



Institut de statistique

Annual Report

Year 2006

UCL
Université catholique de Louvain

Contents

1. Foreword	2
2. Staff	4
3. Research Activities	10
3.1. Presentation	
3.2. Research contracts	11
3.2.1. Research projects under contracts and cooperation projects	
3.2.2. Applied research contracts	13
3.3. Reprints	15
3.4. Discussion papers	17
3.5. Books published by the Members of the Institute	18
3.6. Visitors	19
3.7. Statistics Seminars	20
3.8. Organization of scientific meetings.....	22
3.9. Academic visits	27
3.10. Participation to conferences and statistics meetings	29
4. Teaching Activities and Educational Activities	32
4.1. Bachelor level	
4.2. Minor in statistics	
4.3. Master and DEA	
4.4. Towards the new Master programs in 2007.....	33
4.5. Certificate	
4.6. Doctoral School	34
4.6.1. Presentation	
4.6.2. Doctoral dissertations and Ph. D. in progress	35
4.6.3. Doctoral seminars	36
4.6.4. Short courses	37
4.7. Internationalization	38
4.7.1. Bilateral agreements	
4.7.2. Cooperation programs	
4.7.3. Mandate of invited professors	
5. Support for the University and Consulting Activities	39
5.1. Presentation	
5.2. Consulting activities	
5.3. Applied Statistics Workshops	40
6. Appendix	
6.1. Abstracts of the discussion papers	41
6.2. Abstracts of the consulting reports	49

1. Foreword

This year 2006 was another very productive and successful one for the Institute of Statistics of the Université catholique de Louvain (UCL, Louvain-la-Neuve).

Since its founding by Professor Léopold Simar in the early 1990's, the UCL Institute of Statistics has undoubtedly become the largest and most comprehensive statistics faculties within a single department in Belgium, and one of the largest in the EU. The Institute has as its purpose (i) to provide excellent education in probability and statistics, (ii) to extend the frontiers of the knowledge in probability and statistics through pure and applied research, (iii) to collaborate with researchers in other disciplines to investigate important scientific issues, and (iv) to serve the UCL community in all areas related to probability and statistics.

In 2006 the teaching faculty comprised a full time staff of 10 professors and 16 teaching and research assistants. Several invited professors broaden the scope of expertise of the Institute in specific areas. The Institute also accommodates post-doc researchers and short-term as well as long-term foreign visiting scholars. The administrative, library, and technical staff consists of 9 persons.

The Institute of Statistics is responsible for all the probability and statistics courses offered in the university. Students in many programmes, in all the faculties, take the ancillary statistics courses taught by members of the Institute who always place strong emphasis on quality teaching. Besides these courses taught in the different faculties of the university, the Institute of Statistics continued strengthening its master and doctoral programmes. Increasing number of students both pass our courses and obtain degrees, see Section 4 of this report. That our teaching has become more innovative and varied is undoubtedly responsible for this excellent result.

Members of the Institute have a wide variety of research interests, ranging from theoretical to applied topics and over a broad spectrum of methodological topics. The main research areas at the Institute are mathematical statistics (semi- and non-parametric statistics, Bayesian statistics, multivariate analysis, regression estimation, mixtures models, survival analysis, time series, ...) with important implications in various fields of application (actuarial science, biostatistics, econometrics, industrial statistics, transportation problems,...).

Members of the Institute actively collaborate with scientists from other disciplines on important research investigations. In particular, the application of probability and statistics to solving problems in insurance, pension and risk management is conducted in the Institute of actuarial science that is run jointly by the UCL business school and the Institute of Statistics, combining so an unique expertise as well in deep mathematical background as in business aspects. The activities in this field are described in a separate report.

Research activity continued to flourish in 2006. During 2006, members of the Institute published 25 papers in refereed journals. Research to appear soon, or surveyed, was reported in 25 discussion papers.

The Institute of Statistics is the leader of an IAP research network in statistics (Interuniversity Attraction Pole, Phase VI 2007-2011) on the theme "Statistical Analysis of Association and Dependence in Complex Data". The network involves 9 research teams, of which five are Belgian and 4 are European, non-Belgian partners. The current network is a continuation of the network "Statistical techniques and modeling for complex substantive questions with complex data", financed from 2002 till 2006 by BELSPO. In addition to the IAP, several research projects are funded by public and private bodies.

Members of the Institute serve on the editorial boards of several prestigious journals. These editorial roles for international journals held by Statistics staff is a clear mark of international regard.

There are three regular seminars held at the Institute. At the statistics seminars, invited speakers present their research results. A diversity of subjects is presented at this seminar by foreign visitors of the Institute. From time to time, a joint statistics and econometrics seminar, organised in collaboration with CORE, takes place. The "Atelier de statistique appliquée/Applied Statistics Workshop" is organised by the Institute on a regular basis. It focuses on problem driven statistics, where a real world problem of substantial practical interest is treated. At the "Séminaires des jeunes chercheurs/Doctoral Seminar", graduate and undergraduate students in statistics present their recent research work. These doctoral seminars constitute an extra stimulant for Ph.D. students and other young researchers.

The Graduate School in Statistics and Actuarial Science has been created in 2006, under the patronage of the "Fonds National de la Recherche Scientifique (FNRS)". It is also associated to the Graduate Colleges in Medicine, Agronomie, Economical Sciences, and Engineering, as Statistics and Actuarial Science have links to very many other disciplines.

The Graduate School gathers students from various French-speaking Belgian universities, and is chaired by the Institute of Statistics.

The Institute of Statistics offers a consulting service to the UCL community. Researchers of the others faculties receive advice concerning appropriate methodologies and suitable statistical packages for their specific problems. The scientific members and the computer scientists of the Institute are developing their knowledge about the evolution of statistical software and they give advice in this context. The Institute hosts the "Statistical application server" used daily by UCL researchers and students for their statistical computations. It has been decided to merge these two services in 2007 to further increase the quality of the assistance to the UCL community.

In 2006, Dr Johan Segers was appointed in the Institute of Statistics after an international search. The former Head of the Institute, Professor Jean-Marie Rolin, retired at year's end. It was the Institute's great loss that Professor Philippe Lambert left to join ULg in December 2006. We would like to take this opportunity to publicly thank Philippe for his efforts to increasing the reputation and activity of the Institute in biostatistics during the 7 years he was here.

We would like to thank the Université catholique de Louvain for supporting the activities of the Institute, as well as all the members of the Institute for developing a friendly, open and stimulating research and teaching environment. In the future, we hope to further increase the strengths of the Institute in education, research, and service to the community.

Louvain-la-Neuve, April 18, 2007

Michel Denuit
Chairman

Sébastien Van Bellegem
Academic Secretary

2. Staff

Permanent Academic Members



Michel Denuit

Professor, Université catholique de Louvain (Institut de statistique).

Ph. D. Université Libre de Bruxelles, 1997.

Stochastic inequalities, mathematical risk theory, statistics applied to insurance.

Associate editor of the *Australian and New Zealand Journal of Statistics* (since 2003),

Proceedings Editor for *Insurance : Mathematics and Economics* (since 1999).

Co-author of 7 books, including "Modern Actuarial Risk Theory" (Kluwer Academic Publishers, 2001), "Actuarial Theory for Dependent Risk" (John Wiley & Sons, New York, 2005), "Construction de Tables de Mortalité Périodiques et Prospectives"

(Collection Audit-Actuariat-Assurance, Economica, Paris, 2005).

The work of Michel Denuit has been awarded several times. In 2006, he received the Giuseppe Ottaviani Prize in Insurance by the Italian Institute of Actuaries.

Michel Denuit is chairman of the Institute of Statistics (2006-2009) and Academic Secretary of the Institute of Actuarial Science.

Homepage : <http://www.actu.ucl.ac.be/staff/denuit/mdenuit.html>



Dominique Deprins

Part-time professor, Université catholique de Louvain (Institut de statistique), full-time professor, Facultés Universitaires Saint-Louis, Bruxelles.

Ph. D. Université catholique de Louvain, 1989.

Epistemological work about Statistics and Probabilities : application in the field of mental health and, more generally, in the field of human sciences.

Academic secretary of Facultés Universitaires Saint-Louis (1999-2005), pedagogic adviser (since 2004), President of the Committee responsible for internal organization of education in mathematics, statistics, sciences and computer sciences (since 2000) at Facultés Universitaires Saint-Louis.



Bernadette Govaerts

Professor, Université catholique de Louvain (Institut de statistique).

Ph. D. Université catholique de Louvain, 1987.

Industrial statistics, chemometrics, experimental design, statistical quality control, pre-clinical biostatistics, development, validation and monitoring of laboratory analytical methods.

Academic secretary of the Institute of Statistics (2000-2002).

Bernadette Govaerts is responsible of the master and certificate of statistics and coordinator of the statistical consulting and computing services.

Homepage : <http://www.stat.ucl.ac.be/ISpersonnel/govaerts/govaerts.html>



Christian Hafner

Professor, Université catholique de Louvain (Institut de statistique).

Ph. D. Humboldt-Universität Berlin, Germany, 1996.

Nonparametric and semiparametric statistics, time series, volatility models, financial econometrics.

Associate editor of *Computational Statistics* (since 2004), *Eurasian Review of Econometrics*

(since 2006), invited co-editor of a special issues of the journal *Studies in Nonlinear Dynamics Econometrics* (2006).

“Nonlinear Time Series Analysis with applications to Foreign Exchange Rate Volatility”, (Physica Verlag, Heidelberg, 1998).

“Einfuehrung in die Statistik der Finanzmaerkte” (Springer Verlag, 2001 – second edition 2003).

Member of the Econometrics and Empirical Economics Programme Committee of the Econometric Society European Meeting (Vienna, Austria 2006).

Core fellowship (1996-1997). Head of the Econometrics group at Electrabel (1999-2002).

Christian Hafner is responsible for the web site of the Institute of Statistics, coordinator for the Statistics Minor program and organiser of the Statistics Seminars.

Homepage : <http://www.stat.ucl.ac.be/ISpersonnel/hafner/>



Philippe Lambert

Professor, Université catholique de Louvain (Institut de statistique).

Ph. D. Universiteit Hasselt, 1995.

Quantitative methods in social sciences, biostatistics, Bayesian statistics, repeated measurements, dependence modelling and copulas, discrete data, statistical tools for the analysis of brain signals, MCMC methods in practice.

President of the biostatistic section of the Belgian Statistical Society (2004-2006).

Member of the board of the Belgian Statistical Society.

Homepage : <http://www.stat.ucl.ac.be/ISpersonnel/lambert/lambert.html>



Christian Ritter

Part-time Professor, Université catholique de Louvain (Institut de statistique).

Ph. D. University of Wisconsin, Madison, USA, 1992.

Industrial statistics, statistical consulting, statistics in spreadsheets.

Homepage : <http://www.stat.ucl.ac.be/ISpersonnel/ritter/Newritter.html>



Johan Segers

Professor, Université catholique de Louvain (Institut de statistique).
Ph. D. Leuven, 2001.

Extreme value theory, Nonparametric curve estimation, Dependence modelling via copulas, Edgeworth expansions.

Associate editor of the *Journal of the Royal Statistical Society, Series B* (from 2005).

Co-author of "Statistics of Extremes : Theory and Applications" (John Wiley & Sons).

Extramural fellow of CenTER, Tilburg University, The Netherlands (2006-2011).

Secretary of the Section Mathematical Statistics of the Netherlands Society for

Statistics and Operations Research (2005-2008).

Homepage : <http://www.stat.ucl.ac.be/ISpersonnel/segers/index.html>



Léopold Simar

Professor, Université catholique de Louvain (Institut de statistique), Part-time professor, Facultés Universitaires Saint-Louis, Bruxelles.

Ph. D. Université catholique de Louvain, 1974 (Docteur en Sciences Appliquées).

Frontier estimation, resampling methods, multivariate statistical techniques.

Associate editor, *Journal of Productivity Analysis* (since 2003), Member of the

European Network of Excellence Prime (2004-2009), Associate partner of the

Scuola Superiore Santa Anna, Pisa, Italy - project AQUAMETH, Dean, Faculté des

Sciences Economiques, Sociales et Politiques, FUSL (1978-1990), President of the

Belgian Statistical Society (1999-2002), Founder President of the Institut de Statistique, UCL (1992-2004).

Co-author of "Applied Multivariate Statistical" (Springer-Verlag, Berlin, 2003)

Homepage : <http://www.stat.ucl.ac.be/ISpersonnel/simar/index.html>



Sébastien Van Belleghem

Professor, Université catholique de Louvain (Institut de statistique).

Ph. D. Université catholique de Louvain, 2003.

Econometric theory and microeconometrics, semi and nonparametric statistics, local stationary time series.

Associate editor of *Eurasian Review of Econometrics* (since 2006).

The work of Sébastien Van Belleghem has been awarded several times. In 2005, he received the Prix Marie-Jeanne Laurent-Duhamel from the French Statistical Society for the best Ph. D. Thesis in statistics.

Sébastien Van Belleghem is Academic secretary of the Institute of Statistics (2006-2009).

Homepage : <http://www.stat.ucl.ac.be/~vanbelle>



Ingrid Van Keilegom

Professor, Université catholique de Louvain (Institut de statistique).

Ph. D. Limburge Universitair Centrum, 1998.

Non and semiparametric regression, Survival analysis, Bootstrap methods, Mathematical statistics, Empirical likelihood methods, Inverse problems, Goodness-of-fit problems.

