who apply and possibly develop statistical models and who wish to learn more about the basic concepts of numerical techniques, with an emphasis on optimization problems, and their use in statistics.

Prerequisites

Participants should have a basic knowledge of the principles of statistical inference. This includes some familiarity with the concept of a likelihood function and likelihood-based inference for linear, binomial, multinomial, and logistic regression models. Readers should also have a basic understanding of matrix algebra. A working knowledge of the basic elements of univariate calculus is also a prerequisite, including (the concepts of continuity of a function, derivative and integration).

Presenters

Geert Molenberghs, Francis Tuerlinckx, Katrijn Van Deun, Tom Wilderjans

Course materials

The course material will be made available.

Background reading

Introduction to Statistics with JMP

Course outline

Basic two-days hands-on course to give a broad introduction on JMP, with special focus on graph building and interactive features of JMP.

DAY 1: DATA MANAGEMENT AND DESCRIPTIVE STATISTICS WITH JMP
- Basic static and dynamic graphs,
- Maps,
- Descriptive statistics,
- Managing data (filtering, coding, merging, …),
- Hypothesis testing,
- Power calculations

DAY 2: BASIC REGRESSION AND ANOVA WITH JMP
- Simple and multiple linear regression,
- Model building (incl. stepwise regression),
- Visualization of regression modelling results (e.g., quadratic effects, interaction effects, …),
- One-way and multi-way analysis of variance (ANOVA)
- Logistic regression

Target audience

Everyone who wants to use JMP for data visualization, descriptive statistics and model fitting.

Prerequisites

Participants are familiar with basic statistical modeling, regression and analysis of variance concepts (see topics described in Fundamental Statistical Methods and Regression and Analysis of Variance). If you have basic knowledge in statistics and you want to know how to use JMP, then this is a course for you. If you miss this basic knowledge, then you can first follow ‘Fundamentele Statistische methoden’ or ‘Regression and Analysis of Variance’.

Presenter

José Nunez Ares

Course materials

The course material will be made available.

Textbook (not compulsory): Statistics with JMP: Graphs, Descriptive Statistics and Probability (Peter Goos, David Meintrup)
Niet-parametrische statistiek

Beschrijving
Deze cursus behandelt een aantal statistische technieken - analoog aan parametrische statistiek (bv. t-test, variantieanalyse) - waarbij de klassieke onderstellingen uit de parametrische statistiek niet hoeven gemaakt te worden (distributievrije technieken), technieken gebaseerd op ‘ordeningen’ of ‘rankings’, alsook technieken specifiek geschikt voor nominale gegevens.

INHOUD VAN DE CURSUS
• Chi- kwadraat goodness of fit testen
• Testen mbt verschil tussen twee onafhankelijke steekproeven
• Testen mbt verschil tussen twee afhankelijke steekproeven
• Testen mbt verschil tussen meerdere onafhankelijke steekproeven
• Testen mbt verschil tussen meerdere afhankelijke steekproeven
• Kengetallen mbt de samenhang tussen variabelen

Doelgroep
Gebruikers van basis statistische technieken (t-test – variantie-analyse)

Voorkennis
Cursisten dienen vertrouwd te zijn met de methodiek aangebracht in ‘Fundamentele Statistische technieken’ en variantie analyse.

Lesgever
Marlies Lacante

Cursusmateriaal
Cursusmateriaal wordt ter beschikking gesteld.

Datum
• 14 december 2017, van 9 u tot 12 u.

Taal
• Nederlands

Prijs
• (PhD) studenten van KU Leuven en Associatie KU Leuven € 25
• Personeel KU Leuven en Associatie KU Leuven, andere (PhD) studenten € 40
• Non profit/sociale sector € 62,50
• Private sector € 150
Applied Spatial Modelling with R

Course outline
This course is useful for data scientists and data analysts which work frequently with data with a spatial component (data with latitude/longitude information). It gives an introduction to the numerous spatial facilities of R and some standard spatial statistical models. The following items are covered during the course.

- The sp package to handle spatial data (spatial points, lines, polygons, spatial data frames)
- Importing spatial data and setting the spatial projection
- Plotting spatial data on static and interactive maps
- Adding graphical components to spatial maps
- Manipulation of geospatial data, geocoding, distances, ...
- Density estimation, kriging and spatial point pattern analysis
- Spatial regression

Target audience
The course is for R users in industry/academics who work frequently with spatial data and R and who want to either plot results on spatial graphs or used spatial modelling techniques on their data.

Prerequisites
Initial experience in R ranging from a few weeks to several years (at least understanding of the course ‘Essential Tools for R’ is needed, preferably also ‘Advanced R programming topics’). Some practical experience in regression modelling, handling of correlations and using R graphics.

