Courses in Statistics and Mathematics

Gabriel Bénédict

December 16, 2018

Contents

1 BA in Economics (St Gallen) 2
- 1.1 Statistics [3,220] . . . 2
- 1.2 Empirical Economic Research [4,222] . . . 3
- 1.3 Mathematics for Economists [5,255] . . . 3
- 1.4 Stochastic Methods in Finance [4,252] . . . 4

2 Msc in Statistics (KU Leuven) 6
- 2.1 Multivariate Statistics [D0M62C] . . . 6
- 2.2 Fundamental Concepts of Statistics [G0A17A] . . . 7
- 2.3 Generalized Linear Models [G0A18A] . . . 8
- 2.4 Statistical Software [G0A21A] . . . 8
- 2.5 Robust Statistics [G0B16A] . . . 9
- 2.6 Statistical Analysis of Reliability and Survival Data [G0B67A] 9
- 2.7 Experimental Design [G0B68A] . . . 10
- 2.8 Sampling Theory [G0B72A] . . . 10
- 2.9 Concepts of Bayesian Data Analysis [G0B74A] . . . 10
- 2.10 Meta Analysis [G0B75A] . . . 11
- 2.11 Concepts of Multilevel, Longitudinal and Mixed Models [G0B76A] 12
- 2.12 Regression Analysis [G0S75A] . . . 12
- 2.13 Analysis of Variance [G0S76A] . . . 13
- 2.14 Avanced Non-Parametric Statistics [G0A23A] . . . 13
- 2.15 Support Vector Machines [H02D3A] . . . 14

3 Semester Exchange (Columbia University) 16
- 3.1 Applied Data Science [GR5243] . . . 16
- 3.2 Spatial Statistics [GR5070] . . . 17
- 3.3 Regression Modelling of Temporal Processes [G5016] . . . 17

1 BA in Economics (St Gallen)

Min Grade: 1, Max Grade: 6

1.1 Statistics [3,220]

Department Faculty of Mathematics and Statistics

Instructor Prof. Roger Baumann

grade 4.5

Course Content

- Probability theory: the building blocks
  - Random experiments
  - Probability models
  - Probability computation rules
  - Basic theorems
- Combinatorial methods
- Random variables: definition and properties
- Special distributions
- Multivariate random variables
  - Joint, marginal, and conditional distributions
  - Expectation, variance, and correlation
  - Sums and sample means of random variables
- The Central Limit Theorem (CLT)
- Descriptive statistics
- Estimation of unknown parameters
- Confidence intervals

Textbooks

- J. Schira, "Statistische Methoden der VWL und BWL: Theorie und Praxis", 3. Auflage, Pearson Studium (2009) - in German

1.2 Empirical Economic Research [4,222]

Department Faculty of Economics

Instructor Prof. Anthony Strittmatter

grade 4.5

Course Content

- Univariate linear regression model
- Inference and hypothesis testing
- Asymptotic Theory
- Multivariate linear regression model
- Inference and hypothesis testing
- Model specification: Non-linearity and misspecification
- Heteroscedasticity
- Autocorrelation
- Time series analysis

Textbooks

- Wooldridge, Jeffrey (2009): Introductory Econometrics - A Modern Approach; THOMSON South Western

1.3 Mathematics for Economists [5,255]

Department Faculty of Economics

Instructor Prof. Mahmoud Ola

grade 5

Course Content

- Part I: Probability Theory
  - Foundations of probabilities
  - Conditional probability and independence

  - Random variables, distributions, expectations
  - Convergence, law of large numbers, central limit theorem, characteristic functions
  - Introduction to advanced topics: martingales, Markov chains, Brownian motion

- Part II: Dynamic Systems and Stability
  - Linear difference equations with constant coefficients
  - General properties of differential equations
  - Linear and separable differential equations
  - Stability
  - Systems of differential equations in two variables; Lyapunov Stability Theorem
  - Systems of linear differential equations with constant coefficients