Associate editor of the *Journal of the Royal Statistical Society – Series B* (2002-2006), *Scandinavian Journal of Statistics* (2004-2007), *Annals of the Institute of Statistical Mathematics* (2006-2009), *The Annals of Statistics* (2007-2009).

President of the Three Country Corner (Local group of the Royal Statistical Society),

Committee member of the « European Regional Committee » of Bernoulli Society (since 2004),

Member of the steering group of the « Séminaire Européen de Statistique » (since 2004).

Ingrid Van Keilegom follows the developments on various research funding programmes for the Institute of Statistique, she is Secretary of the Jury for the Master in statistics, and academic responsible for the library CORE/STAT.

Homepage : <http://www.stat.ucl.ac.be/ISpersonnel/vankeile/vankeile.html>



Rainer von Sachs

Professor, Université catholique de Louvain (Institut de statistique)

Ph. D. Heidelberg, Germany, 1991.

Mathematical statistics, nonparametric curve estimation, Analysis of (nonstationary) time series, spectral density estimation, Statistical signal processing, biomedical time series, financial time series, Wavelets and related localization methods.

Associate editor of the *Journal of the Royal Statistical Society, Series B* (2003-2006).

Rainer von Sachs is President of the FNRS Graduate School in Statistics and Actuarial Sciences of the Communauté française (as of April 2006).

Homepage : <http://www.stat.ucl.ac.be/rvs>

Associate Academic Members

Luc Bauwens, Faculty of Economical, Social and Political Sciences
Patrick Bogaert, Faculty of Biological, Agronomic and Environmental Engineering
Pierre Devolder, Institute of Actuarial Sciences
Marie-Paule Kestemont, Faculty of Economical, Social and Political Sciences
Eric Le Boulengé, Faculty of Sciences
Guy Lories, Faculty of Psychology
Annie Robert, Faculty of Medicine
Michel Verleysen, Faculty of Applied Mathematics

Invited Academic Members

Yves Berger
Libei Chen
Isabelle De Macq
Cédric Heuchenne

Emeritus Professors

Michel Mouchart
José Paris
Jean-Marie Rolin (since October 1, 2006)

Researchers and Doctoral Students

Carlos Almeida
Hilmar Böhm
Jean-Philippe Boucher
Cindy Courtois
Alexandra Daskovska
Anouar El Ghouch
Gery Geenens
Julien Hunt
Astrid Jullion
Thomas Laloux
Céline Le Bailly de Tillegem
Giovanni Motta
Sandra Pitrebois
Olga Reznikova (since October 1, 2006)
Réjane Rousseau
Bianca Teodorescu

Postdoctoral Researchers

Cédric Heuchenne (until September 15, 2006)
Davit Varron (until August 31, 2006)

Associate Researchers

Taoufik Bouezmarni
Céline Bugli
Nancy François
Maria Key Prato
Philippe Vanden Eeckaut

Researchers as visitors

Alessandro Fiorini, Università di Roma, Roma, Italy (01/10/06-31/01/07)
Elisa Maria Molanes Lopez, Universidade da Coruña, A Coruña, Spain (02/10/06-05/11/06)
Jorge Luis Ojeda Carbrera, Universidad de Zaragoza, Zaragoza, Spain (20/09/06-20/12/06)

Doctors honoris causa

Luc Devroye, Mc Gill University, Montréal, Canada (2002)

Peter Hall, Australian National University, Canberra, Australia (1997)

Computer Scientists

Laurent Buset

Jean-Luc Marrison (Responsible)

Jean-Marie Zélis (Responsible of the University statistical server)

Statistician - Consultant

Alain Guillet

Administrative Responsible

Christine Denayer

Administrative Staff

Anne-Marie Bellemans

Monique Descamps

Marguerite-Marie Hanon

Sophie Malali

3. Research Activities

3.1 Presentation

The research areas in which the members of the Institute of statistics are working are very diverse. The main areas of expertise are non and semi-parametric regression techniques, time series analysis, survival analysis, industrial statistics, extreme value analysis and statistics for the actuarial sciences.

In the context of non- and semi-parametric regression, the Institute is a leading expert in the area of frontier estimation, and in particular in the application of non- and semi-parametric approaches for this problem in the context of efficiency analysis. A lot of research is also carried out in the context of inverse problems in econometrics, and its applications in instrumental regression and deconvolution problems. The study of semi-parametric regression models (e.g. single index models, partial linear models) and of nonparametric location-scale models, is another area in which the Institute is taking a leading role. Denoising and functional image clustering by means of wavelets are also highly studied.

The analysis of time series is a second cornerstone of the research activities at the Institute. The focus lies on the modelling and analysis of non-stationary time series, multivariate (high-dimensional) time series, factor models, volatility models, spectral density estimation and goodness-of-fit methods. Also, applications in statistical signal processing, and biomedical, economical and financial time series are studied.

The analysis of data coming from medical or industrial studies is also a research topic to which much attention is paid at the Institute. Medical data are often subject to censoring (survival analysis). The non- and semi-parametric modelling of this type of data is studied in detail, both the asymptotics for these models, as the application to medical data. In the context of industrial statistics, the focus lies on experimental design and multicriteria optimisation with applications in drug discovery, and on the analysis and modelling of time intensity curves in sensometrics.

The development of methods for extreme value analysis is also of interest at the Institute : modelling of extremes in univariate and multivariate time series, and in particular in Markov chains.

The modelling of dependence by means of copulas, the development of stochastic inequalities (stochastic orders, stochastic extrema), and the application of statistical concepts to the research area of insurance are also highly studied.

3.2 Research contracts

3.2.1 Research projects under contracts and cooperation projects

This section discusses ongoing research projects and cooperation projects at the Institute de statistique that are financed by outside agencies in the form of grants and contracts.

Statistical techniques and modelling for complex substantive questions with complex data (2002-2006)

Financing : Interuniversity Attraction Poles Programmes, Belgian Science Policy, Brussels, Belgium

Coordinators : L. Simar and I. Van Keilegom

STAT Researchers : C. Heuchenne (until September 15), H. Böhm and D. Varron (until August 31)

Partners Institutions :
Katholieke Universiteit Leuven, Belgium
UHasselt, Belgium
Université Libre de Bruxelles, Belgium
Aachen Technical University, Germany
Université Joseph Fourier, Grenoble, France

The point of departure of the network activities is that of a broad range of complex substantive data sets and questions arising in various disciplines (including psychology, biomedical sciences, economics and climatology).

The overall aim of our project is to develop appropriate statistical models and techniques to deal with these data and questions.

Local modelling and estimation in volatile highdimensional statistical systems (2004-2006)

Financing : Fonds Spéciaux de recherche (FSR)

Grant holder : R. von Sachs

STAT Researcher : H. Böhm

The goal of this project is to develop statistical methods to model and estimate multivariate statistical signals with a time-varying variance-covariance structure.

In particular the problem to dimension reduction is addressed, where Principal Component Analysis is an important tool. In order to derive a sound mathematical theory of treating the non-stationary nature of the observed data, the concept of local stationarity is employed. Applications to classical PCA in the time domain as well as to multivariate analysis of non-stationary EEG signals in the frequency domain will be treated subsequently.

Policies for research and innovation in the move towards the european research area (2005-2008)

Financing : European Union

Grant holder : L. Simar

STAT Researcher : L. Simar

Using advanced quantitative methods for evaluating the performance of Public Sector Research. Analysing in particular the system of the European universities and the availability of national data at the microlevel.

Adaptive modelling of the dependence in multivariate time series (2006-2007)

Financing : Fonds spéciaux de recherche (FSR)

Grant holder : C. Hafner
Principal investigators : S. Van Bellegem, I. Van Keilegom

STAT Researcher : O. Reznikova

The project intends to investigate theoretically and empirically a very flexible class of models for multivariate time series. We would like to continue multivariate volatility models with a copula approach for the dependence of the error term. Estimation of the dependence parameter should be done an adaptive way.

Implementation of a quality control plan for the monitoring of blood treatment processes for the Belgian Red Cross (2006-2007)

Financing : Belgian Red Cross

Grant holder : B. Govaerts

STAT researchers : A. Guillet and A. Benoit

The Belgian Red Cross has to develop quality control procedures to monitor its blood treatment processes in order to be in conformity with the Belgian legislation. This project implies the adaptation of statistical quality control techniques like control charts, capability analysis and reception sampling to the particular problematic of blood treatment. The project aims to develop the necessary tools and implement them in an adapted computing tool.

Development of the industrial field of renewable energy in the Walloon region : set-up of an inventory of existing initiatives (2006)

Financing : ECODE

Grant holders : B. Govaerts and M.-P. Kestemont

STAT participant : A. Guillet

Nowadays, renewable energies are at the centre of politic of energy decisions. In particular, the Belgian Walloon region aims at developing their use in Wallonia and at promoting the development of regional industries and jobs in this area. In order to orient political decisions, it has charged the society ECODE to set-up an inventory of existing initiatives in the domain in Wallonia. In the context, Ecode has collaborated with the Institute of statistics of UCL for the design and analysis of a survey.

Modelling dynamics and volatilities of multivariate economic time series (2005-2007)

Financing : Fonds spéciaux de recherche (FSR)

Grant holders : R. von Sachs, L. Simar, S. Van Bellegem and L. Bauwens

STAT Researcher : A. Daskovska

Through a multivariate modelling framework, we aim to understanding and prediction the temporal dependence in (co)-volatilities of financial time series.

Two main questions are addressed : How can we multivariate models be sufficiently flexible to incorporate dynamics of high-dimensional econometric data changing smoothly or abruptly over time ?

How can the rapidly growing model complexity (dimensionality, number of parameters) be efficiently controlled ?

3.2.2 Applied research contracts

The Institute of statistics is developing many collaborations within the Université catholique de Louvain and with several companies in the field of applied statistics. In addition to seminars organized weekly (see point 5.3), the members of the Institute participate to research contracts in applied statistics and offer consulting services to other departments and institutions of the University (see point 5.2). They also offer some courses of continued education (see point 6) at the University and in companies.

Statistical analysis of PET scan data (2003-2007)

Financing : Eli Lilly (July 2003 till June 2004), FSR (July 2004 to June 2005), Eli Lilly (July 2005 till June 2006)

Grant holder : Ph. Lambert

STAT Researcher : A. Jullion

The goal of this research is to build statistical longitudinal models for the analysis of Positron Emission Tomography (PET) scan data. In particular, we plan to set up methods leading to new definitions of receptor occupancy and enabling to quantify associated measures of precision.

Modern risk management models for insurance companies and pension funds (2004-2009)

Financing : Communauté française Wallonie-Bruxelles, ARC project

Grant holders : M. Denuit, P. Devolder, J-M. Rolin and Y. Smeers

Researchers : C. Courtois, D. Hainaut, A. Miller and J-P. Boucher

Project description :

1. Generalized life insurance policies
2. Risk theory under partial information
3. Unification of actuarial and financial pricing techniques
4. Risk management in incomplete markets

Development of computer solutions in the field of risk management (2003-2006)

Financing : FIRST SPIN-OFF Programme, Walloon Ministry

Grant holder : M. Denuit

Researcher : X. Maréchal

The project aims to develop realistic valuation models different from the usual actuarial vision which looks for explicit solutions in quite simple models. Numerical techniques will be used (simulations, MCMC methods) and proved to be more accurate than classical analysis techniques. Applications will deal with mortality projections, non linear effects in IARD tarification, an predictive laws in credibility theory.

Functional Statistical Inference. New approaches in Finance, Environmental Science, Industry and Economy
(2005-2008)

Financing : Spanish Ministry of Education and Science

Researcher : I. Van Keilegom

Many efforts have been devoted, in recent times, to the study of statistical inference on curves. In this project, our aim is to develop new theoretical methodologies in order to make inference on curves, with applications in different frameworks : finite populations, goodness-of-fit testing, neural networks, machine learning, functional data analysis, set estimation, spatial and spatio-temporal statistics and finance series. Many of the new methods to be developed, will be implemented for the first time. They may also represent modifications on previous versions and their validity will be tested by theoretical results and simulation studies. Applications will be made on different settings as environmental science (SO₂ predictions in a power plant, concentration of heavy metals in mosses in Galicia, ...), Industry (control of wastewater treatment by anaerobic processes) and Finance (semiparametric inference in portofolio design, ..)