Presenter
Jan Wijffels

Course materials
The course material will be made available.
Robust Statistics

Course outline

It is well-known that real data often contain outliers (also called anomalies), which are data points that do not follow the same pattern(s) as the majority of the data. Classical statistical methods may be strongly affected by outliers, often to the extent that the outliers are masked, in the sense that they go undetected. On the other hand, robust statistical methods are more resistant to outliers, which can often be detected by their large deviation (distance or residual) from the robust fit. The course gives a general introduction to robust methods, starting with univariate data (median, median absolute deviation, M-estimators, ...) and describes some basic tools such as the influence function and the breakdown value.

The next chapters deal with robust estimation of location and covariance matrices of multivariate data, robust regression methods, and principal component analysis. Throughout the course the methods are illustrated with R code.

Target audience

The course is targeted to people who are interested in methods for analyzing real data.

Prerequisites

Some familiarity with basic statistical methods is assumed.

Presenter

Peter Rousseeuw

Course materials

The course material will be made available.

Dates

- 14, 21 and 28 February, 7, 14, 21 and 28 March, 18 and 25 April, 2, 9, 16 and 23 May 2018, from 10.30 to 12.30
- Exercises on 1, 15 and 29 March and 26 April 2018 from 14.00 to 16.00

Language

- English

Price

- (PhD) students KU Leuven and Association KU Leuven € 275
- Staff KU Leuven and Association KU Leuven, other (PhD) students € 440
- Non profit/social sector € 690
- Private sector € 1650
Analytics for Big Data World

Course outline

We aim to bridge the gap between technical concepts and business applications of analytics techniques and big data. Participants will gain an insight in challenges and issues related to the storage and processing of large data sets, will be able to indicate which technologies, approaches and techniques are applicable for different types of data and objectives, will gain an insight in how advanced analytics can be used to optimize business decision in e.g. marketing, finance, logistics, HR, etc. We do so by taking a deep dive into "advanced analytics" on the one hand, i.e. advanced supervised and unsupervised modelling techniques as well as exploring "big data platforms and technologies" on the other hand, including Hadoop, Spark, stream processing engines, etc. The following topics will be discussed during the course:

- General introduction to data science, AI, predictive modeling, and everything in between
- The data science process
- Supervised modeling: classification and regression
- Unsupervised modeling: clustering, association rules and anomaly detection
- Model evaluation
- Advanced techniques: ensemble modeling, SVMs, artificial (deep) neural networks
- The data science tools ecosystem
- Hadoop and MapReduce
- Spark and SparkSQL
- Spark streaming and other stream processing engines
- NoSQL, Neo4j and Cypher
- Applications and use cases (special topics): text mining, social network mining, recommender systems, web mining
- Assignments will be provided throughout the course which can optionally be executed by the participants

Target audience

Example target profiles are: chief data officer, chief analytics officer, data architects, data project managers, (junior) industry data scientists, PhD and postdoctoral researchers

Prerequisites

We’ll go fast through the general introduction and basics of analytics, so a background knowledge of basic modeling techniques (clustering, regression, decision trees) is helpful. Some programming skills (R, Python and/or Java) are also helpful for participants willing to join in the assignments.

Presenter

Prof. Seppe vanden Broucke

Course Materials

The course material will be made available.
Experimental Design

Course outline

This course discusses the design of factorial experiments. Initially, the focus is on completely randomized experimental designs. Next, the focus shifts to experimental designs involving a restricted randomization. First, the concept of blocking is discussed. Next, split-plot and strip-plot designs are studied.

The emphasis in the course is on the optimal design of experiments. In optimal design of experiments, the experimental design is tailored to the problem at hand (unlike classical experimental design, where standard designs from catalogs are chosen).

The course builds on concepts from regression and analysis of variance, such as fixed and random effects, power calculations, variance inflation factors, multicollinearity, confidence intervals, prediction and lack-of-fit tests.

Every topic in the course is introduced and illustrated by means of a case study from industry. The case studies are realistic in the sense that they involve quantitative and qualitative experimental factors, experimenters have to deal with limited budgets and difficulties to randomize, and forbidden combinations of factor levels. In each of the case studies, the goal is to enhance the performance of a process or a product.

The statistical software package used is JMP.

Target audience

The target audience for the course is master students or PhD students in statistics, engineers and engineering students planning to perform experiments. The course is ideal for Six Sigma black belts, quality managers and people working in environments where collecting data is expensive.