- Part III: Topics in Mathematical Analysis and Optimization
  - Taylor polynomials in one and two variables, Taylor's Theorem
  - Linear Algebra Review
  - Quadratic forms and definite matrices
  - Complex numbers
  - Method of Lagrange multipliers for optimization problems with several constraints
  - Implicit function theorem
  - Envelope Theorem
  - Convex analysis and Kuhn-Tucker Theorem

1.4 Stochastic Methods in Finance [4,252]

Department Faculty of Economics

Instructor Prof. Ortega Juan-Pablo

grade 5

Course Content

- Probability Theory

  - Introduction
  - Distribution functions
  - Normal distribution
  - Multivariate normal distribution
  - Lognormal distribution
  - Binomial distribution
- Pricing and No-arbitrage
  - Binomial model
  - Fundamental asset pricing theorem
- Ito's lemma and Stochastic Integrals
  - Random walk and Brownian motion
  - Ito processes and Ito lemma
  - Derivative pricing
  - Partial differential equations
  - Stochastic differential equations
- Risk Neutral Valuation
  - Discrete model
  - Lognormal model
  - Extensions
- Markowitz Portfolio Theory
  - Markowitz approach
  - Asset liability approach
  - Shortfall constraint
- Arbitrage Pricing Theory Model
  - Discussion of the model
  - Mathematics properties of the model
- Portfolio Theory in Continuous Time
  - Definition - Extensions

Textbooks

- Lamberton, D. and Lapeyre, B. "Introduction to stochastic calculus applied to finance". Chapman and Hall/CRC, 2008.
- Neftci, S. N. "An Introduction to the Mathematics of Financial Derivatives". Academic Press, 2000.
- Watsham, T. J. and Parramore, K. "Quantitative Methods in Finance". International. Thomson Business Press, 1997.

2 Msc in Statistics (KU Leuven)

Min Grade : 0, Max Grade 20

2.1 Multivariate Statistics [D0M62C]

Department Faculty of Statistics

Instructor Prof. Martina Vandebroek

grade 14

Course Content

- Principal component analysis
- Factor analysis
- Confirmatory factor analysis
- Structural equation models
- Cluster analysis
- Discriminant analysis
- Multivariate analysis of variance
- Logistic regression
- Canonical correlation
- Multidimensional scaling
- Correspondence analysis

2.2 Fundamental Concepts of Statistics [G0A17A]

Department Faculty of Statistics

Instructor Prof. Ortega Juan-Pablo

grade 10

Course Content

1. Probability.
2. Random variables and distributions
3. Multivariate distributions
4. Transformation of random variables
5. Convergence of random variables
6. Sample statistics and limit theorems
7. Estimation
8. Confidence intervals
9. Hypothesis testing
10. Bootstrap

Textbooks

- Mathematical Statistics and Data Analysis, J.A. Rice (Duxbury Press)
- Probability and Statistics, M.H. DeGroot and M.J. Schervish (Pearson Education)
- Mathematical Statistics with Applications, D. Wackerly, W. Mendenhall and R.L. Scheaffer (Thomson Brooks/Cole)
- A Course in Probability, N.A. Weiss (Pearson Education)

2.3 Generalized Linear Models [G0A18A]

Department Faculty of Statistics

Instructor Prof. Emmanuel Lesaffre

grade 13

Course Content In this course an overview of the generalized linear model is presented as the unifying framework for many commonly used statistical models. The emphasis is on the methods in categorical data analysis, model building and interpretation. We start with recapitulation on previous knowledge on the description and inference for contingency tables. The generalized linear model framework is presented. Examples are the Poisson model for counts and the logit model for binary outcome variables. Overdispersion is discussed in the Poisson model. We continue with the loglinear model and the link between the logit and the loglinear model. Special cases include the linear by linear association model for ordinal variables. The binary logit model is extended to multicategory logit models (nominal and ordinal).

