



**Institut de Statistique**

**Annual Report**

**Year 2007**

**UCL**  
**Université catholique de Louvain**



# Contents

---

1. Foreword .....	3
2. Staff .....	5
3. Research Activities .....	11
3.1. Presentation .....	11
3.2. Research contracts .....	12
3.2.1. Research projects under contracts and cooperation projects .....	12
3.2.2. Applied research contracts .....	15
3.3. Reprints .....	16
3.4. Discussion papers .....	18
3.5. Books .....	20
3.6. Visitors .....	21
3.7. Statistics Seminars .....	22
3.8. Organization of scientific meetings.....	24
3.9. Academic visits .....	25
3.10. Participation to conferences and statistics meetings.....	27
4. Teaching Activities and Educational Activities .....	30
4.1. Bachelor level .....	30
4.2. Minor in Statistics .....	30
4.3. New Master programs in Statistics .....	30
4.4. Certificate .....	30
4.5. Doctoral School .....	31
4.5.1. Presentation .....	31
4.5.2. Doctoral dissertations and PhD in progress .....	32
4.5.3. Doctoral seminars .....	34
4.5.4. Short courses .....	35
5. Support for the University and Consulting Activities .....	36
5.1. Presentation .....	36
5.2. SMCS activities .....	36
5.3. Applied Statistics Workshops .....	37
6. Appendix .....	39
6.1. Abstracts of the discussion papers .....	39
6.2. Abstracts of the consulting reports .....	45

# 1. Foreword

---

Since its founding by Professor Léopold Simar in the early 1990s, the UCL Institute of Statistics has undoubtedly become the largest and most comprehensive statistics group 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 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 2007 the teaching faculty comprised a full time staff of 9 professors and 20 teaching and research assistants. Dr Catherine Legrand was appointed as full-time professor in the Institute of Statistics after an international search. She is expected to further increase the reputation and activity of the Institute in the area of biostatistics. Several part-time and invited professors broadened the scope of expertise of the Institute in specific topics, like data mining or surveys. The Institute also accommodated post-doc researchers and short-term as well as long-term foreign visiting scholars. The administrative, library and technical staff consisted 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 schools, take of ancillary statistics courses taught by members of the Institute who always place strong emphasis on quality teaching. Besides these courses taught in the different schools of the university, the Institute of Statistics continued strengthening its master and doctoral programmes. An 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. Continuing professional development is also offered through various short courses and through a university certificate run with the collaboration of IUFC.

Members of the Institute have a wide variety of research interests, ranging from theoretical are developed topics and over a broad spectrum of methodological topics. The main research areas at the Institute develop along the three following main directions: (i) Mathematical Statistics (semi- and non-parametric statistics, Bayesian statistics, multivariate analysis, regression, mixture models, survival analysis, time series, ...), (ii) Biostatistics, and (iii) Quantitative Analysis of Insurance and Finance Risks.

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, including scientific publications, seminars, and research projects, are described in a separate report.

Research activity continued to flourish in 2007. During 2007, members of the Institute published 25 papers in refereed journals. Research to appear soon, or under evaluation, was reported in 28 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 nine research teams, of which five are Belgian and four 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 recognition.

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 "Young Researchers Day - YRD" organized twice a year, PhD students in statistics present their recent research work. These doctoral seminars constitute an extra stimulant for PhD 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. It offers various activities, such as short courses, workshops and an annual colloquium, some of them held in UCL under the auspices of the Institute, others in partner universities.

The Institute of Statistics offers a consulting service to the UCL community. Researchers of the other 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. These two services have been merged in 2007 to form the SMCS task force (acronym for the French "Support en Méthodologie et Calcul Statistique"). This initiative is meant to further increase the quality of the assistance to the UCL community. A detailed account of all the SMCS activities can be found in a separate specific report.

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, September 15, 2008

Michel Denuit  
Chairman

Sébastien Van Bellegem  
Academic Secretary

## 2. Staff

---

### Permanent Academic Members



#### Michel Denuit

Professor, Université catholique de Louvain (Institute of Statistics).

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 (Institute of Statistics), 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.

Member of the College of the O.P.P (Observatoire du Principe de Précaution).



#### Bernadette Govaerts

Professor, Université catholique de Louvain (Institute of Statistics).

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.

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

President of the executive committee of the SMCS and of the CORE-STAT-ACTU computing commission.