Prerequisites

Prerequisites for the course are knowledge of basic statistics, regression analysis (least squares, multicollinearity), and matrix algebra (matrix products, inverse matrices, determinants).

Presenter

Peter Goos

Course Materials

The course material will be made available.

Textbook: Optimal Design of Experiments: A Case Study Approach (Peter Goos & Bradley Jones)
Chemometrics

Course outline
The goal of the course is to teach students how to perform multivariate sensor calibration. Students will become familiar with the use of statistical concepts in chemometric applications. Most attention will be given to the ideas underlying the different methods and the application of these methods to realistic examples. Theoretical considerations and equations will be limited to what is needed to have sufficient insight to properly use the methods. Most examples will be related to spectroscopy and analytical chemistry, but the scope is broader. By using a combination of lectures, computer sessions and take-home assignments, the students will really learn how to apply the chemometric methods.

The following aspects of chemometrics will be handled in this course:

- Classical modelling concepts for quantitative calibration: Classical Least Squares (CLS), Inverse Least Squares (ILS), Multivariate Linear Regression (MLR), Principle Component Regression (PCR) and Partial Least Squares (PLS).
- Necessary steps for the creation and successful deployment of calibrations: selection of calibration standards and assessment of the reliability of the models: (Test set validation vs. Cross-validation, model statistics). Special attention will be given to the methods for the selection of the number of principle components or latent variables in the projection methods.
- Methods for data pre-processing with special attention for the phenomena of light scattering and instrument drift and the methods to deal with these phenomena: derivatives, standard normal variate (SNV), multiplicative signal correction (MSC) and extended multiplicative signal correction (EMSC).
- Variable selection in a chemometric context and some commonly used methods for this.
- Qualitative analysis in a chemometric context: discrimination and classification
- New trends in chemometrics such as functional data analysis and augmented classical least squares (ACLS).

Target audience
The intended target audience includes PhD students and researchers in a variety of fields, including statistics, chemistry, biosciences and engineering. We aim at readers who wish to learn more about multivariate calibration of sensor systems and the use of statistical concepts in chemometric applications.

Prerequisites
Knowledge of basic concepts of statistics and linear algebra is required. Some notions of analytical chemistry, sensor technology and multivariate statistics are a plus.

Presenter
Wouter Saeys
Advanced R programming

Course outline

R is the lingua franca of statistical research and data analysis. But in order to get you up and running with R, and to get over the steep learning curve, you need to know how to use it efficiently.

This course is a hands-on course covering the basic toolkit you need to have in order to use R efficiently for data analysis tasks.

It is an intermediate course aimed at users who have the knowledge from the course 'Essential tools for R' and who want to go further to improve and speed up their data analysis tasks.

The following topics will be covered in detail:

- The apply family of functions and basic parallel programming for these, vectorisation, regular expressions, string manipulation functions and commonly used functions from the base package. Useful other packages for data manipulation.
- Making a basic reproducible report using Sweave and knitr including tables, graphs and literate programming
- If you want to build your own R package to distribute your work, you need to understand S3 and S4 methods, you need the basics of how generics work as well as R environments, what are namespaces and how are they useful. This will be covered to help you start up and build an R package.
- Basic tips on how to organise and develop R code and test it.

Target audience

People who have had their initial use of R and want to go one step further. This covers people using R for a few months already to several years. And more specifically users who want to extend their data manipulation techniques to speed up their day-to-day data analysis tasks.

Researchers from the university interested in making reproducible research reports or users who want to use R as a report generating tool.

R users interested in getting the fundamentals you need to know before you can create your own R package.

Business users who want to learn how to get the maximum out of R by speeding up their code, learn vectorisation, execute the basics of parallel programming and want to learn how to build methods and code which is reproducible in production environments.

Prerequisites

Initial experience in R ranging from a few weeks to several years.

Presenter

Jan Wijffels

Course materials

The course material will be made available.
Cluster analysis, principal component analysis and exploratory factor analysis with SAS Eguide, SPSS and R

Course outline

Multivariate data consist of observations on two or more variables for each individual or unit. The variables will be generally correlated, and a variety of techniques are available to analyse these data.

The objective of cluster analysis is to form groups of observations such that each group is as homogeneous as possible with respect to certain characteristics. The groups are as different as possible.

Principal component analysis is one of the popular tools to summarize quantitative multivariate data. During this course, PCA and exploratory factor analysis, will be introduced and the relation between them examined.

The emphasis of the course will be on the interpretation of the example data and on the results through the Biplot. Mathematical details are kept to a minimum.

For the exercises, participants can choose to use the statistics package SAS (through Enterprise Guide), SPSS or R.