3.3 Reprints

- RP255 Bonaccorsi, A., Daraio, C. and L. Simar,
Advanced indicators of productivity of universities. An application of robust nonparametric methods in Italian data.
Scientometrics, 66, pp. 389-410, 2006.
- RP256 Cao, R. and I. Van Keilegom,
Empirical likelihood tests for two-sample problems via nonparametric density estimation.
The Canadian Journal of Statistics, 34, pp. 61-77, 2006.
- RP257 San Martin, E., Mouchart, M. and J.-M. Rolin,
Ignorable common information, null sets and Basu's first theorem.
Sankhya : the Indian Journal of Statistics, 67, pp. 674-698, 2005.
- RP258 Patilea, V. and J.-M. Rolin,
Product-limit estimators of the survival function with twice censored data.
The Annals of Statistics, 34, pp. 925-938, 2006
- RP259 François, N., Guyot-Declerck, C., Hug, B., Callemien, D., Govaerts, B. and S. Collin,
Beer astringency assessed by time-intensity and quantitative descriptive analysis : influence of pH and accelerated aging,
Food Quality and preference, 17, pp. 445-452, 2006.
- RP260 Pardo-Fernandez, J.C. and I. Van Keilegom,
Comparison of regression curves with censored responses.
Scandinavian Journal of Statistics, 33, pp. 109-434, 2006.
- RP261 Zelenyuk, V.
Aggregation of Malmquist productivity indexes.
European Journal of Operational Research, 174, pp.1076-1086, 2006.
- RP262 Delouille, V., Jansen, M. and R. von Sachs,
Second-generation wavelet denoising methods for irregularity spaced data in two dimensions.
Signal Processing, 86, pp. 1435-1450, 2006.
- RP263 Courtois, C., Denuit, M. and S. Van Belleghem,
Discrete s -convex extremal distributions : theory and applications.
Applied Mathematics Letters, 19, pp. 1367-1377, 2006.
- RP264 Denuit, M., Purcaru, O. and I. Van Keilegom,
Bivariate archimedean copula models for censored data in non-life insurance.
Journal of Actuarial Practice, 13, pp. 5-32, 2006.
- RP265 Daraio, C. and L. Simar,
A robust nonparametric approach to evaluate and explain the performance of mutual funds.
European Journal of Operational Research, 175, pp. 516-542, 2006.
- RP266 Jeong, S.O. and L. Simar,
Lineary interpolated FDH efficiency score for nonconvex frontiers.
Journal of Multivariate Analysis, 97, pp. 2141-2161, 2006.
- RP267 Patilea, V. and J.-M. Rolin,
Product-limit estimators of the survival function for two modified forms of current-status data.
Bernoulli, 12, pp. 801-819, 2006.

- RP268 Rozet, E., Mertens, B., Dewe, W., Ceccato, O., Govaerts, B., Boulanger, B., Chiap, P., Strel, B., Crommen, J. and Ph. Lambert,
The transfer of a LC-UV method for the determination of fenofibrate and fenofibric acid in Lidoses : use of total error as decision criterion.
Journal of Pharmaceutical and Biomedical Analysis, 42, pp. 64-70, 2006.
- RP269 Oulhaj, A. and M. Mouchart,
The role of exogenous randomness in the identification of conditional models.
Metron – International Journal of Statistics, LXIV, pp. 253-271, 2006.
- RP270 Bugli, C. and Ph. Lambert,
Comparison between principal component analysis and independent component analysis in electroencephalograms modelling.
Biometrical Journal, 48, pp. 1-16, 2006.
- RP271 Eysers, L., Smoot, J.C., Smoot, L.M., Bugli, C., Urakawa, H., MacMurry, Z., Siripong, S., El Fantroussi, S., Lambert, Ph., Agathos, S. and D.A. Stahl,
Discrimination of shifts in a soil microbial community associated with TNT-contamination using a functional ANOVA of 16S rRNA hybridized to oligonucleotide microarrays.
Environmental Science & Technology, 40, pp. 5867-5873, 2006.
- RP273 Van Bellegem, S. and R. Dahlhaus,
Semiparametric estimation by model selection for locally stationary processes.
Journal of the Royal Statistical Society, Series B, 68, pp. 721-746, 2006.
- RP276 Geenens, G. and M. Delecroix,
A survey about single-index models theory.
International Journal of Statistics and Systems, pp. 203-230, 2006.
- RP277 Hafner, C. and H. Herwartz,
Volatility impulse response functions for multivariate GARCH models : An exchange rate illustration.
Journal of International Money and Finance, pp. 719-740, 2006.
- RP278 Hafner, C. and H. Herwartz,
A Lagrange multiplier test for causality in variance.
Economics Letters, 93, 137-141, 2006.
- RP279 Hafner, C. and O.B. Linton,
Discussion of quantile autoregression by Koenker and Xiao.
Journal of the American Statistical Association, 101, pp. 998-1001, 2006.

3.4. Discussion papers

The abstracts of the discussion papers have to be found in the Appendix.

- 0601 Russo, F., Mouchart, M., Ghins, M. and G. Wunsch
Statistical modelling and causality in the social sciences
- 0602 D. Varron
A limited in bandwidth uniformity for the functional limit law of the increments of the empirical process
- 0603 D. Varron
A non-standard uniform functional limit law for the increments of the multivariate empirical distribution function
- 0604 Jeong, S.-O., Park, B.U. and L. Simar
Nonparametric conditional efficiency measures : asymptotic properties
- 0605 D. Varron
Some uniform in bandwidth functional results for the tailuniform empirical and quantile processes
- 0606 Park, B.U., Simar, L. and V. Zelenyuk
Local likelihood estimation in truncated regression and its derivatives : theory and application
- 0607 D. Varron
Clustering rates and Chung type laws of the iterated logarithm for empirical and quantile processes
- 0608 Fryzlewicz, P., Nason, G.P. and R. von Sachs
A wavelet-Fisz approach to spectrum estimation
- 0609 Neumeyer, N. and I. Van Keilegom
Change-point tests for the error distribution in nonparametric regression
- 0610 Linton, O., Sperlich, S. and I. Van Keilegom
Estimation of a semiparametric transformation model
- 0611 Boulanger, G., Dewe, W., Gilbert, A., Govaerts, B. and M. Maumy
Risk management for analytical methods based on the local error concept : conciliating the objectives of the pre-study and in-study validation phases
- 0612 Lambert, Ph. and P.H.C. Eilers
Bayesian multi-dimensional density estimation with P-splines
- 0613 Almeida, C. and M. Mouchart
Bayesian encompassing specification test under not completely known partial observability
- 0614 Einmahl, J.H.J. and I. Van Keilegom
Goodness-of-fit tests in nonparametric regression
- 0615 Varron, D. and I. Van Keilegom
Uniform in bandwidth exact rates for a class of kernel estimators
- 0616 Chen, S.X. and I. Van Keilegom
A goodness-of-fit test for parametric and semiparametric models in multiresponse regression
- 0617 Bigot, J. and S. Van Bellegem
Log-density deconvolution by wavelet thresholding
- 0618 Ombao, H. and S. Van Bellegem
Coherence analysis of nonstationary time series : a linear filtering point of view

- 0619 Jullion, A., Lambert, Ph., Beck, B. and F. Vandenhende
Pharmacokinetic parameters estimation using adaptive Bayesian P-splines models
- 0620 El Ghouch, A. and I. Van Keilegom
Nonparametric regression with dependent censored data
- 0621 Bouezmarni, T. and J.V.K. Rombouts
Nonparametric density estimation for positive time series
- 0622 Roberts, G.O., Rosenthal, J.S., Segers, J. and B. Sousa
Extremal indices, geometric ergodicity of Markov chains, and MCMC
- 0623 Antoniadis, A., Bigot, J. and R. von Sachs
A multiscale approach for statistical characterization of functional images
- 0624 Motta, G., Hafner, C. and R. von Sachs
Locally stationary factor models : identification and nonparametric estimation
- 0625 San Martin, E. and M. Mouchart
On joint completeness : sampling and Bayesian versions, and their applications

3.5 Books published by the Members of the Institute

- Hafner, C. (1998)
"Nonlinear Time Series Analysis with Applications to Foreign Exchange Rate Volatility" – 222 p, Physica Verlag Heidelberg
- Florens, J.-P., Mouchart, M. and J.-M. Rolin (1999)
"Elements of Bayesian Statistics" – 544 p, New York : Marcel Dekker, 1990
- Härdle, W. and L. Simar (1993)
"Computer Intensive Methods in Statistics"– 175 pp. Vol I, Statistics and Computing, Physica-Verlag, Berlin, 1993
- Härdle, W., Klinke, S. and B.A. Turlach (1995)
"WploRE : An Interactive Statistical Computing Environment" – 387 p, Statistics and Computing, Springer-Verlag, New-York, 1995
- Fan, J. and I. Gijbels (1996)
"Local Polynomial Modelling and its Applications" – Chapman and Hall : London, 1996
- Kaas, R., Goovaerts, M.J., Dhaene, J. and M. Denuit (2001)
"Modern Actuarial Risk Theory" - Kluwer Academic Publishers, Dordrecht, 2001
- Hafner, C., Franke, J. and W. Härdle (2001)
"Einführung in die Statistik der Finanzmärkte" – 358 pp., Springer Verlag, 2001
- Wunsch, G., Mouchart, M. and J. Duchêne (2002)
"The Life Table : Modelling Survival and Death : Book series : European Studies of Population", vol. 11, Kluwer Academic Publishers, Dordrecht, 2002
- Härdle, W. and L. Simar (2003)
"Applied Multivariate Statistical Analysis" - 486 pp, Springer Verlag Berlin, 2003
- Beirlant, J., Goegebeur, Y., Segers, J. and J. Teugels (2004)
"Statistics of Extremes : Theory and Applications" – John Wiley & Sons
- Delwarde, A. and M. Denuit (2005)
"Construction de Tables de Mortalité Périodiques et Prospectives" - Audit-Actuariat-Assurances, Economica, Paris
- Denuit, M., Dhaene, J., Goovaerts, M.J. and Kaas, R. (2005)
"Actuarial Theory for Dependent Risk : Measures, Orders and Models" - Wiley, New York

3.6. Visitors

Doctoral Visitors

Alessandro Fiorini, Università di Roma, Roma, Italy
Elisa Maria Molanes Lopez, Universidad da Coruña, da Coruña, Spain

Short Terms Visitors

Antoniadis Anestis, Université Joseph Fourier, Grenoble, France
Bai Jushan, University of New-York, New-York, USA
Berthet Philippe, Université Rennes 1, Rennes, France
Berger Yves, University of Reading, Reading, UK
Bigot Jérémie, Université Paul Sabatier, Université Toulouse 3, Toulouse, France
Bouezmarni Taoufik, HEC Montréal, Montréal, Canada
Cabrera Jorge Luis Ojeda, Universidad de Zaragoza, Zaragoza, Spain
Cao Ricardo, Universidad da Coruña, da Coruña, Spain
Cavalier Laurent, Université Aix-Marseille 1, France
Dahlhaus Rainer, Universität Heidelberg, Germany
Delgado Miguel, Universidad Carlos III, Madrid, Spain
Delsol Laurent, Université Toulouse II, Toulouse, France
De Una Alvarez Jacobo, Universidade de Vigo, Vigo, Spain
Eichler Michael, Universiteit Maastricht, Maastricht, The Netherlands
Eilers Paul, Universiteit Leiden, Leiden, The Netherlands
Fève Frédérique, Institut d'Economie Industrielle, Université Toulouse 1, Toulouse, France
Florens Jean-Pierre, Institut d'Economie Industrielle, Université Toulouse 1, Toulouse, France
Franke Juergen, Universität Kaiserslautern, Kaiserslautern, Germany
Fryzlewicz Piotr, Bristol University, Bristol, UK
Gannaz Irène, Université Joseph Fourier, Grenoble, France
Gómez Guadalupe, Universitat de Catalunya, Barcelona, Spain
Hannig Jean, Colorado State University, Fort Collins, USA
Jouneau Frédéric, Université Lille 3, Lille, France
Kauermann Göran, Universität Bielefeld, Bielefeld, Germany
Kenward Mike, London School of Hygiene and Tropical Medicine, London, UK
Klein Thierry, Université Toulouse 3, Toulouse, France
Kneip Alois, Universität Bönning, Bönning, Germany
Leblanc Frédérique, Université Joseph Fourier, Grenoble, France
Long Xiandong, University of Cambridge, Cambridge, UK
Lopez Olivier, ENSAI Rennes, Rennes, France
Mammen Enno, Universität Mannheim, Mannheim, Germany
Meitz Mika, Stockholm School of Economics, Stockholm, Sweden
Ombao Hernando, University of Illinois, USA
Opsomer Jean, Iowa State University, USA
Oulhaj Abderrahim, Oxford University, Oxford, UK
Pardo-Fernandez Juan Carlos, Universidade de Vigo, Vigo, Spain
Proietti Tommaso, Università di Roma, Roma, Italy
Patilea Valentin, CREST-ENSAI, France
Scaillet Olivier, HEC Université de Genève, Genève, Switzerland
Steel Mark, University of Warwick, Warwick, UK
Tadjuidje Joseph, Universität Kaiserslautern, Kaiserslautern, Germany
Valenzuela Eduardo, Pontificia Universidad Católica de Chile, Chile
Vanhems Anne, École Supérieure de Commerce, Toulouse, France
Wilson Paul, University of Texas, Austin, USA
Yao Qiwei, London School of Economics, London, UK
Zakoian Jean-Michel, Université Lille, Lille, France
Zelenyuk Valentin, National University of Kiev, Kiev, Ukrainian

3.7. Statistics seminars

- February 10, 2006 - Special afternoon on « Applications of statistics in psychology »
 - Iven Van Mechelen, KULeuven, Leuven, Belgium
Modelling tools to summarize of individual differences in personality functioning and to understand its process basis (seminar jointly organised with the IAP network)
 - Francis Tuerlinckx, KULeuven, Leuven, Belgium
Diffusion and item response models : some examples of statistical modelling in psychological research (seminar jointly organised with the IAP network)

- February 17, 2006 - Special afternoon on « GARCH processes »
 - Jean-Michel Zakoian, Université Lille 3, Lille, France
Estimating and testing GARCH processes when the parameter is on a boundary
(seminar jointly organised with CORE)
 - Mika Meitz, Stockholm School of Economics, Stockholm, Sweden
Ergodicity mixing and existence of moment of a class of Markov with applications to GARCH and ACD models (seminar jointly organised with CORE)