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



### **Christian Hafner**

Professor, Université catholique de Louvain (Institute of Statistics).  
Ph. D. Humboldt-Universität Berlin, Germany, 1996.  
Non and semiparametric statistics, time series, volatility models, financial econometrics.  
Associate editor of *Computational Statistics* (since 2004), *Eurasian Review of Econometrics* (since 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).

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, pedagogic adviser for the Master program in statistics, member of the SMCS.

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



### **Catherine Legrand**

Professor, Université catholique de Louvain (Institute of Statistics).  
Ph. D. Universiteit Hasselt, 2006.  
Biostatistics, Design and analysis of clinical trials, Survival data and frailty models, Competing risks, Repeated events.  
Board member of the biostatistic section of the Belgian Statistical Society (since 2008).  
Catherine Legrand is pedagogic adviser for the Master in Biostatistics.



### **Christian Ritter**

Part-time Professor, Université catholique de Louvain (Institute of Statistics).  
Ph. D. University of Wisconsin, Madison, USA, 1992.  
Industrial statistics, statistical consulting, statistics in spreadsheets.  
Creation of Ritter and Danielson Consulting sprl (2007) to provide R&D assistance.

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



### Johan Segers

Professor, Université catholique de Louvain (Institute of Statistics).  
Ph. D. Leuven, 2001.

Extreme value theory, Dependence modelling via copulas, Edgeworth expansions. Copulas, curve estimation under shape constraints.

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). Coordinator of the Erasmus program at the Institute of Statistics.

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



### Léopold Simar

Professor, Université catholique de Louvain (Institute of Statistics).

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 Institute of Statistics, UCL (1992-2004).

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

**"Professore di Chiara Fama"**, 2006 and 2007, (Italian Ministry of Research, Scuola Superiore Santa Anna, Pisa & University of Pisa).

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



### Sébastien Van Belleghem

Professor, Université catholique de Louvain (Institute of Statistics).

Ph. D. Université catholique de Louvain, 2003.

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

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

Academic secretary of the Institute of Statistics (2006-2009), President of the Jury for the Master in Statistics.

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



### **Ingrid Van Keilegom**

Professor, Université catholique de Louvain (Institute of Statistics).

Ph. D. Limburgs 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 *Scandinavian Journal of Statistics* (since 2004), *Annals of the Institute of Statistical Mathematics* (2006), *The Annals of Statistics* (since 2007), *Statistics and Probability Letters* (since 2007).

Elected member of the ISI (International Statistical Institute), Ingrid Van Keilegom follows the developments on various research funding programmes for the Institute of Statistics, 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 (Institute of Statistics).

PhD 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.

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  
Guy Lories, Faculty of Psychology  
Annie Robert, Faculty of Medicine  
Michel Verleysen, Faculty of Applied Mathematics

### **Invited Academic Members**

Yves Berger  
Céline Bugli  
Libei Chen  
Isabelle De Macq  
Cédric Heuchenne

### **Emeritus Professors**

Michel Mouchart  
José Paris  
Jean-Marie Rolin

### **Researchers and Doctoral Students**

Carlos Almeida  
Hilmar Böhm  
Jean-Philippe Boucher  
Cindy Courtois  
Alexandra Daskovska  
Anouar El Ghouch  
Rachid El Medhi (since September 2007)  
Jean-Marc Freyermuth (since October 2007)  
Gery Geenens  
Julien Hunt  
Astrid Jullion  
Thomas Laloux  
Céline Le Bailly de Tillegem  
Thomas Meinguet (since September 2007)  
Giovanni Motta  
Sandra Pitrebois  
Olga Reznikova  
Réjane Rousseau  
Bianca Teodorescu  
Catherine Timmermans (since September 2007)

### **Associate Researchers**

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

**Visiting researcher**

Claudia Curi, Università di Roma, Roma, Italy (05/03/07-15/09/07) (18/10/07-12/11/07)

**Doctors Honoris Causa**

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

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

**Computer Scientists SGSI (Service général du système d'information)**

Laurent Buset

Jean-Luc Marrison

**SMCS Staff (Support en Méthodologie et Calcul Statistique)**

Alain Guillet

Jean-Marie Zélis

**Head of Unit (Administration)**

Christine Denayer

**Administrative Staff**

Anne-Marie Bellemans

Marguerite-Marie Hanon

Sophie Malali

Monique Tanga (since October 2007)

## 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 particularly 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 of Statistics that are financed by outside agencies in the form of grants and contracts.