COURSE CONTENT

- Hierarchical cluster analysis
- Nonhierarchical cluster analysis
- Linear combination of variables
- Eigenvectors and eigenvectors
- PCA scores and Factor scores
- What is a loading or the factor pattern?
- Scree plot
- How many components or factors to retain?
- Communals
- The biplot
- The Varimax rotation

Target audience

Data analysts and scientists involved in analysing multivariate data.

Prerequisites

A practical knowledge of basic statistics is assumed, such as standard deviations and correlations.

Presenters

Anne-Marie De Meyer and An Carboniez

Course materials

The course material will be made available
Computer Vision with R and Python

Course outline

This course is a hands-on course covering the use of image analysis. It covers basic image manipulation, feature engineering techniques and finding patterns in images.

The following items are covered:
- image manipulation & adjustments
- finding blobs, corners, gradients, edges & lines
- feature & object detection
- applying filters
- deep learning for image analysis
- image segmentation

Target audience

The course is for R/Python users in industry/academics who are interested in image analysis and have a computer science or data science background.

Prerequisites

You need some experience in R & Python for following the course. For the part on the modelling, knowledge of statistical learning techniques is preferred.

Presenter

Jan Wijffels

Course materials

The course material will be made available.

Dates
- 7 and 8 March 2018, 09.00 – 12.00 and 13.00 – 16.00

Language
- English

Price
- (PhD) students KU Leuven and Association KU Leuven € 100
- Staff KU Leuven and Association KU Leuven, other (PhD) students € 160
- Non profit/social sector € 250
- Private sector € 600
Fundamental Statistical Methods

Course outline
This basic course in statistics emphasizes on selecting the appropriate statistical method and drawing the right conclusions from the obtained results. Mathematical details will be kept to a minimum. The emphasis will be on understanding the concepts and on practical applications. The adequate use of basis statistical summaries (measures of central tendency, measures of dispersion, box-plots, ...) will be illustrated. The foundations of confidence intervals and of testing hypotheses will be dealt with.

COURSE CONTENT
• Descriptive statistics: graphical and numerical summaries of the data
• Distributions: Binomial, Poisson, Exponential, Normal and t-distribution
• Distribution of the sample mean
• Confidence intervals
• Hypothesis tests for a population mean (one and two samples)
• Paired t-test
• Estimating and testing for proportions

Target audience
Anyone who wishes to understand basic statistical techniques more thoroughly.

Prerequisites
There are no prerequisites

Presenter
Marlies Lacante

Course material
The course material will be made available.

Dates
• 13, 14 and 15 March 2018, 09.00 – 12.00

Language
• English

Price
• (PhD) students KU Leuven and Association KU Leuven € 75
• Staff KU Leuven and Association KU Leuven, other (PhD) students € 120
• Non profit/social sector € 187,50
• Private sector € 450
Concepts of Multilevel, Longitudinal and Mixed models

Course outline
Starting from ANOVA models with random factor levels, the concepts of mixed models are introduced and the basics about inference in random-effects models will be explained. Afterwards, the mixed ANOVA model is extended to general linear mixed models for continuous data. Finally, extensions to models for binary or count data will be briefly discussed. Omitting all theoretical details, sufficient background will be given such that practicing statisticians can apply mixed models in a variety of contexts, know how to use up-to-date software, and are able to correctly interpret generated outputs. Many applications, taken from various disciplines, will be discussed.

Target audience
The targeted audience includes methodological and applied statisticians and researchers in industry, public health organizations, contract research organizations and academia. Important: this course will also serve the MSc in Statistics students.

Prerequisites
The student knows the basics of statistical inference, and statistical modeling (regression, Anova and generalized linear models).

Presenter
Geert Verbeke

Course material
The course material will be made available.
Inleiding tot enquêtering

Beschrijving


Tevens zal uitsluitend voor KU Leuven deelnemers een half dagdeel besteed worden aan de enquêteservice aan de KU Leuven, die gebaseerd is op de open-source software “Limesurvey”. Deze software laat gebruikers toe om snel zeer krachtige online enquêtes te ontwikkelen.

INHOUD VAN DE CURSUS

• Analyse van de onderzoeksvraag: wat wil men te weten komen?
• Verzamelen van de gevraagde informatie
• Welke regels moet men in acht nemen bij het formuleren van de vragen? (invloed van de vraagstelling op het antwoord, betrouwbaarheid en validiteit)
• Methoden van steekproeftrekkingen
• Verwerken van de gegevens
• Rapportering
• Hoe werkt de enquêteservice van de KU Leuven
• Algemene instellingen voor de enquête
• Beschikbare vraagtypes en hun mogelijkheden
• Werken met tokens
• Uitnodigen van de respondenten en opvolgen van de responses
• Exporteren van de resultaten naar statistische pakketten

Doelgroep

Gebruikers van vragenlijstonderzoek

Voorkennis

Cursisten dienen vertrouwd te zijn met de methodiek aangebracht in 'Fundamentele Statistische technieken' en de cursus 'Regressie- en variantie analyse'.