- March 3, 2006 - Special afternoon on "Model Selection"
 - Laurent Cavalier, Université Aix-Marseille 1, Marseille, France
Risk hull Method in inverse problems
 - Arie Preminger, CORE, UCL, Louvain-la-Neuve, Belgium
A model selection method for S-estimation (seminar jointly organised with CORE)

- March 15, 2006
 - Tommaso Proietti, Università di Roma, Roma, Italy
Band spectral estimation for signal extraction (seminar jointly organised with CORE)

- March 24, 2006
 - Yves Berger, University of Reading, Reading, UK
Adjusted Jackknife for imputation under unequal probability sampling without replacement
 - Anestis Antoniadis, Université Joseph Fourier, Grenoble, France
Dimension reduction in functional regression with applications
(seminar jointly organised with the IAP network)

- April 21, 2006 - Special afternoon on « Empirical Processes »
 - Philippe Berthet, Université Rennes 1, Rennes, France
An empirical processes approach of sample fluctuations and asymptotic statistics
 - Thierry Klein, Université Toulouse 1, Toulouse, France
Concentration inequalities

- May 3, 2006
 - Mark Steel, University of Warwick, Warwick, UK
Bayesian nonparametric inference in time series (seminar jointly organised with CORE)

- May 5, 2006
 - Miguel Delgado, Universidad Carlos III, Madrid, Spain
Nonparametric tests for conditional symmetry in dynamic models
 - Roel Braekers, Universiteit Hasselt, Hasselt, Belgium
A conditional Koziol-Green model under dependent censoring
(seminar jointly organized with the IAP network)

- September 22, 2006 - Special afternoon on « Regularization methods in functional estimation »
 - Olivier Scaillet, HEC Genève, Genève, Switzerland
Tikhonov regularization for functional minimum distance estimators
(seminar jointly organised with CORE)
 - Christine De Mol, ULB, Bruxelles, Belgium
Sparsity-enforcing regularization (seminar jointly organised with the IAP network)

- October 6, 2006 – Special afternoon in the honour of Jean-Marie Rolin
 - Valentin Patilea, CREST-ENSAI, Rennes, France
Some recent challenges in nonparametric regression checks
 - Jean-Marie Rolin, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Some new product – limit estimators in survival analysis

- October 27, 2006
 - Michael Eichler, Universiteit Maastricht, Maastricht, The Netherlands
Identification and likelihood for recursive linear models with correlated errors
 - Frédéric Jouneau, Université Lille 3, Lille, France
Uniformly consistent and exact tests for semi-parametric single-index models

- November 10, 2006
 - Jan Hannig, Colorado State University, Fort Collins, USA
On fiducial inference – the good, the bad and the ugly

- November 24, 2006 – Special afternoon on « Incomplete Data »
 - Guadalupe Gómez, Universitat de Catalunya, Barcelona, Spain
Analysis of sequential times subject to dependent censoring
(seminar jointly organised with the IAP network)
 - Jorge Cabrera, Universidad de Zaragoza, Spain
A comparison of two wild-bootstrap schemes on goodness of fit tests in presence of selection bias (seminar jointly organised with the IAP network)

- December 8, 2006 - Special afternoon on « Time series » (organized with 3CC)
 - Xiangdong Long, University of Cambridge, Cambridge, UK
Estimation of dynamic conditional covariance : a semiparametric multivariate model
 - Qiwei Yao, London School of Economics, London, UK
Modelling multiple time series via common factors (seminar jointly organised with CORE)

- December 13, 2006
 - Jushan Bai, University of New York, New York, USA
Instrumental variable estimation in a data rich environment (seminar jointly organised with CORE)

3.8 Organization of Scientific Meetings

Fifth IAP Workshop on Flexible Statistical Analysis to Complex Data Structures

Louvain-la-Neuve, Belgium, March 16-17, 2006

Organizers : Anestis Antoniadis, Université Joseph Fourier, Grenoble, France
Christian Hafner, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Cédric Heuchenne, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Sébastien Van Belleghem, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Ingrid Van Keilegom, Institut de statistique, UCL, Louvain-la-Neuve, Belgium

Program

Thursday, March 16

Invited talk

Ricardo Cao, Universidad da Coruña, Spain
Relative curve estimation in survival analysis

Contributed talks

Anouar El Ghouch, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Nonparametric censored regression with dependent data

Frédérique Leblanc, Université Joseph Fourier, Grenoble, France
Maximum Likelihood estimation in Poisson regression via wavelet model selection

Davit Varron, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Asymptotic confidence bands for the conditional c.d.f. empirical likelihood

Invited talk

Mike Kenward, London School of Hygiene and Tropical Medicine, London, UK
The problem of missing data and some practical applications and issues with multiple imputation

Contributed talks

Silvia Cecere, KULeuven, Leuven, Belgium
Analyzing the emergence times of permanent teeth : an example of Bayesian modelling of a Covariance matrix with interval-censored data

Harriet Namata, Universiteit Hasselt, Hasselt, Belgium
Estimation of microbial risk using AIC model averaging approach

Kristel Van Steen, Universiteit Gent, Belgium
Genomic screening in family based association testing using the same data set and the multiple testing problem

Invited talk

Hernando Ombao, University of Illinois, Urbana-Champaign, USA
Localized feature selection for discrimination and classification of non-stationary signals

Contributed talks

Guy Mélard, ULB, Bruxelles, Belgium
One-line estimation of the parameters of a time model with applications

Dirk Depril, ULB, Bruxelles, Belgium
One-mode additive clustering for two-way two-mode data : a comparison of algorithms

Juergen Franke, Universität Kaiserslautern, Kaiserslautern, Germany
Mixture of stationary nonlinear time series

Friday, March 17

Invited talk

Enno Mammen, Universität Mannheim, Mannheim, Germany
Partial identification and nonparametric estimation of nonseparable, nonmonotonic functions

Contributed talks

Alejandro JaraVallejos, KULeuven, Leuven, Belgium
Semiparametric Bayesian analysis and missclassified binary data

Irène Gannaz, Université Joseph Fourier, Grenoble, France
Robust M-estimation and wavelet thresholding in partially binary data

Gery Geenens, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Index coefficients estimation in single-index models : the generalized maximum rank correlation Estimator

Invited talk

Göran Kauermann, Universität Bielefeld, Bielefeld, Germany
Mixed models and smoothing

Contributed talks

Wim Van den Noortgate, KULeuven, Leuven, Belgium
Assessing and explaining differential item functioning using logistic mixed models

John T. Maringwa, Universiteit Hasselt, Hasselt, Belgium
Analysis of a cardiovascular safety experiment with longitudinal data using penalized splines

Benoit Beck, Eli Lilly & Company, Belgium
Bayesian evolving sample size evaluation for screening experiments

Invited talk

Alois Kneip, Universität Bonn, Bonn, Germany
Common functional principal components

Contributed talks

Laurent Delsol, Université Toulouse III, Toulouse, France
Asymptotic normality in nonparametric time series analysis

Zita Oravecz, KULeuven, Leuven, Belgium
Modelling core affect space trajectories with the Ornstein-Uhlenbeck diffusion process

Christine De Mol, ULB, Bruxelles, Belgium
Forecasting using a large number of predictors

14th Annual Meeting of the Belgian Statistical Society

Houffalize, Belgium, October 11-12-13, 2006

Local organising Committee : Bernadette Govaerts, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Adelin Albert, ULg, Liège, Belgium
Christine Denayer, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Monique Descamps, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Christian Hafner, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Céline Le Bailly de Tillegem, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Régiane Rousseau, Institut de statistique, UCL, Louvain-la-Neuve, Belgium

Scientific Committee : Jean-Marie Rolin, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Christian Ritter, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Sébastien Van Bellegem, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Hilmar Böhm, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Cécile Dubois, UCB Pharma, Biostat section of SBS, Belgium
Peter Goos, UA, B-ENBIS, Belgium
Annie Robert, UCL, Quetelet Society, Belgium
Ingrid Van Keilegom, Institut de statistique, UCL, Louvain-la-Neuve, Belgium

Program

Wednesday, October 11

Is there a life after the Ph.D. in statistics ?

Karliën Vanden Brande, consultant at a DG, European Commission
Gunter Matthys, quantitative analyst structured credit products at Fortis
François Vandenhende, senior researcher at Eli Lilly, Louvain-la-Neuve

Thursday, October 12

Sessions RS3cc and Methodology

Invited speaker : Jean Opsomer, Iowa State University, USA
Applications of nonparametric regression in survey estimation

Delphine Cassart, ULB, Bruxelles, Belgium
A class of multivariate parametric and non parametric tests for symmetry

Dean S. Barron, California State University, USA
Kolmogoroff-Smirnoff enhancement

Julia Dony, VUB, Bruxelles, Belgium
Uniform in bandwidth consistency of conditional U-statistics

Huijuan Ding, KEL, Leuven Belgium
Lack-of-fit tests using wavelet regression in semiparametric mixed models

David Veredas, ULB, Bruxelles, Belgium
Indirect estimation in elliptical fatailed distributions

Dimitris Rizopoulos, KUL, Leuven, Belgium
Sensitivity analysis for the shared parameter models using copulas

Roula Tsonaka, KUL, Leuven, Belgium
Shared parameter models with a flexible random effects distribution

An Creemers, Universiteit Hasselt, Belgium
A new nonparametric approach to weighted estimating equations for missing data

Tim Verdonck, Universiteit Antwerp, Belgium
Principal component analysis for data containing outliers and missing elements

Peter Goos, Universiteit Antwerp, Belgium
Bayesian Analysis of data from split-plot designs

Session Biostatistical Section of BSS and Methodology

Invited speaker : Dave Collett, UK Transplant, UK
Some statistical problems in organ donation and transplantation

Beatrijs Moerkerke, Universiteit Gent, Belgium
A powerful strategy for detecting differentially expressed genes

Dirk Valkenborg, Universiteit Hasselt, Belgium
Monoisotopic peak validation using a Poisson approximation in a COFRADIC setting

Tom Jacobs, Universiteit Hasselt, Belgium
In-vivo – In-vitro correlation data modelling and random effects : a case study

Nezar Bennala, ULB, Bruxelles, Belgium
Sign and rank optimal test for error components models

Hilmar Böhm, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Shrinkage as a tool in multivariate spectral analysis

Geert Dhaene, KUL, Leuven, Belgium
Jackknife bias reduction for nonlinear dynamic panel data models with fixed effects

Astrid Jullion, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
PK parameters estimation using adaptive Bayesian P-splines models

Julie Francart, EORTC Data Center, Belgium
Prognostic factors in malignant mesothelioma of progression-free survival : a better approach for phase II mesothelioma trials

Katrien Antonio, KUL, Leuven, Belgium
Dependence modelling with multilevel data on fleet policies

Etienne Cuvelier, FUNDP, Namur, Belgium
Probability distribution and density for functional random variable. A new tool for functional data analysis

Invited speaker : Erich Neuwirth, Universität Wien, Wien, Austria
When methodology and data influence each other : The Pisa study

Friday, October 13

Session B-ENBIS and Quetelet Society

Invited speaker : Bart Baessens, KUL, Leuven, Belgium
New trends and applications in data mining

Serge Waterschoot, Banksys, Belgium
Fraud detection and reaction at Banksys

Nicolas Glady, KUL Leuven, Leuven, Belgium
Churn prediction using customer lifetime value : a business case in retail banking

Inge Koch, Universiteit Antwerp, Antwerpen, Belgium
An alternative measure for positive dependency useful in insurance and finance

Julie Perelman, UCL, Louvain-la-Neuve, Belgium
Deriving a risk-adjustment formula for hospital financing integrating the impact of socio-economic status and severity on length of stay

Simon De Jaeger, ULg, Liège, Belgium
Do voluntary agreements with municipalities help to reduce residential garbage ? A case study for Flanders

Sophie Vanbelle, ULg, Liège, Belgium
Comparing correlated kappa coefficients by the bootstrap method

Manuel Piette, Fortis, Belgium
The logit target mean approach for binary classification

Ann De Schepper, Universiteit Antwerp, Antwerpen, Belgium
Upper and lower bounds for the value at risk under partial information

Guy Brys, Statistics Belgium, Belgium
The GGP project

Philippe Haldermans, Universiteit Hasselt, Hasselt, Belgium
Using linear mixed models for normalization of cDNA microarrays

Invited speaker : Ziv Shkedy, Universiteit Hasselt, Hasselt, Belgium
Statistical and mathematical models in epidemiology

Satellite workshop presented by Christian Ritter, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
On Deployment of Statistics in an Office Environment – a hands-on-workshop

3.9 Academic visits

The members of the Institute visited other institutions and most of them presented seminars.