2.4 Statistical Software [G0A21A]

Department Faculty of Statistics

Instructor Prof. An Carbonez

grade 20

Course Content

- With the SAS software:
  - Write a SAS program
  - Work with libraries and SAS help
  - Work with basic SAS procedures
  - Do some data handling by using the DATA step and Proc SQL
  - Use ODS
- With the R software:
  - Write an R script
  - Work with data structures

- Create an R function
- Create graphs
- Use R functions to produce basic statistical results
- Install R packages and work with R help

2.5 Robust Statistics [G0B16A]

Department Faculty of Statistics

Instructor Prof. Peter Rousseeuw

grade 12

Course Content
The course offers an introduction to the field of robust statistics, which comprises the study of statistical methods that are more resistant to outlying observations than classical methods. It introduces the most basic robust methods such as M-estimators, trimmed estimators and depth-based methods in several statistical models. Their main properties (such as breakdown value and influence function) are discussed, as well as their computation. Students are also introduced to recent scientific papers and research results. In computer sessions, robust methods will be applied to real data sets and the results will be interpreted. Some properties of the estimators will be verified empirically, for instance by Monte Carlo simulation.

2.6 Statistical Analysis of Reliability and Survival Data [G0B67A]

Department Faculty of Statistics

Instructor Prof. Roel Braekers

grade 15

Course Content
The concept of censoring and different types of censoring will be discussed. Hereby it will be argued that censoring cannot be ignored in the statistical analysis of reliability and survival data. Starting from lifetable analyses, the Kaplan-Meier estimator will be introduced, and methods will be discussed for the comparison of multiple survival distributions. Next different parametric models as the accelerated failure time model and the proportional hazard model will be introduced and comparisons between these models are made. Finally, the semi-parametric Cox proportional hazards model will be introduced, discussed and illustrated.

2.7 Experimental Design [G0B68A]

Department Faculty of Statistics

Instructor Prof. Eric Schoen

grade 12

Course Content
This course deals with modern methods for setting up highly informative experiments. More specifically, it provides an in-depth treatment of the optimal experimental design approach, which is extremely flexible and can handle all kinds of practical constraints that may occur in the planning phase of an experiment. During the course, the students learn how to use the JMP software, which is state of the art for the optimal design of experiments. The focus is on factorial experiments, i.e., experiments studying multiple treatment factors.
Textbooks

- "Optimal design of experiments: A case study approach", co-authored by Peter Goos and Bradley Jones.

2.8 Sampling Theory [G0B72A]

Department Faculty of Statistics

Instructor Prof. Geert Molenberghs

grade 16

Course Content
Different methods for selecting a (survey) sample from an existing population will be considered. Problems arising in the sampling designs will be discussed.

2.9 Concepts of Bayesian Data Analysis [G0B74A]

Department Faculty of Statistics

Instructor Prof. Christel Faes

grade 12

Course Content
This course will give a broad introduction to basic concepts of Bayesian analysis. Posterior summary measures, predictive distributions and Bayesian hypothesis tests will be contrasted with the frequentist approach. Simulation methods such as Markov chain Monte Carlo (MCMC) enable the Bayesian analysis . An introduction to algorithms like Gibbs sampling and Metropolis-Hastings will be explained and illustrated. Various medical case studies will be considered.

2.10 Meta Analysis [G0B75A]

Department Faculty of Statistics

Instructor Prof. Wim Van Den Noortgate

grade 11

Course Content

- Conducting a statistical meta-analysis:
  - Effect size measures: definition, use and comparability
  - Combining effect sizes compared with narrative reviews, vote-counting methods and combining p-values
  - Assessing between-study heterogeneity
  - Fixed, random and mixed effects models for integrating study outcomes
  - Assessing and accounting for publication bias and study quality
  - Sensitivity analyses
- Specific applications and extensions:
  - Assessing and accounting for study quality
  - Meta-analysis of individual patient data
  - Meta-analysis of epidemiological and other observational studies
  - Bayesian methods in meta-analysis
  - Meta-analysis of multiple and correlated outcome measures
  - Cumulative meta-analysis

2.11 Concepts of Multilevel, Longitudinal and Mixed Models [G0B76A]

Department Faculty of Statistics

Instructor Prof. Geert Verbeke

grade 16

Course Content
Starting from ANOVA models with random factor levels, the concepts of mixed models is introduced and the basics about inference in random-effects models will be explained. Afterwards, the mixed ANOVA models is extended to general linear mixed models for continuous data. Finally, extensions to models for binary or count data will be briefly discussed.