#### **Statistical analysis of association and dependence in complex data** (2007-2011)

Financing: Interuniversity Attraction Pole Programme, Belgian Science Policy, Brussels, Belgium

Coordinators: L. Simar and I. Van Keilegom

STAT Researchers: H. Böhm and A. Daskovska

Partners Institutions: Katholieke Universiteit Leuven, Belgium  
Universiteit Hasselt, Belgium  
Universiteit Gent, Belgium  
Université Joseph Fourier, Grenoble, France  
Universiteit Utrecht, Netherlands  
Universidad de Santiago de Compostela, Spain  
London School of Hygiene and Tropical Medicine, UK

One key aim of statistics is to analyse in an appropriate way the dependence and association present in a dataset. The data collected to analyse these dependence structures, are often of a complex nature and also the research questions are of an ever increasing complexity. This requires the construction of new models, or the adaptation of existing models which is a challenging task. The development of new methods and intensive interaction between experts will also be required to cope with these complex data. The global objective of the network is to develop new models and methodological tools to infer and to analyse these complex data structures.

#### **Econometric modelling of multivariate financial time series** (2007-2012)

Financing : Action de Recherche Concertée (ARC)

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

STAT Researchers: A. Daskovska, J.M. Freyermuth and O. Reznikova

This interdisciplinary research project deals with modelling, estimation and prediction of the dynamics and the temporal dependence in the mean and the variance-covariance structure of multivariate time series data arising in economical and financial applications. Particular emphasis is put on questions such as dimension reduction (factor approach, modelling of co-movements), non-stationary behaviour over time, modelling of structural breaks (regime-switching), volatilities with and without jump behaviour, etc... .

These questions are addressed by a number of econometricians and statisticians using and comparing a series of modern approaches in parametric, semi-parametric and non-parametric statistics. Applications to real data will help to assess the quality of the proposed models and estimation procedures.

**Policies for research and innovation in the move towards the European research area** (2004-2008)  
(PRIME Network of Excellency)

Financing : European Union

Grant holder : L. Simar, Associate partner of the Scuola Superiore Santa Anna, Pisa, Italy

STAT Researcher : L. Simar

Using advanced quantitative methods for evaluating the performance of Public Sector Research (AQUAMETH). 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

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 in an adaptive way.

**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 (SO<sub>2</sub> predictions in a power plant, concentration of heavy metals in mosses in Galicia, ...), Industry (control of wastewater treatment by anaerobic processes) and Finance (semi parametric inference in portofolio design).

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

Financing : Fonds spéciaux de recherche (FSR)

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

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 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?

## **Adaptative regularization in nonparametric instrumental regression (2007-2009)**

Financing : Fonds spéciaux de recherche (FSR)

Grant holder : S. Van Bellegem

We focus on the nonparametric estimation of a regression function in endogenous models. The regression function is identified through a set of instrumental variables. Stability problems occur in the estimation process, leading to an ill-posed inverse problem. The general objective is to develop adaptive methods of inversion (of regularization), based on the theory of adaptive, nonparametric, local inference.

### 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 bi-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 on continued education (see point 6) at the University and in companies.

#### **Statistical analysis of PET scan data (2003-2007)**

Financing : Eli Lilly (07/2003 - 06/2004), FSR (07/ 2004 - 06/ 2005), Eli Lilly (07/ 2005 - 06/ 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.

#### **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

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 at developing the necessary tools and implementing them in an adapted computing tool.