Lesgevers

Marlies Lacante en Kristel Hoydonckx

Cursusmateriaal

Cursusmateriaal wordt als .pdf file ter beschikking gesteld.

Datum

• 20 en 21 maart 2018, van 09.00 tot 12.00
• Enkel voor KU Leuven deelnemers: extra sessie Limesurvey op 22 maart 2018, van 13u tot 16u

Taal

• Nederlands

Prijs

• (PhD) studenten van KU Leuven en Associatie KU Leuven € 100
• Personeel KU Leuven en Associatie KU Leuven, andere (PhD) studenten € 160
• Non profit/sociale sector € 187,50
• Private sector € 450
Graphics with R

Course outline

R has great features to make graphics and produce attractive figures. This workshop is intended for R users who want to take full advantage of the graphical possibilities of the R package ggplot2 and the R package plotly.

DAY 1: GG PLOT 2
- Grammar of ggplot2
- Build a plot layer by layer by ggplot()
- Working with scales, colour, shape and titles
- Extra facilities
- Adding statistical summaries
- Applications
- Exercises

The content of day 1 is based on the book “ggplot2, Elegant graphics for data analysis” by Hadley Wickham, Springer

DAY 2: PLOTLY
- Introduction
- Turn ggplot2 objects into plotly objects with ggplotly()
- Construct plotly objects directly with plot_ly()
- Basic chart types and graphing features
- Other chart types
- Interactive graphing features
- Beyond plotly

Target audience

Everybody who is interested to improve their graphical skills in R and produce high-quality graphs.

Prerequisites

Participants should be able to read and manipulate data in R. If you have no idea what the following commands mean, the course is too advanced for you.

mydata <- read.table("~/temp/rawdata.txt", header=T, dec="",)
sub.males<-mydata[mydata$sex=="male",]
sub.females<-mydata[mydata$sex=="female",]
mydata$pass<-as.factor(ifelse(mydata$examresult<10,0,1))
levels(mydata$pass)<-c("fail","pass")
table(mydata$pass)

No prior knowledge on statistics is assumed.

Presenters

An Carbonez and Dominik Sznajder

Course materials

The course material will be made available.
Logistic Regression Models, applications with SAS and SPSS

Course outline

The focus is on the statistical model where the response variable is categorical.

A categorical response variable can be a binary variable, an ordinal variable or a nominal variable and each type requires a different model to describe its relationship with the predictor variables.

We will define, interpret and illustrate the models for each type of outcome and place the models in the framework of the Generalized Linear Model.

SAS (we make use of SAS Enterprise Guide) or SPSS will be used in the applications in the afternoon. Applications with R will be on 20 April (see further, you’ll need to register separately).

• Introduction
• Binary Logistic Regression
• Multinomial Logistic Regression for nominal outcome variables
• Proportional Odds Model - Ordinal Logistic Regression
• Logistic regression in the framework of the Generalized Linear Model

Target audience

Data analysts in all disciplines

Prerequisites

• Fundamental Statistical Methods
• Knowledge of the linear regression model

Presenters

Anne-Marie De Meyer and N

Course materials

The course material will be made available.
Logistic Regression Models, applications with R

Course outline
The focus is on the statistical model where the response variable is categorical.

A categorical response variable can be a binary variable, an ordinal variable or a nominal variable and each type requires a different model to describe its relationship with the predictor variables.

We will define, interpret and illustrate the models for each type of outcome and place the models in the framework of the Generalized Linear Model.

The theoretical part on Thursday morning 19 April will be jointly with participants to the course “Logistic regression models, applications with SPSS and SAS”. R will be used in the applications on Friday morning 20 April.

• Introduction
• Binary Logistic Regression
• Multinomial Logistic Regression for nominal outcome variables
• Proportional Odds Model - Ordinal Logistic Regression
• Logistic regression in the framework of the Generalized Linear Model

Target audience
Data analysts in all disciplines

Prerequisites
• Fundamental Statistical Methods
• Knowledge of the linear regression model

Presenters
Anne-Marie De Meyer and N

Course materials
The course material will be made available.