January 2006

Sébastien Van Bellegem, "Instrumental regression in partially linear models", Université Paul Sabatier, Toulouse, France (January, 10)

Ingrid Van Keilegom, "Estimation of a semiparametric transformation model", University of Minneapolis, USA (January, 19)

Ingrid Van Keilegom, "Estimation of a semiparametric transformation model", Iowa State University, USA (January, 25)

February 2006

Sébastien Van Bellegem, "Instrumental regression in partially linear models", KULeuven, Leuven, Belgium (February, 24)

Ingrid Van Keilegom, "Estimation of a semiparametric transformation model", Columbia University, USA, (February, 2)

Rainer von Sachs, "A multiscale approach for statistical characterization of brain response images", Fachbereich Mathematik, Universität Kaiserslautern, Kaiserslautern, Germany (February, 1-28)

Rainer von Sachs, "A multiscale approach for statistical characterization of brain response images", Institut für Angewandte Mathematik, Universität Heidelberg, Heidelberg, Germany (February, 16)

March 2006

Ingrid Van Keilegom, "Estimation of a semiparametric transformation model", ENSAI Rennes, Rennes, France (March, 10)

April 2006

Philippe Lambert, Actuarial Mathematics and Statistics School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, Scotland (April, 3-7)

Léopold Simar, "Efficiency Analysis : The Econometric Approach", Department of Statistical Science, Università di Bologna, Bologna, Italy (April, 2-7)

Sébastien Van Bellegem, Institut für Angewandte Mathematik, Universität Heidelberg, Heidelberg, Germany (April, 3-7)

Rainer von Sachs, "A multiscale approach for statistical characterization of brain response images", Statistics Department, University of Illinois, Urbana-Champaign, USA (April, 30 – May 31)

May 2006

Léopold Simar, Institut d'Economie Industrielle (IDEI), Université des Sciences Sociales, Toulouse I, France

Sébastien Van Bellegem, "Local adaptive estimation for nonstationary time series", Department of Economic Statistics and Decision Support, Stockholm School of Economics, Stockholm, Sweden (May, 18)

Ingrid Van Keilegom, "Estimation of a semiparametric transformation model", Universidad de Vigo, Vigo, Spain (May, 25)

Philippe Lambert, "Bayesian multidimensional estimation with P-splines", Département de Mathématiques, Université de Liège, Liège, Belgium (May, 26)

Giovanni Motta, invited speaker, "Dynamic Factor Models and Dimension Reduction", Göteborg, Sweden (May, 18-25)

July 2006

Léopold Simar, Laboratory of Economics and Management, Santa-Anna School of Advanced Studies and, Università di Pisa, Faculty of Engineering, Pisa, Italy

Léopold Simar, Summer School on "Advanced Quantitative Methods for Productivity and Efficiency Analysis", Università di Pisa, Faculty of Engineering, Pisa, Italy (July, 10-14)

Ingrid Van Keilegom, "Estimation of a semiparametric transformation model", Weierstrass Institute, Berlin, Germany (July, 5)

Ingrid Van Keilegom, "Change-point tests for the error distribution in nonparametric regression", Universidad da Coruña, da Coruña, Spain (July, 10)

September 2006

Léopold Simar, Institute of Economics and Social Sciences, Universität Bonn, Bonn, Germany (September 26 – October, 10)

Sébastien Van Belleghem, Université Toulouse I, IDEI, France (September, 23-30)

October 2006

Sébastien Van Belleghem, Department of Statistics and Center of Measurement, Pontificia Universidad Católica de Chile, Santiago, Chile (October, 25 – November, 4)

November 2006

Léopold Simar, Lectures on "An invitation to the Bootstrap : Panacea for Statistical Inference ?" and "Recent Advances in Multivariate Data Analysis" Laboratory of Economics and Management, Santa-Anna School of Advanced Studies and Facoltà di Ingegneria , Università di Pisa, Pisa, Italy (November, 17-23)

Léopold Simar, Institut für Statistik und Ökonometrie, Humboldt-Universität zu Berlin, Germany (November, 2-9)

December 2006

Christian Hafner, Tinbergen Institute, "Asymptotic Theory of Multivariate GARCH Models", Amsterdam, The Netherlands

Ingrid Van Keilegom, "A goodness-of-fit test for semiparametric models in multiresponse regression", University of Neuchâtel, Neuchâtel, Switzerland (December, 19)

3.10 Participation to conferences and scientific meetings

The members of the Institute assisted and/or participated to the following conferences.

March 2006

Gery Geenens, Workshop V PAI, talk, "Index coefficients estimation in single-index models : the Generalized Maximum Rank Correlation Estimator", UCL, Louvain-la-Neuve, Belgium

Céline le Bailly de Tillegem, Workshop V PAI, poster, "Uncertainty propagation in multiresponse optimization using a desirability index", UCL, Louvain-la-Neuve, Belgium

Anouar El Ghouch, Workshop V PAI, contributed talk, "Nonparametric censored regression with dependent data", UCL, Louvain-la-Neuve, Belgium

April 2006

Bernadette Govaerts, 8th Chemometrics symposium of the Belgian Chemometrics society, poster, "A statistical global view of a Time-Intensity case study" (with N. François and C. Guyot), Bruxelles, Belgium

May 2006

Bernadette Govaerts, 4th International chemometrics research meeting, poster, "A statistical global view of a Time-Intensity case study" (with N. François and C. Guyot) , Veldoven, The Netherlands

June 2006

Almeida Carlos, ISBA 8th World Meeting on Bayesian Statistics, poster, "Bayesian encompassing test under partial observability", Benidorm, Spain

Michel Denuit, International Conference on Mathematical and Statistical Modelling , ICMSM 2006, speaker, "Hurdle and zero-inflated models for panel data", Ciudad Real, Spain

Philippe Lambert, ISBA 8th World Meeting on Bayesian Statistics, Benidorm, Spain

Céline le Bailly de Tillegem, BNBIS Workshop, "Taking into account model uncertainty in multicriteria optimization", Antwerpen, Belgium

Giovanni Motta, 50 Years of Econometrics, Econometric Institute, Rotterdam, The Netherlands

Léopold Simar, Institute for Innovation and Knowledge Management, Universidad Polytechnica de Valencia, Spain

Léopold Simar, AQUAMETH meeting, PRIME network of Excellence, 6th EC Framework Programme "Efficiency and Productivity in European Universities" , speaker

Léopold Simar, North American Productivity Workshop IV, Member of the scientific Committee and invited discussant, "Improving the Performances of Parametric Estimators of Frontiers : from Nonparametric to Parametric Models" , Stern School of Business, New-York University, New York, USA

Rainer von Sachs, 4th Graybill Conference on "Multiscale methods and Statistics – a productive marriage", keynote speaker, Fort Collins, Colorado, USA

July 2006

Michel Denuit, 10th International Congress on Insurance : Mathematics and Economics, speaker, "Dynamic convex bounds, with applications to pricing in incomplete markets", Leuven, Belgium

Philippe Lambert and P.H.C. Eilers - "Bayesian multi-dimensional density estimation with P-splines". 21th International Workshop on Statistical Modelling. ISBN 1-86220-180-3. Galway, Ireland

Bianca Teodorescu, XXVI European Meeting of Statisticians, "Goodness-of-fit tests for conditional models with time-dependent coefficients under censoring and truncation", Torun, Poland

Ingrid Van Keilegom, XXVI European Meeting of Statisticians, invited, "Estimation of a semiparametric transformation model, Torun, Poland

August 2006

Christian Hafner, 61st European Meeting of the Econometric Society, "Testing for vector autoregressive dynamics under heteroskedasticity", Vienna, Austria

Léopold Simar, 61st European Meeting of the Econometric Society, Vienna, Austria

Sébastien Van Bellegem, 61st European Meeting of the Econometric Society, "Projection estimation in nonparametric instrumental regression", Vienna, Austria

September 2006

Michel Denuit, 4th Samos Conference in Actuarial Science & Finance, speaker, "S-(convex-concave) orderings on an arbitrary grid with applications to finance an actuarial science", Samos, Greece

Bernadette Govaerts, "Non-clinical statistics Conference - Analytical method validation based on the total error concept : comparison of alternative statistical approach", oral presentation, Postdam, Germany

Bernadatte Govaerts, "Analyzing the results of a designed experiment when the response is a curve", oral presentation, Southampton, UK

Christian Hafner, Journées Modélisation Aléatoire et Statistique de la Société de Mathématiques Appliquées et Industrielles, "Asymptotic Theory for a Factor GARCH Model", invited speaker, Lille, France

Céline le Bailly de Tillegem, Annual Non-Clinical Statistics Conference, oral presentation, "Taking into account prediction error in the virtual optimization of molecules using QSAR models", Potsdam, Germany

Céline le Bailly de Tillegem, Designed Experiments : Recent Advances in Methods and Applications, "Uncertainty Propagation in Multiresponse Optimization using a Desirability Index", Southampton, UK

Christian Ritter, Interdisciplinary workshop on statistical consulting at the ZIF, Universität Bielefeld, Bielefeld, Germany

Christian Ritter, Enbis Conference, Workshop on presentation on data and models, invited, Wroclaw, Poland

Léopold Simar, 4th Workshop on Indicators in the Knowledge Economy, invited speaker, "Robust alternatives to Estimate Benchmark Frontiers", KULeuven, Center of Economic Studies, Belgium.

October 2006

Hilmar Böhm, 14th Annual meeting of the Belgian Statistical Society, chairman fo the Young Researcher's afternoon, Houffalize, Belgium

Anouar El Ghouch, 14th Annual Meeting of the Belgian Statistical Society, poster, "Bayesian data fusion in a spatial prediction context : a general formulation", Houffalize, Belgium

Gery Geenens, 14th Annual Meeting of the Belgian Statistical Society, poster, "Nonparametric test for conditional independence in two-way contingency tables", Houffalize, Belgium

Christian Hafner, 14th Annual Meeting of the Belgian Statistical Society, poster, "Locally stationary factor models", Houffalize, Belgium

Astrid Jullion, 14th Annual Meeting of the Belgian Statistical Society, oral presentation, "PK parameters estimation using adaptive Bayesian P-splines models", Houffalize, Belgium

Céline Le Bailly de Tillegem, 14th Annual Meeting of the Belgian Statistical Society, poster, "Uncertainty propagation in multiresponse optimization using a desirability index", Houffalize, Belgium

Ingrid Van Keilegom, Journées d'Etudes en Statistique, invited, "Nonparametric approaches in regression", Marseille-Luminy, France

Ingrid Van Keilegom, 14th Annual Meeting of the Belgian Statistical Society, chairman of a session, Houffalize, Belgium

Ingrid Van Keilegom, Workshop on Nonparametric Statistics, invited, "Change-point tests for the error distribution in nonparametric regressions", Rennes, France

November 2006

Léopold Simar, Università della Svizzera italiana, Lugano, Italy
AQUAMETH Meeting on "Efficiency and Productivity in European Universities" and PRIME Conference on "Indicators on Science, Technology and Innovation. History and new Perspectives"

Ingrid Van Keilegom, Workshop on "Qualitative Assumptions and Regularization in High-Dimensional Statistics", invited, "A goodness-of-fit test for semiparametric models and models with shape constraints in multiresponse regression", Oberwolfach, Germany

December 2006

Bernadette Govaerts, Chimométrie 2006, "Analytical method validation based on the total error concept : comparison of alternative statistical approach", Paris, France

4. Teaching and Educational Activities

One major mission of the Institute of Statistics is educational and teaching services for the university. This activity concerns all levels of the statistical education and is described below. All professors and research assistants offer a strong investment in developing various pedagogical approaches that are adapted to the very large, heterogeneous public of the university. Moreover, because statistics is a very active field both in theory and applications, our courses are incessantly updated.

More specifically, our action in statistical education is organized at the following levels.

4.1 Bachelor level

The Institute of statistics is responsible for advising all Faculties about the education in statistics and data analysis in their Bachelor programs. On most of these programs, the Institute makes proposals about the content and the level of the course, and our team is responsible for teaching and grading in these courses. In 2006, the Institute was involved in the Bachelor course of about 3.350 students, with a total of about 900 hours of teaching in statistics.

4.2 Minor in statistics

Most students of UCL at the Bachelor level can choose an option in statistics called *Minor in statistics*. The goal of this option of 30 ECTS is to provide the basic knowledge in applied statistics that are useful in their specific field. That program is also a perfect preparation to the Master in statistics.

All information about the minor in statistics can be found at the page <http://www.ucl.ac.be/etudes/2006/mineures/minstat.html>

4.3 Master and DEA

Since 2005, the Institute organizes a *Master in statistics* that is accessible to students who have a degree at the 2nd level (Master, Licence), and who want to master the statistical tools that are useful in their main field. In order to fit to the variety of scientific fields using statistics, the master is divided into 5 specialities: (i) statistics in science and technology; (ii) Statistics in marketing and survey; (iii) statistics in economics and insurance; (iv) mathematical methods in statistics and (v) data management and data mining. The last speciality is organized jointly with the society SAS®. In 2006, 33 students attended the Master.

The Institute of statistics also proposes a DEA (Diplôme d'études approfondies), which is an advanced one-year degree in statistics. The DEA does not only train for professional life as a statistician, but also prepares to enter the Doctoral Program (see below). That program is organised in four themes: (i) Biostatistics and Epidemiology; (ii) Applied statistics; (iii) Statistical methodology and (iv) Statistics and Econometrics. In 2006, 16 students attended this program.

Both programs present a combination of courses, seminars, and research projects within the UCL or in collaboration with private organizations.

4.4 Towards the new Master programs in 2007

During 2006, the team of the Institute of statistics also prepared the Bologna reform and the new version of the master that will start in September 2007. An updated information about that new program can be found at: <http://www.uclouvain.be/66436.html>

4.5 Certificate

The Institute of statistics also proposes a *Certificate in statistics* that is mostly devoted to adult students. That program aims to provide the necessary tools for the professional in order to update his knowledge and techniques in statistics. Six orientations are proposed in order to fit to the field and the level in statistics of the interested candidate. That program is jointly organized with IUFC (Institut universitaire de formation continue) and more details can be found at <http://www.uclouvain.be/formationcontinue-statistique.html> . In 2006, 8 students attended the certificate in statistics.