2.12 Regression Analysis [G0S75A]

Department Faculty of Statistics

Instructor Prof. Mia Hubert

grade 15

Course Content

- Simple linear regression
- The general linear model
- Multiple linear regression: inference
- Polynomial regression
- Categorical predictors
- Transformations (including Box-Cox, weighted regression)
- Variable selection methods
- Multicollinearity and its remedials (PCR, ridge regression)
- Robust regression
- Nonlinear regression
- Nonparametric regression

2.13 Analysis of Variance [G0S76A]

Department Faculty of Statistics

Instructor Prof. Abad Ariel Alonso

grade 18

Course Content

- Statistical design of scientific studies
- One-way ANOVA: F-test
- Factor levels effects and multiple comparison
- Diagnostics and remedial measures
- Model reformulation and regression approach
- Two-way ANOVA: Equal sample size
- Two-way ANOVA: Unequal sample size
- Estimable Functions
- Type I, II and III Sum of Squares

Textbooks

- Applied Linear Statistical Models, 5th Edition, Kutner et al. (2005)

2.14 Avanced Non-Parametric Statistics [G0A23A]

Department Faculty of Statistics

Instructor Prof. Clément Cerovecki

grade TBA

Course Content

- overview of nonparametric methods for estimating a density: kernel estimation methods, nearest-neighbour methods, maximum-likelihood-based methods, orthogonal series method, wavelets, ...

- kernel estimators of densities: basic properties (bias, variance, mean squared error), asymptotic properties, asymptotic normality, rates of convergence (and their meaning/interpretation), selection of smoothing parameters (via cross-validation, plug-in, bootstrap or resampling procedures, ...).
- nonparametric estimation of a regression function: the cases of fixed and random design, homoscedasticity and heteroscedasticity, Nadaraya-Watson estimator, Gasser-Müller estimator, weighted least-squares methods, local polynomial fitting, splines, P-splines, wavelets, ... The impact and choices of parameters in each of these techniques will be discussed.
- nonparametric estimation of hazard functions and applications (e.g. in survival analysis).
- multivariate regression models: additive modelling and backfitting algorithms, dimension reduction techniques.
- nonparametric smoothing and deconvolution problems (e.g. measurement errors).
- nonparametric estimation of boundaries and frontiers with applications in image analysis and econometrics (for example).
- modelling dependencies and nonparametric techniques, for example, use of nonparametric techniques in time series context.
- other applications of nonparametric techniques: classification techniques, neural networks, statistical learning and data mining, modelling dependencies.

2.15 Support Vector Machines [H02D3A]

Department Faculty of Statistics

Instructor Prof. Johan Suykens

grade 5

Course Content

- Basics of statistical decision theory and pattern recognition

- Basics of convex optimization theory, Karush-Kuhn-Tucker conditions, primal and dual problems
- Maximal margin classifier, linear SVM classifiers, separable and non-separable case
- Kernel trick and Mercer theorem, nonlinear SVM classifiers, choice of the kernel function, special kernels suitable for textmining
- Applications: classification of microarray data in bioinformatics, classification problems in biomedicine
- VC theory and structural risk minimization, generalization error versus empirical risk, estimating the VC dimension of SVM classifiers, optimal tuning of SVMs
- SVMs for nonlinear function estimation
- Least squares support vector machines, issues of sparseness and robustness, Bayesian framework, probabilistic interpretations, automatic relevance determination and input selection, links with Gaussian processes and regularization networks
- Applications: time-series prediction, finance
- Kernel versions of classical pattern recognition algorithms, kernel Fisher discriminant analysis
- Kernel trick in unsupervised learning: kernel based clustering, SVM and kernel based density estimation, kernel principal component analysis, kernel canonical correlation analysis
- Applications: datamining, bioinformatics
- Methods for large scale data sets, approximation to the feature map, estimation in the primal
- From statics to dynamics: SVM extensions to recurrent models and control