Dates
• 19 April, from 09.00 to 12.00 (theory)
• 20 April 2018, from 09.00 to 12.00 (applications with R)

Language
• English

Price
• (PhD) students KU Leuven and Association KU Leuven € 50
• Staff KU Leuven and Association
KU Leuven, other (PhD) students € 80
• Non profit/social sector € 125
• Private sector € 300
Python for Finance

Course outline

Python is an open-source programming language which brings a lot of added value for anybody working in the financial industry with a quantitative background. The rising popularity of machine learning and artificial intelligence has given this relatively young language a serious boost. This practical course covers several aspects of the Python programming language with a focus on financial topics (risk, insurance, trading, …).

In a first step the delegate will be introduced into the main packages making up the backbone of Python in a financial context: NumPy, Pandas and Sci-Kit Learn. Topics such as broadcasting and universal functions, which make up the DNA of Python, will be explained. After this introductory step, Python will be put at work through several case studies that will be solved during the course.

The case studies and programs will be developed in Jupyter Notebooks.

CASE STUDIES:
- Monte Carlo simulation
- Building an option pricer
- Kernel density estimation
- Time series analysis
- Asset allocation

Target audience

- Portfolio managers
- Risk managers
- Professionals looking to introduce data-mining concepts in their day-to-day tasks
- IT developers
- Statisticians
- Quant analysts
- Financial engineers

Prerequisites

- Basic notions of mathematical Finance
- No prior knowledge of Python is required

Presenters

Lieven Op de Beeck and Jan De Spiegeleer

Course material

The course material will be made available.
Survey Sampling Methods

Course outline

THIS IS A WEB LECTURE WITH FACE-TO-FACE AND Q & A SESSIONS.

Different methods for selecting a (survey) sample from an existing population will be considered. Problems arising in the sampling designs will be discussed. The focus will be on the concepts rather than on the formulas. Nevertheless, attention will be paid at the estimation of the population parameters of interests.

Target audience

Everyone with an interest in sampling theory, from an applied and/or methodological point of view.

Prerequisites

Participants should have an intimate knowledge of basic concepts of descriptive and inductive statistics.

Presenter

Geert Molenberghs

Course material

The course material and the link to the web lectures will be made available.

Dates

- 24 April 2018, 16.30 – 20.00 (face-to-face lecture)
- 8 May 2018, 18.00 – 20.00 (Face- to-face Q & A session 1)
- 22 May 2018, 18.00 – 20.00 (Face-to-face Q & A session 2)

Language

- English

Price

- (PhD) students KU Leuven and Association KU Leuven € 125
- Staff KU Leuven and Association KU Leuven, other (PhD) students € 200
- Non profit/social sector € 312,50
- Private sector € 750
Statistical consulting Service

Consulting was our historical embryo and remains a core business.

The statistical consulting service center acts as the main pivot to determine the ideal combination between the customer and the most appropriate university entity.

We recommend that you contact us in an early stage of your project and write a short description of your problem and send it to info@lstat.kuleuven.be.

The LStat Statistical consulting Service covers:

- Statistical support for researchers within the university. The LStat provides statistical help for university research groups and for the central administration of the university. We help you with advice and we offer support with the design of your study and with the statistical analysis of your data whether elementary or sophisticated. The first hour of first-line consulting is provided free of charge.
- Statistical service and execution of projects for government and industry, in service or in partnership.

The LStat uses the administrative help of Leuven Research and Development (LRD) in the negotiation of the contracts with industries and the private sector.

We have experience with major financial companies, international institutions, manufacturers, medical organizations, marketing companies, FMCG as well as small and medium-sized enterprises.

Our solutions range from basic regression, multivariate techniques, analysis of variance, mixed models, data mining, process control, to risk theory, categorical data analysis, longitudinal data analysis and tailor-made simulations and calculations.

If you have any questions, do not hesitate to contact us at: an.carbonez@kuleuven.be and take a look at our website: http://lstat.kuleuven.be
Practical Matters

Registration costs

The indicated prices correspond to the training for one person.

There are several fee categories:

- PhD students and students KU Leuven and Association KU Leuven: €50/full day
- Staff from KU Leuven and Association KU Leuven, (PhD) students from other universities: €80/full day
- Non-profit sector, Social sector: €125/full day
- Private sector: €300/full day

Prices include all course material.

Location and timing of the courses

All courses will be held in Leuven centre and/or at Heverlee Arenberg Campus. For exact and updated information, always check our website [https://lstat.kuleuven.be/training/Timetable-2017-2018](https://lstat.kuleuven.be/training/Timetable-2017-2018).

Discount

When you register for several courses, you can get a discount of 10% if the total number of full training days equals or exceeds 5 days per person. A discount of 20% is attributed if you follow courses for at least 10 full days.