4.6 Doctoral School

4.6.1. Presentation

In 2006, the Institute of Statistics succeeded to widen the scope of its doctoral school: The FNRS Graduate School in Statistics and Actuarial Sciences is now an interuniversity school with partners in 9 departments at 5 French-speaking universities (ULB, UCL, ULg, FUNDP, FUSAG), and the Institute of Statistics continues to play the leading role in the administration and organization of this School.

The "Bureau" of the School, composed of President (Rainer von Sachs, Institute of Statistics), Vice-President, Academic Secretary and Administrative Secretary (Sophie Malali, Institute of Statistics), meets regularly and is completed by two annual meetings of the Scientific Committee of the School – with representatives of both senior members and doctoral students of all partners of the School. Every academic member of the Institute of Statistics participates in the Graduate School, and of course every PhD student working at the Institute of Statistics on his or her thesis.

One of the objectives of the School is to offer our PhD students a high-quality training to and by research, under the guidance of their supervisors and a follow-up committee.

In 2006 the Graduate School has organized a growing number of short courses given by international visitors in various fields of methodological and applied statistics. These courses are at the heart of the training of our students to complete their education in statistics and its neighbouring fields, at an advanced level of graduate teaching ("3e cycle"). This offer is completed by regular research seminar series, young researchers days, workshops and summer schools, some of which organized by our PhD students themselves. Our students are encouraged to collaboration with international colleagues who regularly visit the Institute of Statistics, and (also financially supported) to present their work on behalf of national and international meetings.

These activities create for our students the opportunity to work in a stimulating environment, open to discussion and exchanges, in order to develop inter-disciplinary work based on a common methodology in a variety of fields in statistics, biostatistics and actuarial sciences.

As a consequence of the interuniversity character of our Graduate School, we are in the process of developing further our web page (currently at www.stat.ucl.ac.be/edt) – primary means of communicating information among our members about our activities.

Thanks to our administrative secretary maintaining this web page relevant information is spread out quickly in order to reach all the members of the School.

In its first year of existence, after a successful opening afternoon with short but very informative presentations of every research group, the new interuniversity school is still developing, hoping to attract new PhD students (partially from abroad). To this end a colourful information brochure has been published in September 2006 (available at <http://www.stat.ucl.ac.be/edt/flyer.pdf>).

4.6.2 Doctoral dissertations and Ph. D. thesis in Progress

Doctoral Dissertations

Céline Bugli (June 2006)

Statistical tools for the analysis of event-related potentials in electroencephalograms

Supervisor : Philippe Lambert

Ph.D. thesis in progress

Carlos Almeida

Structural equation modeling with categorical ordered variables

Supervisor : M. Mouchart

Hilmar Böhm

Model selection in principal component analysis for multivariate time series

Supervisor : Rainer von Sachs

Jean-Philippe Boucher

Non-life ratemaking with panel data

Supervisor : Michel Denuit

Cindy Courtois

Risk theory under partial information and dependence

Supervisor : Michel Denuit

Alexandra Daskovska

Dynamical analysis of productivity index

Supervisors : Léopold Simar and Sébastien Van Bellegem

Anouar El Ghouch

Empirical likelihood with incomplete data

Supervisor : Ingrid Van Keilegom

Nancy François

Statistical analysis of time intensity curves in sensory analysis

Supervisors : Bernadette Govaerts and Philippe Lambert

Gery Geenens

Conditional independence tests for two-way contingency tables

Supervisor : Léopold Simar

Astrid Jullion

Statistical analysis of PET (Positron Emission Tomography) scan data

Supervisor : Philippe Lambert

Maria Key Prato

Detection and quantification of treatment effect on blood pressure profile curves

Supervisor : Philippe Lambert

Céline le Bailly de Tillegem

Multiobjective optimization by computer simulations

Supervisors : Bernadette Govaerts and Léopold Simar

Giovanni Motta

Nonparametric estimation of time-varying covariance matrices for multivariate nonstationary time series

Supervisor : Rainer von Sachs

Sandra Pitrebois

Bonus-malus scales and segmented tariffs

Supervisors : Michel Denuit and Jean-François Walhin

Réjane Rousseau
Outils statistiques pour identification de biomarqueurs de toxicité métabonomiques
Supervisors : Bernadette Govaerts and Michel Verleysen

Bianca Teodorescu
Generalized linear conditional models under left truncation and right censoring
Supervisor : Ingrid Van Keilegom

Olga Reznikova
Adaptive modelling of the dependence in multivariate time series
Supervisor : Christian Hafner

4.6.3. Doctoral Seminars

- February 9, 2006
Thomas Laloux, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Matlab implementation of a specific pharmacokinetic model applied to medical imaging (PET technology)
- February 23, 2006
Gery Geenens, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Nonparametric testing for conditional independency in two-way contingency tables
- March 9, 2006
Almeida Carlos, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Bayesian encompassing specification test under partial observability
- March 10, 2006
Alexandra Daskovska, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Dynamical forecasting of Malmquist productivity index
- March 30, 2006
Giovanni Motta, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Locally stationary factor models
- March 31, 2006
Hilmar Böhm, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Enhanced estimation of Kullback-Leibler distance for multivariate time series
- April 20, 2006
Réjane Rousseau, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Statistical methods to identify metabonomics biomarkers in H-NMR stectroscopy : an application to an artificial database
- May 4, 2006
Céline Le Bailly de Tillegem, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Taking into account model uncertainty in multicriteria optimization : an application to molecules optimization
- May 18, 2006
Astrid Jullion, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
PK parameters estimation using adaptative Bayesian P-splines models

- October 19, 2006
Elisa María Molanes López, Universidade da Coruña, da Coruña, Spain
Bandwidth Selection for the Kernel relative density estimator
- November 30, 2006
Hilmar Böhm, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Shrinkage estimation in the frequency domain
- December 7, 2006
Alexandra Daskovska, Institut de statistique, UCL, Louvain-la-Neuve, Belgium
Dynamical analysis of the Malmquist productivity index

4.6.4 Short courses

- Anestis Antoniadis, Université J. Fourier, Grenoble, France
March 6-8-13-15, 2006
Support Vector Machines and Statistical Learning
- Hernando Ombao, University of Illinois, USA
June 21-23-28-29, 2006
Topics in Applied Time Series Analysis

4.7 Internationalization

4.7.1 Bilateral agreements

In 2006, a bilateral agreement with the Dipartimento di Scienze Statistiche of the Università di Bologna (Italy). The following exchanges were organized :

- One course by Professor A. Guizzardini from the Università di Bologna on *Business forecasting* was organized in Louvain-la-Neuve
- One course by Professor L. Simar (UCL) on *Productivity and Efficiency Analysis : the statistical approach* (April 2006)

We also concluded two new agreements with the following universities :

- Fachbereich Statistik in Universität Dortmund (Germany)
- Ecole nationale de la statistique et de l'analyse de l'information (ENSAI, Rennes, France)

4.7.2 Cooperation programs

- Sébastien Van Bellegem, course on *Analisis de datos de encuestas transversaels repetidas* (15h), Universidad Mayor de San Simón, Cochabamba, Bolivia (November 2006)

4.7.3. Mandate of invited professors

Michel Denuit, Groupe Consultatif Summer School 2006

Modelling mortality dynamics for pensions and annuity business under the auspices of the University of Parma, Italy (August 2006)

Forecasting mortality : Applications and Examples (September 2006)

Michel Denuit, Warsaw Actuarial Summer School 2006, Department of Economics, University of Warsaw, Poland

Projected life tables for insurance companies and pension funds (September 2006)

Léopold Simar, Professore di Chiara Fama, Pisa, Italian Ministry of Research

Sébastien Van Bellegem, Guest professor, Université Libre de Bruxelles, Bruxelles, Belgium (2006-2007)

Ingrid Van Keilegom, Guest professor, Universidade de Vigo, Vigo, Spain

Doctoral course on *Survival analysis* (May 2006)

5. Support for the University and internal Consulting Activities

5.1 Presentation

One fundamental mission of the Institute of statistics consists in offering the university community a service in statistical software and applied statistics. The main goal of the service is to encourage the use of modern and "fit for purpose" statistical tools in the university's various areas of applied research and stimulate interactions between researchers.

In this context, the Institute of statistics offers several services and activities.

The consulting service offers advice and collaboration to researchers and members of the university community concerning the appropriate methodology and suitable statistical package for their specific problems.

The statistical software server allows the members of UCL to manage their data bases and perform their statistical computations and simulations on an efficient and well maintained computing tool.

Training courses in statistical software are also proposed to inform researchers of existing statistical software and train them to their practical use.

Finally, in the context of its Master programs, the Institute of statistics is also organising a Statistical Workshop. Twice every month, speakers are invited to present applications of statistical methods in various fields. Some presentations take also the form of tutorials related to subjects of interest for people working in applied statistics.

5.2 Consulting activities

In 2006, the consulting service has offered numerous advices to members of the university community. Among them, more than 70 have been identified as "conseils" (advices of more than 2 hours of consulting work) and are listed on the consulting service webpage : <http://www.stat.ucl.ac.be/Sconsultation>. Clients come from various faculties and departments of the university (e.g. AGRO/BAPA, AGRO/MILA, ESPO/IAG, ESPO/SPED, FSA/MAPR, MD/ESP, MD/GYNE, MD/IEPR, PSP/CODE, PSP/ECSA, SC/BIOL, SC/CHIM, SC/GEOG...) and application fields are very stimulating for the statistical consultants.

In 2006, the consulting service has also conducted an analysis and comparison of several statistical software with the aim of replacing the SPSS software at UCL and the "Académie Louvain". JMP, Minitab, NCSS, SAS Enterprise Guide, SPSS, Stata and Statistica were tested and demonstrated to users of the different faculties. This evaluation conducted to a reinforcement of the use of SAS/Enterprise guide and a reduction of the SPSS site licence to basic modules. The "Académie Louvain" has now signed two new contracts with SAS and SPSS in the line of the recommendations made by the Institute of Statistics in this context and substantially reduced the related total costs.

5.3 Applied Statistics Workshops

- February 9, 2006
Fabienne Nackers, Unité Epidémiologie, UCL Woluwé, Belgium
Mesure de l'efficacité vaccinale du BCG pour la prévention de l'ulcère de Buruli au Bénin : Illustration de l'application de la statistique en l'épidémiologie
- February 16, 2006
Philippe Baret, Unité de Génétique, UCL, Louvain-la-Neuve, Belgium
Application of bioinformatics to the annotation of genomes : do we need statistical methods ?
- March 9, 2006
Bruno Schoumaker, Institut de Démographie, UCL, Louvain-la-Neuve, Belgium
L'analyse statistique des biographies en démographie : exemples à partir d'une enquête biographique sur les migrations au Burkina Faso
- March 30, 2006
Véronique Delouille, Royal Observatory of Belgium, Brussels, Belgium
Multifractal analysis of extreme ultraviolet images and motion detection using optical flow
- April 20, 2006
Laurence Collette, EORTC Data Center Belgium
Surrogate endpoints in cancer clinical trials
- May 4, 2006
Frédérique Fève and Jean-Pierre Florens, Institut d'Economie Industrielle, Université Toulouse 1, Toulouse, France
Economie et statistique du don des cellules souches hématopoïétiques
- May 11, 2006
Christian Valenduc and Frédéric Halleux, SPF Finance
Les facteurs explicatifs des déductions fiscales
- September 29, 2006
 - Christian Ritter, Institut de statistique, UCL, Belgium
On the presentation of quantitative information
 - Christian Ritter, Institut de statistique, UCL, Belgium
On the golden triangle of data analysis : a spreadsheet, a database, and a statistical package
- October 20, 2006
Linda Danielson, IDDI, Belgium, Bruno Boulanger, Eli Lilly, Belgium and François Beckers, GSK Biologicals, Belgium, Christian Ritter – Bernadette Govaerts, Institut de statistique, UCL, Belgium
The steps from university studies to working in/for the pharmaceutical industry : a panel discussion
- November 17, 2006
Benoît Beck, Eli Lilly, Belgium
Variation around the use of Sigmoidal curves for in Vivo Data Evaluations
- December 15, 2006
Benoît Rihoux and Airo Hino, Unité de science politique et de relations internationales, UCL, Belgium
QCA (Qualitative Comparative Analysis) : the overview and confrontation with statistical tools

6. Appendix

6.1. Abstracts of the discussion papers

DP 0601 – Statistical modelling and causality in the social sciences

Russo, F., Mouchart, M., Ghins, M. and G. Wunsch

Philosophers and statisticians have been debating on causality for a long time. However, these discussions have been led quite independently from each other. An objective of this paper is to pursue a fruitful dialogue between philosophy and statistics. As is well known, at the beginning of the 20th century, some philosophers and statisticians dismissed the concept of causality altogether. It will suffice to mention Bertrand Russell (1913) and Karl Pearson (1911). Almost a hundred years later, causality still represents a central topic both in philosophy and statistics.