Textbooks

- J.A.K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, J. Vandewalle, Least Squares Support Vector Machines, World Scientific, Singapore, 2002 (ISBN 981-238-151-1)

- Cristianini N., Shawe-Taylor J., An introduction to support vector machines, Cambridge University Press, 2000.
- Schoelkopf B., Burges C., Smola A., Advances in Kernel Methods: Support Vector Learning, MIT Press, Cambridge, 1998.
- Schoelkopf B., Smola A., Learning with Kernels, MIT Press, Cambridge, 2002
- Vapnik V., Statistical learning theory, John Wiley, New-York, 1998.

3 Semester Exchange (Columbia University)

A-F Grading

3.1 Applied Data Science [GR5243]

Department Faculty of Statistics

Instructor Prof. Ying Liu

grade TBA

Course Content
This course will be a project-based hands-on course in data science. No formal instruction on statistics, data science, machine learning will be given. Project cycles run every 2-3 weeks, where we will have mini- group data projects. Groups will be formed randomly and project products will be peer-reviewed, in addition to evaluation by the instructional team.
Textbooks

- Mount and Zumel (2014) Practical data science with R.
- Segaran (2007) Programming collective intelligence: building smart web 2.0 applications.
- Tuffe (2001) The visual display of quantitative information.
- Fung (2013) Numbersense: how to use big data to your advantage.
- Wickham (2017) R for Data Science `http://r4ds.had.co.nz/`

3.2 Spatial Statistics [GR5070]

Department Faculty of Quantitative Methods in the Social Sciences

Instructor Prof. Michael Parrott

grade TBA

Course Content
The goal of this course is to provide an overview of, and an in- troduction to statistical techniques used in, the analysis of spa- tial/geographic data in the social sciences. It covers introductory concepts and tools related to Geographic Information Systems (GIS); including spatial data acquisition, spatial data management, and spatial data analysis.
Textbooks

- Steven J. Steinberg and Sheila L. Steinberg. 2006. GIS: Geographic Information Systems for the Social Sciences. Thousand Oaks, CA: Sage Publications.
- Fotheringham, A. Stewart, Chris Brunsdon and Martin Charlton. 2000. Quantitative Geography: Perspectives on Spatial Data Analysis. London, UK: Sage Publications
- Mitchell, Andy. 1999. The ESRI Guide to GIS Analysis, Volume 1: Geographic Patterns and Relationships. Redlands, CA: ESRI Press.
- Mitchell, Andy. 2005. The ESRI Guide to GIS Analysis, Volume 2: Spatial Measurements and Statistics. Redlands, CA: ESRI Press.
- Lloyd, Christopher D. 2010. Spatial Data Analysis: An Introduction for GIS Users. Oxford, UK: Oxford University Press.
- Howell, Frank M., Jeremy R. Porter and Stephen Matthews. 2016. Recapturing Space: New Middle Range Theory in Spatial Demography. Springer Publications

3.3 Regression Modelling of Temporal Processes [G5016]

Department Faculty of Quantitative Methods in the Social Sciences

Instructor Prof. Gregory M. Eirich

grade TBA

Course Content
This course is meant to introduce students to the main concepts and methods behind regression analysis of temporal processes. Incorporating time into our analyses complicates and deepens our understanding of social and economic dynamics, but at the same time, properly dealing with time requires attention to a number of new issues not found in simple OLS regression of cross-sectional data.

- ARIMA
- GARCH
- Machine Learning Time Series Methods

Textbooks

- "Introductory Econometrics: A Modern approach," 4th Edition (2009), by Jeffrey M. Wooldridge. The ISBN is 9780324581621.
- "Introduction to Time Series Analysis," by Pickup, Mark (Newbury Park, Calif.: Sage Publications, 2015). The ISBN is 9781452282015.