How to register

- If you are a KU Leuven staff member: via KU Loket (Personnel – Staff training – Research - Statistics)
- For all others: complete the online registration form available on our website: [https://lstat.kuleuven.be/forms/Registration 2017-2018](https://lstat.kuleuven.be/forms/Registration 2017-2018). You can register for several courses at the same time, or complete different forms throughout the year for one or more courses.

Confirmation and payment

You will receive a confirmation mail with payment details and additional information upon receipt of your registration. Please contact us in case you do not receive any confirmation mail.

Payment needs to be settled before the start of the course.

Cancellation

- If you are unable to attend a course for which you have registered, a colleague can take your place. Kindly inform us timely of any change.
- Cancellation without charges is possible up to 2 weeks before the start of the course. After that, the full course fee will be due and invoiced.
- We reserve ourselves the right to cancel a course if the minimum number of participants is not reached or in exceptional circumstances (i.e. illness presenter, extreme weather conditions etc). Registered candidates will be informed timely of any cancellation. In such case, registration fees will not be due and reimbursed if already paid.

Information

For questions on registration and extra information, please contact:

- tel. +32 16 32 88 75
- Eveline.declercq@kuleuven.be
- www.lstat.kuleuven.be
## Course timetable 2017-2018

<table>
<thead>
<tr>
<th>DATE</th>
<th>TITLE</th>
<th>PRESENTERS</th>
<th>LEVEL AND LANGUAGE</th>
<th>MORE ON PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SEPTEMBER 2017</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26, 27, 28 September 2 and 3 October</td>
<td>Essential Tools for R</td>
<td>Dominik Sznajder</td>
<td>Basic (English)</td>
<td>7</td>
</tr>
<tr>
<td><strong>OCTOBER 2017</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3, 4, 5 October</td>
<td>Fundamentele statistische methoden</td>
<td>Marlies Lacante</td>
<td>Basis (Nederlands)</td>
<td>8</td>
</tr>
<tr>
<td>9 October</td>
<td>Fundamental Statistical Methods, applications with R</td>
<td>N.</td>
<td>Basic (English)</td>
<td>9</td>
</tr>
<tr>
<td>9 and 17 October 8 November 1 December</td>
<td>Models for Longitudinal and Incomplete Data (Web lecture with introductory and Q&amp;A sessions)</td>
<td>Geert Molenberghs Geert Verbeke</td>
<td>Advanced (English)</td>
<td>10</td>
</tr>
<tr>
<td>18 and 19 October</td>
<td>Statistical Machine Learning with R</td>
<td>Jan Wijffels</td>
<td>Advanced (English)</td>
<td>12</td>
</tr>
<tr>
<td>19 October</td>
<td>Fundamentele statistische methoden, toegepast met SPSS</td>
<td>Marlies Lacante</td>
<td>Basis (Nederlands)</td>
<td>13</td>
</tr>
<tr>
<td>25, 26 and 27 October</td>
<td>Regression and analysis of variance</td>
<td>N.</td>
<td>Basic (English)</td>
<td>14</td>
</tr>
<tr>
<td><strong>NOVEMBER 2017</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 and 10 November</td>
<td>Text Mining with R</td>
<td>Jan Wijffels</td>
<td>Intermediate (English)</td>
<td>15</td>
</tr>
<tr>
<td>8, 9, 10, 13 and 14 November</td>
<td>Essential tools for R</td>
<td>Dominik Sznajder</td>
<td>Basic (English)</td>
<td>7</td>
</tr>
<tr>
<td>13 and 14 November</td>
<td>Regression and analysis of variance, applications with SPSS</td>
<td>An Carbenez</td>
<td>Basic (English)</td>
<td>16</td>
</tr>
<tr>
<td><strong>NEW</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13, 14 and 15 November</td>
<td>Bayesian Data Analysis with SAS</td>
<td>Emmanuel Lesaffre Kris Bogaerts</td>
<td>Advanced (English)</td>
<td>17</td>
</tr>
<tr>
<td>16 and 17 November</td>
<td>Regression and Analysis of Variance, applications with R</td>
<td>N.</td>
<td>Basic (English)</td>
<td>18</td>
</tr>
<tr>
<td>21 and 22 November</td>
<td>Optimization and Numerical Methods in Statistics (Web lecture with Q&amp;A sessions)</td>
<td>Geert Molenberghs Francis Tuerlinckx Katrijn Van Deun Tom Wilderjans</td>
<td>Advanced (English)</td>