In the social sciences, including research on public health, most studies are concerned with the possible causes, determinants, factors, etc. of a set of observations. In particular, for planning or policy reasons, it is important to know what causes which effects. In order to attain causal knowledge, many social scientists appeal to statistical modelling to confirm or disconfirm their hypotheses about possible causal relations among the variables they consider, taking care of controlling for relevant covariates and especially for possible confounding factors.

To what extent can a statistical model say something about causal relations among variables? In this paper, we will attempt an answer by examining a special class of statistical models, i.e. structural models. The discussion, however, will not be confined to a mere examination of statistical methods, since a considerable effort will be made to consider causality from an epistemological perspective. To put it otherwise, this paper does not address the nature of causation itself, nor the analysis of various causal structures, nor the elaboration of complex causal structures; rather, we will be concerned with the question of how we come to uncover causal relations by means of statistical modelling.

The practice of statistical modelling raises substantial issues of ontological nature. The latter is not the purpose of this paper. Nevertheless, as statistical models are a common way of gaining scientific knowledge, we begin in section I by stating a moderate realist position concerning the relationship between reality and observation. This point of view is not necessarily shared by all philosophers and scientists.

DP 0602 – A limited in bandwidth uniformity for the functional limit law of the increments of the empirical process

D. Varron

Consider the following local empirical process indexed by $K \in \mathcal{G}$, for fixed $h > 0$ and $z \in \mathbb{R}^d$:

$$G_n(K, h, z) := \sum_{i=1}^n K\left(\frac{Z_i - z}{h^{1/d}}\right) - \mathbb{E}\left(K\left(\frac{Z_i - z}{h^{1/d}}\right)\right),$$

where the Z_i are i.i.d. on \mathbb{R}^d . We provide an extension of a result of Mason. Namely, under mild conditions on \mathcal{G} and on the law of Z_1 , we establish a uniform functional limit law for the collections of processes

$$\{G_n(\cdot, h_n, z), z \in H, h \in [h_n, \mathfrak{h}_n]\},$$

where $H \subset \mathbb{R}^d$ is a compact set with nonempty interior and where h_n and \mathfrak{h}_n satisfy the Csörgő-Révész-Stute conditions.

DP 0603 – A nonstandard uniform functional limit law for the increments of the multivariate empirical distribution function

D. Varron

Let $(Z_i)_{i \geq 1}$ be an independent, identically distributed sequence of random variables on \mathbb{R}^d . Under mild conditions on the density of Z_1 , we provide a nonstandard uniform functional limit law for the following processes on $[0,1]^d$:

$$\Delta_n(z, h_n, \cdot) := s \mapsto \frac{\sum_{i=1}^n 1_{[0, s_1] \times \dots \times [0, s_d]} \left(\frac{Z_i - z}{h_n^{1/d}} \right)}{c \log n}, \quad s \in [0, 1]^d$$

along a sequence $(h_n)_{n \geq 1}$ fulfilling $h_n \downarrow 0, nh_n \uparrow, nh_n / \log c \rightarrow 0$.

Here z ranges through a compact set of \mathbb{R}^d . This result is an extension of a theorem of Deheuvels and Mason to the multivariate, non uniform case.

DP 0604 – Non parametric conditional efficiency measures : asymptotic properties

Jeong, S.-O., Park, B.U. and L. Simar

Daraio and Simar (2005a, b) developed a conditional frontier model which incorporates the environmental factors into measuring the efficiency of a production process. They also provided the corresponding nonparametric efficiency measures: conditional FDH estimator, conditional DEA estimator and conditional order- m estimator. The aim of this paper is to provide an asymptotic analysis of the first two estimators.

DP 0605 - Some uniform in bandwidth functional results for the tail uniform empirical and quantile processes

D. Varron

For fixed $t \in [0,1)$ and $h > 0$, consider the local uniform empirical process

$$D_{n,h,t}(s) := n^{-1/2} \left[\sum_{i=1}^n 1_{[t, t+hs]}(U_i) - hs \right], \quad s \in [0, 1]$$

where the U_i are independent and uniformly distributed on $[0,1)$. We investigate the functional limit

behaviour of $D_{n,h,t}$ uniformly in $h_n \leq h \leq h_n$ when $nh_n / \log \log n \rightarrow \infty$ and $h_n \rightarrow 0$.

DP 0606 – Local likelihood estimation of truncated regression and its partial derivatives : theory and application

Park, B.U., Simar, L. and V. Zelenyuk

In this paper we propose a very flexible estimator in the context of truncated regression that does not require parametric assumptions. To do this, we adapt the theory of local maximum likelihood estimation. We provide the asymptotic results and illustrate the performance of our estimator on simulated and real data sets. Our estimator performs as good as the fully parametric estimator when the assumptions for the latter hold, but as expected, much better when they do not (provided that the curse of dimensionality problem is not the issue). Overall, our estimator exhibits a fair degree of robustness to various deviations from linearity in the regression equation and also to deviations from the specification of the error term. So the approach shall prove to be very useful in practical applications, where the parametric form of the regression or of the distribution is rarely known.

DP 0607 – Clustering rates and Chung type laws of the iterated logarithm for empirical and quantile processes

D. Varron

Following the works of Berthet, we first obtain exact clustering rates for the functional law of the iterated logarithm (FLIL) for empirical processes and for their increments. In a second time, we obtain functional Chung-type limit laws for the local empirical process in the spirit of Deheuvels.

DP 0608 – A wavelet-Fisz approach to spectrum estimation

Fryzlewicz, P., Nason, G.P. and R. von Sachs

We suggest a new approach to wavelet threshold estimation of spectral densities of stationary time series. It is well known that choosing appropriate thresholds to smooth the periodogram is difficult because non-parametric spectral estimation suffers from problems similar to curve estimation with a highly heteroscedastic and non-Gaussian error structure. Possible solutions that have been proposed are plug-in estimation of the variance of the empirical wavelet coefficients or the log-transformation of the periodogram. In this paper we propose an alternative method to address the problem of heteroscedasticity and non-normality. We estimate thresholds for the empirical wavelet coefficients of the (tapered) periodogram as appropriate linear combinations of the periodogram values similar to empirical scaling coefficients. Our solution permits the design of "asymptotically noise-free thresholds", paralleling classical wavelet theory for nonparametric regression with Gaussian white noise errors. Our simulation studies show promising results that clearly improve the classical approaches mentioned above. In addition, we derive theoretical results on the near-optimal rate of convergence of the minimax mean-square risk for a class of spectral densities, including those of very low regularity.

DP 0609 – Change-point tests for the error distribution in nonparametric regression

Neumeyer, N. and I. Van Keilegom

In this paper several testing procedures are proposed that can detect change-points in the error distribution of nonparametric regression models. Different settings are considered where the change-point either occurs at some time point or at some value of the covariate. Fixed as well as random covariates are considered. Weak convergence of the suggested difference of sequential empirical processes based on nonparametrically estimated residuals to a Gaussian process is proved under the null hypothesis of no change point. In the case of testing for a change in the error distribution that occurs with increasing time in a model with random covariates the test statistic is asymptotically distribution-free and the asymptotic quantiles can be used for the test. This special test statistic can also detect a change in the regression function. In all other cases the asymptotic distribution depends on unknown features of the data generating process and a bootstrap procedure is proposed in these cases. The small sample performances of the proposed tests are investigated by means of a simulation study, and the tests are applied to a data example.

DP 0610 – Estimation of a semiparametric transformation model

Linton, O., Sperlich, S. and I. Van Keilegom

This paper proposes consistent estimators for transformation parameters in semiparametric models. The problem is to find the optimal transformation into the space of models with a predetermined regression structure like additive or multiplicative separability. We give results for the estimation of the transformation when the rest of the model is estimated non or semi-parametrically and fulfills some consistency conditions. We propose two methods for the estimation of the transformation parameter: maximizing a profile likelihood function or minimizing the mean squared distance from independence. First the problem of identification of such models is discussed. We then state asymptotic results for a general class of nonparametric estimators. Finally, we give some particular examples of nonparametric estimators of transformed separable models. The theoretical results as well as the small sample performance are studied by several simulation exercises.

DP 0611 – Risk management for analytical methods based on the total error concept : conciliating the objectives of the pre-study and in-study validation phases

Boulanger, B., Dewe, W., Gilbert, A., Govaerts, B. and M. Maumy

In industries that involve either chemistry or biology, analytical methods are necessary to keep an eye on all the material produced. If the quality of an analytical method is doubtful, then the whole set of decisions based on those measures is questionable. For this reason, being able to assess the quality of an analytical method is far more than a statistical challenge; it is a matter of ethics and good business practices. The validity of an analytical method must be assessed at two levels. The “pre-study” validation aims to show, by an appropriate set of designed experiments, that the method is able to achieve its objectives. The “in-study” validation is intended to verify, by inserting QC samples in routine runs, that the method remains valid over time. At these two levels, the total error approach considers a method as valid if a sufficient proportion of analytical results are expected to lie in a given interval around the (unknown) nominal value. This paper discusses two methods, based on this total error concept, of checking the validity of a measurement method at the pre-study level. The first checks whether a tolerance interval for hypothetical future measurements lies within given acceptance limits; the second calculates the probability of a result lying within these limits and computes whether it is greater than a given acceptance level. For the “in-study” validation, the paper assesses the properties of the $s-n-\lambda$ rule recommended by the FDA. The properties and respective advantages and limitations of these methods are investigated. A crucial point is to ensure that the decisions taken at the pre-study stage and in routine use are coherent. More precisely, a laboratory should not see its method rejected in routine use when it has been proved to be valid and remains so. This paper shows how this goal may be achieved by choosing compatible validation parameters at both pre- and in-study levels.

DP 0612 – Bayesian multi-dimensional density estimation with P-splines

Lambert, P. and P. Eilers

Polytomous logistic regression combined with spline smoothing gives a powerful tool for Bayesian density estimation. Using fast array algorithms, multiple dimensions can be handled in a fast and uniform way. The Langevin-Hastings algorithm allows efficient sampling from the associated (reparameterized) posterior distribution. Illustrations of density estimation are provided, as well as a new approach to smooth quantile regression.

DP 0613 – Bayesian encompassing specification test under not completely known partial observability

Almeida, C. and M. Mouchart

We present a Bayesian specification test in presence of partial observability and using the encompassing principle. The general theory is developed where the partial observability is known up to a Euclidean parameter, to be estimated from data. This general development is illustrated with an example where only a linear combination of a latent vector is observed; thus, in the example, the partial observability is known up to the vector defining the observed linear combination. Some identifiability issues are treated and the example shows the operability and the pitfall of the proposed test.

DP 0614 – Goodness-of-fit tests in nonparametric regression

Einmahl, J.H.J. and I. Van Keilegom

Consider the location-scale regression model $Y = m(X) + \sigma(X)\mathcal{E}$, where the error \mathcal{E} is independent of the covariate X , and m and σ are smooth but unknown functions. We construct tests for the validity of this model and show that the asymptotic limits of the proposed test statistics are distribution free. We also investigate the finite sample properties of the tests through a simulation study, and we apply the tests in the analysis of data on food expenditures.

DP 0615 – Uniform in bandwidth exact rates for a class of kernel estimators

Varron, D. and I. Van Keilegom

Given an i.i.d sample (Y_i, Z_i) , taking values in $\mathbb{R}^{d'} \times \mathbb{R}^d$, consider the quantities

$$W_n(g, h, z) := f_Z(z)^{-1/2} \sum_{i=1}^n \left[\left(\langle c_g(z), g(Y_i) \rangle + d_g(z) \right) K\left(\frac{Z_i - z}{h}\right) - \mathbb{E} \left(\left(\langle c_g(z), g(Y_i) \rangle + d_g(z) \right) K\left(\frac{Z_i - z}{h}\right) \right) \right],$$

where, z belongs to a compact set $H \subset \mathbb{R}^d$, f_z is the density of Z_1 , K is a kernel, $h > 0$ a bandwidth, g a Borel function on $\mathbb{R}^{d'}$ and $c_g(\cdot), d_g(\cdot)$ are continuous functions on \mathbb{R}^d .

Given two bandwidth sequences $h_n < h_n$ fulfilling mild conditions, we prove that, for an explicit constant, $\mathfrak{C}(\mathcal{G}, K)$ we have almost surely :

$$\lim_{n \rightarrow \infty} \sup_{z \in H, g \in \mathcal{G}, h_n \leq h \leq h_n} \frac{|W_n(g, h, z)|}{\sqrt{2nh^d \log(h^{-d})}} = \mathfrak{C}(\mathcal{G}, K)$$

under mild conditions on the density f_Z , the class \mathcal{G} , the kernel K and the functions $c_g(\cdot), d_g(\cdot)$.

We apply the result to the context of empirical likelihood, where regressions parameters are estimated with a smoothed version of empirical likelihood, involving a kernel K and a bandwidth h . Namely, we proved that smoothed empirical likelihood can be used to build confidence intervals for conditional probabilities $P(Y \in C | Z = z)$, that hold uniformly in $z \in H$, $C \in \mathcal{C}$, $h \in [h_n, h_n]$. Here \mathcal{C} is a Vapnik-Cervonenkis consistency.

DP 0616 – A goodness-of-fit test for parametric and semiparametric models in multiresponse regression

Chen, S.X. and I. Van Keilegom

We propose an empirical likelihood test that is able to test the goodness-of-fit of a class of parametric and semiparametric multiresponse regression models. The class includes as special cases fully parametric models, semiparametric models, like the multi-index and the partially linear models, and models with shape constraints.