#### **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

### 3.3 Reprints

- RP0701 Charpentier, C. and J. Segers  
Lower tail dependence for Archimedean copulas: characterizations and pitfalls.  
*Insurance: Mathematics and Economics*, 40, 525-532, 2007
- RP0702 Haeusler, E. and J. Segers  
Assessing confidence intervals for the tail index by Edgeworth expansions for the Hill estimator.  
*Bernoulli*, 13, 175-194, 2007
- RP0703 Wang, L. and I. Van Keilegom  
Nonparametric test for the form of parametric regression with time series errors.  
*Statistica Sinica*, 17, 369-386, 2007
- RP0704 Boulanger, B., Dewé, W., Gilbert, A., Govaerts, B. and M. Maumy-Bertrand  
Risk management for analytical methods based on the total error concept: conciliating the objectives of the pre-study and in-study validation phases.  
*Chemometrics and Intelligent Laboratory System*, 86, 198-207, 2007
- RP0705 Almeida C., and M. Mouchart  
Bayesian encompassing specification test under not completely known partial observability  
*Bayesian Analysis*, 2, 303-318, 2007
- RP0706 Van Keilegom, I. and R. J. Carroll  
Backfitting versus profiling in general criterion functions.  
*Statistica Sinica*, 17, 797-816, 2007
- RP0707 Heuchenne, C. and I. Van Keilegom  
Polynomial regression with censored data based on preliminary nonparametric estimation.  
*Annals of the Institute of Statistical Mathematics*, 59, 273-297, 2007
- RP0708 Pardo-Fernández, J. C., Van Keilegom, I. and W. González-Manteiga  
Goodness-of-fit tests for parametric models in censored regression.  
*The Canadian Journal of Statistics*, 35, 249-264, 2007
- RP0709 François, N., Govaerts, B. and C. Guyot-Declerck  
Inferential non-centred principal curve analysis of time-intensity curves in sensory analysis: the methodology and its application to beer astringency evaluation.  
*Journal of Chemometrics* 21, 187-197, 2007
- RP0710 Cosma, A., Scaillet, O. and R. von Sachs  
Multivariate wavelet-based shape preserving estimation for dependent observations.  
*Bernoulli*, 13, 2, 301-329, 2007
- RP0711 Bonaccorsi, A., Daraio, C., and L. Simar  
Efficiency and productivity in European Universities - Exploring trade-offs in the strategic profile.  
Bonaccorsi, A. and C. Daraio, eds, *Universities as strategic units. Productivity and Efficiency Patterns in the European University System*, Edward Elgar PRIME collection, 2007
- RP0712 Bonaccorsi, A., Daraio, C., Rätty, T. and L. Simar  
Efficiency and university size: discipline-wise evidence from the European universities.  
*Hyvintöintipalvelujen Tuottavuus: Tuloksia opintien varrelta*, VATT Publications, Helsinki, Finland, 2007
- RP0713 Simar, L. and P. Wilson  
Estimation and inference in two-stage, semi-parametric models of production processes.  
*Journal of Econometrics*, 136, 31-64, 2007

- RP0714 Park, B., Sickles, R. and L. Simar  
Semiparametric efficient estimation in dynamic panel data models.  
*Journal of Econometrics*, 136, 281-301, 2007
- RP0715 Kumbhakar, S. C., Park, B. U., Simar, L. and E. G. Tsionas  
Nonparametric stochastic frontiers: a local likelihood approach.  
*Journal of Econometrics*, 137, 1-27, 2007
- RP0716 Daouia, A. and L. Simar  
Nonparametric efficiency analysis: a multivariate conditional quantile approach.  
*Journal of Econometrics*, 140, 375-400, 2007
- RP0717 Daraio, C. and L. Simar  
Conditional nonparametric frontier models for convex and non convex technologies: a unifying approach.  
*Journal of Productivity Analysis*, 28, 13-32, 2007
- RP0718 Simar, L.  
How to improve the performances of DEA/FDH estimators in the presence of noise.  
*Journal of Productivity Analysis*, 28, 183-201, 2007
- RP0719 Simar, L. and V. Zelenyuk  
Statistical inference for aggregates of Farrell-type efficiencies.  
*Journal of Applied Econometrics*, 22, 1367-1394, 2007
- RP0720 Hall, P. and I. Van Keilegom  
Two-sample tests in functional data analysis starting from discrete data.  
*Statistica Sinica*, 17, 1511-1532, 2007
- RP0721 Heuchenne, C. and I. Van Keilegom  
Nonlinear regression with censored data.  
*Technometrics*, 49, 34-44, 2007
- RP0722 Pardo-Fernández, J. C., Van Keilegom, I. and W. González-Manteiga  
Testing for the equality of k regression curves.  
*Statistica Sinica*, 17, 1115-1137, 2007
- RP0723 Dette, H., Neumeyer, N. and I. Van Keilegom  
A new test for the parametric form of the variance function in nonparametric regression.  
*Journal of the Royal Statistical Society. Series B*, 69, 903-917, 2007
- RP0724 Heuchenne, C. and I. Van Keilegom  
Location estimation in nonparametric regression with censored data.  
*Journal of Multivariate Analysis*, 98, 1558-1582, 2007
- RP0725 Mouchart, M. and M. Vandresse  
Bargaining powers and market segmentation in freight transport.  
*Journal of Applied Econometrics*, 22, 1295-1313, 2007

### 3.4. Discussion papers

The abstracts of the discussion papers can be found in the Appendix.