<td>19</td>
</tr>
<tr>
<td><strong>NEW</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22 and 24 November</td>
<td>Introduction to Statistics with JMP</td>
<td>José Nunez Ares</td>
<td>Basic (English)</td>
<td>20</td>
</tr>
<tr>
<td><strong>DECEMBER 2017</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 December</td>
<td>Niet-parametrische statistiek</td>
<td>Marlies Lacante</td>
<td>Basis (Nederlands)</td>
<td>21</td>
</tr>
<tr>
<td>19 and 20 December</td>
<td>Applied spatial modelling with R</td>
<td>Jan Wijffels</td>
<td>Intermediate (English)</td>
<td>22</td>
</tr>
<tr>
<td>DATE</td>
<td>TITLE</td>
<td>PRESENTERS</td>
<td>LEVEL AND LANGUAGE</td>
<td>MORE ON PAGE</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------------------------------------------------------------------</td>
<td>-----------------------------------</td>
<td>--------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>FEBRUARY 2018</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEW 14, 21, 28 February</td>
<td>Robust Statistics</td>
<td>Peter Rousseeuw</td>
<td>Advanced (English)</td>
<td>23</td>
</tr>
<tr>
<td>1, 7, 14, 15, 21, 28, 29 March</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18, 25, 26 April 2, 9, 16, 23 May</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEW 15, 22 February</td>
<td>Analytics for Big Data World</td>
<td>Seppe vanden Broucke</td>
<td>Advanced (English)</td>
<td>24</td>
</tr>
<tr>
<td>1,8,15,22,29 March</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19, 26 April 2, 9, 16, 23 May</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2, 9, 16, 23 May</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEW 19, 20, 21, 26 and 27 February</td>
<td>Essential tools for R</td>
<td>Dominik Sznajder</td>
<td>Basic (English)</td>
<td>7</td>
</tr>
<tr>
<td>20, 27 February 6, 13, 20, 27 March</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17, 24 April 8, 15 May</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEW 23 February 2, 9, 16, 23 March</td>
<td>Chemometrics</td>
<td>Wouter Saeys</td>
<td>Advanced (English)</td>
<td>26</td>
</tr>
<tr>
<td>20 and 21 February</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEW 20 and 21 February</td>
<td>Advanced R programming</td>
<td>Jan Wijffels</td>
<td>Intermediate (English)</td>
<td>27</td>
</tr>
<tr>
<td>MARCH 2018</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6, 8 and 9 March</td>
<td>Cluster analysis, principal component analysis and exploratory factor analysis with SAS Eguide, SPSS and R</td>
<td>Anne-Marie De Meyer An Carbonez N.</td>
<td>Intermediate (English)</td>
<td>28</td>
</tr>
<tr>
<td>NEW 7 and 8 March</td>
<td>Computer Vision with R and Python</td>
<td>Jan Wijffels</td>
<td>Intermediate (English)</td>
<td>29</td>
</tr>
<tr>
<td>13, 14 and 15 March</td>
<td>Fundamentals of statistical methods</td>
<td>Marlies Lacante</td>
<td>Basic (English)</td>
<td>30</td>
</tr>
<tr>
<td>13, 20, 27 March 17, 24 April</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20, 21 and 22 March</td>
<td>Inleiding tot enquêtering</td>
<td>Marlies Lacante Kristel Hoydonckx</td>
<td>Basis (Nederlands)</td>
<td>32</td>
</tr>
<tr>
<td>22 and 23 March</td>
<td>Text Mining with R</td>
<td>Jan Wijffels</td>
<td>Intermediate (English)</td>
<td>15</td>
</tr>
<tr>
<td>NEW 29 and 30 March</td>
<td>Graphics with R</td>
<td>An Carbonez Dominik Sznajder</td>
<td>Intermediate (English)</td>
<td>33</td>
</tr>
<tr>
<td>APRIL 2018</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19 April</td>
<td>Logistic regression models with SAS and SPSS</td>
<td>Anne-Marie De Meyer N.</td>
<td>Intermediate (English)</td>
<td>34</td>
</tr>
<tr>
<td>NEW 19 and 20 April</td>
<td>Logistic regression models with R</td>
<td>Anne-Marie De Meyer N.</td>
<td>Intermediate (English)</td>
<td>35</td>
</tr>
<tr>
<td>23 April</td>
<td>Python for Finance</td>
<td>Jan De Spiegeleer</td>
<td>Intermediate (English)</td>
<td>36</td>
</tr>
<tr>
<td>24 April 8 and 22 May</td>
<td>Survey Sampling Methods</td>
<td>Geert Molenberghs</td>
<td>Intermediate (English)</td>
<td>37</td>
</tr>
</tbody>
</table>