Another feature of the test is that it allows both the response variable and the covariate be multivariate, which means that multiple regression curves can be tested simultaneously. The test also allows the presence of infinite dimensional nuisance functions in the model to be tested. It is shown that the empirical likelihood test statistic is asymptotically normally distributed under certain mild conditions and permits a wild bootstrap calibration. Despite that the class of models which can be considered is very large, the empirical likelihood test enjoys good power properties against departures from a hypothesized model within the class.

DP 0617 – Log-density deconvolution by wavelet thresholding

Bigot, J. and S. Van Bellegem

This paper proposes a new wavelet-based method for deconvolving a density. The estimator combines the ideas of nonlinear wavelet thresholding with Meyer wavelets and estimation by information projection. It is guaranteed to be in the class of density functions, in particular it is positive everywhere by construction. The theoretical optimality of the estimator is established in terms of rate of convergence of the Kullback-Leibler discrepancy over Besov classes. Finite sample properties is investigated in detail, and show the excellent practical performance of the estimator, compared with other recently introduced estimators.

DP 0618 – Coherence analysis of nonstationary time series : a linear filtering point of view

Ombao, H. and S. Van Bellegem

Coherence is a widely used measure for characterizing linear dependence between two time series. Classical books on time series analysis present coherence as “the frequency domain analogue of the autocorrelation function” which lacks intuitive appeal. The first goal of this paper is to present a more illuminating and yet still precise interpretation of coherence. Consider a filter whose power transfer function is concentrated on a particular frequency band Ω . We show that coherence at Ω is equivalent to the correlation between the two filtered time series. The second goal of this paper is to develop a novel adaptive statistical procedure for estimating coherence when the time series are non-stationary, that is, the nature of linear dependence between time series may evolve with time. The proposed method for estimating local coherence automatically selects, via repeated tests of homogeneity (in time) of coherence, the optimal width of the time window on which one computes the estimated local coherence. This approach is point-wise adaptive in the sense that the width of the optimal interval is allowed to change across time. Under the locally stationary process framework, we develop a central limit theorem on the Fisher-z transform of our time-localized band coherence. We apply our method to a pair of highly dynamic brain waves signals whose coherence is shown evolve during an epileptic seizure.

DP 0619 – Pharmacokinetic parameters estimation using adaptive Bayesian P-splines models

Jullion, A., Lambert, Ph., Beck, B. and F. Vandenhende

In preclinical experiments, pharmacokinetic (PK) studies are designed to analyse the evolution of drug concentration in plasma over time i.e. the PK profile. Some PK parameters are estimated in order to summarize the complete drug’s kinetic profile : area under the curve (AUC), maximal concentration (C_{max}), time at which the maximal concentration occurs (t_{max}) and half life time ($t_{1/2}$).

Several methods have been proposed to estimate these PK parameters. A first method relies on interpolating between observed concentrations. The interpolation method is often chosen linear. Another method relies on compartmental modelling. In this case, non linear methods are used to estimate parameters of a chosen compartmental model. Two problems can arise with this method. The first one is the difficulty to choose the suitable compartmental model given the small number of data available in preclinical experiment. Secondly, non linear methods may fail to converge with sparse data. Hence, there are some limitations in practice that prevent its use in preclinical PK studies.

In this paper, we propose a Bayesian nonparametric model based on P-splines. Simulations show that the proposed method provides better PK parameters estimations than the interpolation method, both in terms of bias and precision. We extend the basic model to a hierarchical one that treats the case where we have concentrations from different subjects. We are then able to get individual PK parameter estimations. Finally, with Bayesian methods, we can get easily some uncertainty measures by obtaining credibility sets for each PK parameter.

DP 0620 – Nonparametric regression with dependent data

El Ghouch, A. and I. Van Keilegom

Let (X_i, Y_i) ($i = 1, \dots, n$) be n replications of a random vector (X, Y) , where Y is supposed to be subject to random right censoring. The data (X_i, Y_i) are assumed to come from a stationary α -mixing process. We consider the problem of estimating the function $m(x) = \mathbb{E}(\phi(Y)|X = x)$, for some known transformation ϕ . Particular choices of ϕ lead to the conditional moment function of Y given X , or the conditional distribution of Y given X .

This problem is approached in the following way : first, we introduce a transformed variable Y_i^* , that is not subject to censoring and satisfies the relation $\mathbb{E}(\phi(Y_i)|X_i = x) = \mathbb{E}(Y_i^*|X_i = x)$ and then we estimate $m(x)$ by applying local linear regression techniques to the pseudodata (X_i, \hat{Y}_i^*) , where \hat{Y}_i^* is a certain estimator of Y_i^* . As a by-product, we obtain a general result on the uniform rate of convergence of kernel type estimators of functionals of an unknown distribution function, under strong mixing assumptions. This result is of independent interest, and can be applied in a wide variety of contexts.

DP 0621 – Nonparametric density estimation for positive time series

Bouezmarni, T. and J.V.K. Rombouts

The Gaussian kernel density estimator is known to have substantial problems for bounded random variables with high density at the boundaries. For i.i.d. data several solutions have been put forward to solve this boundary problem. In this paper we propose the gamma kernel estimator as density estimator for positive data from a stationary α -mixing process. We derive the mean integrated squared error, almost sure convergence and asymptotic normality. In a Monte Carlo study, where we generate data from an autoregressive conditional duration model and a stochastic volatility model, we find that the gamma kernel outperforms the local linear density estimator. An application to data from financial transaction durations, realized volatility and electricity price data is provided.

DP 0622 – Extremal indices, geometric ergodicity of Markov chains, and MCMC

Roberts, G.O., Rosenthal, J.S., Segers, J. and B. Sousa

We investigate the connections between extremal indices on the one hand and stability of Markov chains on the other hand. Both theories relate to the tail behaviour of stochastic processes, and we find a close link between the extremal index and geometric ergodicity. Our results are illustrated throughout with examples from simple MCMC chains.

DP 0623 – A multiscale approach for statistical characterization of functional images

Antoniadis, A., Bigot, J. and R. von Sachs

In this paper we use an approach of spatial multiscales for an improved characterization of functional pixel intensities of images. Examples are numerous such as temporal dependence of brain response intensities measured by fMRI or frequency dependence of NMR spectra measured at each pixel. The overall goal is to improve the misclassification rate in clustering (unsupervised learning) of the functional image content into a finite but unknown number of classes. Hereby we adopt a non-parametric point of view to reduce the functional dimensionality of the observed pixel intensities, modelled to be of a very general functional form, by a combination of "aggregation" and truncation techniques. Clustering is applied via an EM-algorithm for estimating a Gaussian mixture model in the domain of the discrete wavelet transform of the pixel intensity curves. We show improvements of our multiscale method, based on complexity-penalised likelihood estimation for Recursive Dyadic Partitioning of the image, over existing monoscale approaches, by simulated and real data examples, and we give some theoretical treatment of the resulting misclassification rate in the simplified set-up of the "horizon" model of two classes.

DP 0624 – Locally stationary factor models : identification and nonparametric estimation

Motta, G., Hafner, C. and R. von Sachs

In this paper we propose a new approximate factor model for large cross-section and time dimensions. Factor loadings are assumed to be smooth functions of time, which allows to consider the model as locally stationary while permitting empirically observed time-varying second moments. Factor loadings are estimated by the eigenvectors of a nonparametrically estimated covariance matrix. As is well-known in the stationary case, this principal components estimator is consistent in approximate factor models if the eigenvalues of the noise covariance matrix are bounded. To show that this carries over to our locally stationary factor model is the main objective of our paper. Under simultaneous asymptotics (cross-section and time dimension go to infinity simultaneously), we give conditions for consistency of our estimators. A simulation study illustrates the performance of these estimators.

DP 0625 – On joint completeness : sampling and Bayesian versions, and their applications
San Martín, E. and M. Mouchart

Recently, Cramer, Kamps and Schenk (2002) established conditions under which a family of joint distributions of two independent statistics is complete, and related their result with a previous one of Landers and Rogge (1976). We first propose, within a sampling theory framework, a modification of Cramer, Kamps and Schenk's (2002) generalization of Landers and Rogge's (1976) theorem, paying a particular attention to the concept of completeness of a function of the parameters. Next, after reminding the concept of completeness in a Bayesian framework, we discuss its robustness with respect to the prior specification and its relationship with sampling completeness. It is then shown that Landers and Rogge's (1976) theorem can be extended, and in a sense generalized, to a Bayesian framework. A Bayesian version of Cramer, Kamps and Schenk (2002)'s theorem is also provided. These results are exemplified in both a normal and a discrete Bayesian experiment.

6.2. Abstracts of the consulting reports

CR 0601 – Evaluation statistique de la méthodologie de distribution de médailles dans des concours internationaux de vin : application au Concours Mondial de Vin de Bruxelles

Ritter, C., Baclin, A. and B. Govaerts

Le Concours Mondial de Bruxelles (CMB) est un grand concours international de vins organisée jurys parallèles par la société Vinopres (Bruxelles). Afin de valider une méthode d'alignement de jurys dans la distribution des médailles et explorer des alternatives, Vinopres a collaboré durant deux ans avec une équipe de l'Institut Statistique de l'Université catholique de Louvain. Ce rapport regroupe l'ensemble des documents qui ont été produits durant cette collaboration. Certains se rapportent directement au concours de Bruxelles, d'autres sont d'ordre méthodologique général. Ils sont disponibles sur demande auprès de govaerts@stat.ucl.ac.be.

Liste des documents disponibles :

Rapport final : Ce document contient le rapport final (en résumé) et les éléments principaux des analyses sur l'imprécision des cotes moyennes du concours.

Rapport Février 2005: Ce document contient une analyse faite en Février 2005 sur les données du CMB2004. On y trouve des graphiques temporels par dégustateur qui montrent leur écart par rapport au consensus du jury.

Rapport Mai 2005: Ce rapport examine les résultats du CMB2005 et donne des graphiques temporels par dégustateur (semblables au rapport du mois de février).

Expertise Octobre 2004: Ce document contient une expertise réalisée pour l'accréditation de la formule de normalisation de Vinopres. Proposition de formule d'attribution de médailles. Ce document contient les résultats et développements intermédiaires d'une étude sur l'ajustement des cotes sur base de moyennes sur vins témoins. Cet ajustement dépend du fait de disposer d'au moins deux vins témoins bien différents.

Fiches Dégustateurs: Ce document propose une méthode pour résumer en une fiche les performances de dégustation d'un juge. Cette fiche situe le dégustateur en moyenne et dispersion par rapport aux autres membres de son jury et ce jury par rapport aux autres, ceci sur base de vins témoins et des vins dégustés pour le concours. Etude de la performance des juges en analyse sensorielle, mémoire méthodologique réalisé par Angélique Baclin avec application aux données du CMB.

CR 0602 – Calcul de l'incertitude sur l'estimation du nombre de trous moyens par unité dans des membranes de polymère irradiées par des faisceaux d'ions lourds

B. Govaerts

Dans un but de contrôle de qualité, ce travail a pour but de proposer une méthode pour calculer l'incertitude sur l'estimation du nombre de trous moyens par unité dans des membranes de polymère irradiées par des faisceaux d'ions lourds.

L'intervalle proposé se base sur l'hypothèse d'une distribution de Poisson pour le nombre de trous observés par unité de surface. Il est appliqué à une série de données de fabrication et permet de montrer comment tirer des conclusions concernant la capacité du procédé de fabrication à répondre aux spécifications du client.

CR 0603 – SPSS : doit-on, veut-on le garder à l'UCL ?

Govaerts, B., Baclin, A. and A. Guillet

SPSS Belgique a annoncé pour 2006-2007 une augmentation importante du prix de la licence de site. Cette annonce a amené l'UCL à réaliser une réflexion de fond sur le choix des logiciels mis à disposition des utilisateurs en licence de site. Cette réflexion a été coordonnée par l'Institut de statistique et réalisée en collaboration un groupe de représentants des utilisateurs.

La démarche poursuivie a été la suivante:

1. Recherche d'une liste de logiciels candidats pour remplacer SPSS.
2. Définition des utilisateurs visés.
3. Recherche des besoins des utilisateurs.
4. Tests puis représentations de logiciels aux représentants des utilisateurs.
5. Tests supplémentaires des logiciels par les utilisateurs et réflexions.
6. Négociation avec les firmes de logiciels.
7. Recommandation et mise en oeuvre.

Le choix final a consisté à garder, pour les utilisateurs d'outils de base de statistique, une version très réduite de SPSS et à recommander, pour les utilisateurs d'outils plus avancés ou spécialisés le logiciel SAS (avec son interface SAS Enterprise Guide) ainsi que des logiciels spécifiques (STATA, SPAD, R, JMP, PASSÉ)

CR 0604 – Etude de similitude entre deux formulations chimiques. Proposition et comparaison d'indices statistiques de mesure de similitude

B. Govaerts

Le problème posé consiste à visualiser graphiquement et quantifier les concordances éventuelles existant entre produits commerciaux sur base des formulations (concentrations exprimées en pourcentage de chaque composé chimique) fournies par les fabricants.

Trois approches sont proposées. La première consiste à représenter simplement les deux produits sur un graphe X-Y. Les deux autres sont des indices destinés à quantifier la similitude. Le premier est tout simplement le coefficient de corrélation de Pearson et le second est basé sur le Phi de Cramer. Les avantages et inconvénients de chacun sont discutés et illustrés sur un exemple.