- 0701 Krier, C., Mouchart, M. and A. Oulhaj  
*Neural modelling of ranking data with an application to stated preference data*
- 0702 San Martin, E. and J.-M. Rolin  
*Identifiability and estimability of parametric Rasch-type models*
- 0703 Segers, J.  
*Multivariate regular variation of heavy-tailed Markov chains*
- 0704 Robert, C. Y. and J. Segers  
*Tails of random sums of a heavy-tailed number of light-tailed terms*
- 0705 Rolin, J.-M.  
*Bayesian consistency of density estimators*
- 0706 Böhm, H. and R. von Sachs  
*Shrinkage estimation in the frequency domain of multivariate time series*
- 0707 Almeida, C. and M. Mouchart  
*Testing the normality of the latent variables in the polychoric correlation model*
- 0708 Bouezmarni, T., El Ghouch, A. and M. Mesfioui  
*Gamma kernels estimator for density and hazard rate of right censored data*
- 0709 Mouchart, M., Russo, F. and G. Wunsch  
*Causality, structural modelling and exogeneity*
- 0710 Lebrun, P., Govaerts, B., Debrus, B., Ceccato, A., Caliaro, G., Hubert, P. and B. Boulanger  
*Development of a new predictive modelling technique to find with confidence equivalence zone and design space of chromatographic analytical methods*
- 0711 Rousseau, R., Govaerts, B., Verleysen, M. and B. Boulanger  
*Comparison of some chemometric tools for metabonomics biomarker identification*
- 0712 El Ghouch, A. and I. Van Keilegom  
*Local linear quantile regression with dependent censored data*
- 0713 Van Keilegom, I., Sánchez-Sellero, C. and W. González-Manteiga  
*Empirical likelihood based testing for regression*
- 0714 Bouezmarni, T. and J. Rombouts  
*Semiparametric multivariate density estimation for positive data using copulas*
- 0715 Jullion, A., Lambert, Ph. and F. Vandenhende  
*A non-parametric bayesian method to smooth PET time-activity-curves*
- 0716 Bouezmarni, T. and J. Rombouts  
*Nonparametric density estimation for multivariate bounded data*
- 0717 Basrak, B. and J. Segers  
*Regularly varying multivariate time series*
- 0718 Genest, C. and J. Segers  
*Rank-based inference for bivariate extreme-value copulas*
- 0719 Le Bailly de Tillegem, C. and B. Govaerts  
*A review of quantitative structure-activity relationship (QSAR) models*

- 0720 Dette, H., Pardo-Fernández, J. C. and I. Van Keilegom  
*Goodness-of-fit tests for multiplicative models with dependent data*
- 0721 Florens, J.-P., Johannes, J. and S. Van Belleghem  
*Identification and estimation by penalization in nonparametric instrumental regression*
- 0722 Einmalh, J. H. J., Krajina, A. and J. Segers  
*A method of moments estimator of tail dependence*
- 0723 Bouezmarni, T., Mesfioui, M. and A. Tajar  
*On concordance measures for discrete data and dependence properties of Poisson model*
- 0724 Braekers, R. and I. Van Keilegom  
*Flexible modeling based on copulas in nonparametric regression*
- 0725 Bonaccorsi, A. , Curi, C., Daraio, C., Mancuso, P. and L. Simar  
*Analysing total factor productivity through a bootstrap-based nonparametric approach: the case of italian regions*

### 3.5 Books

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

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

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 R. Kaas, (2005)

"Actuarial Theory for Dependent Risk : Measures, Orders and Models" - Wiley, New York

Daraio, C. and L. Simar (2007)

"Advanced Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications" Springer, New-York

Deprins, D. (2007)

"Le Pari de Pascal, une allégorie du Principe de Précaution", sur site de l'Observatoire du Principe de Précaution

Deprins, D., (2007)

"Quel homme se construit sous le paradigme de Précaution?", sur site de l'Observatoire du Principe de Précaution"

Härdle, W. and L. Simar (2007)

"Applied Multivariate Statistical Analysis" - Second Edition, 458 pp., Springer-Verlag, Berlin

## 3.6. Visitors

### Doctoral Visitors

GAJEWSKI Claudia, Geophysical Institute, Postdam, Germany

### Short Terms Visitors

ANDERSON Amy, Western Washington University, Washington, USA  
BADIN Luiza, ASE Bucarest, Romania  
BIGOT Jérémie, Université Toulouse III, Toulouse, France  
CADARSO SUAREZ Carmen, Universidad de Santiago de Compostela, Spain  
CAVAIGNAC Laurent, Université de Perpignan, Perpignan, France  
CHRISTENSEN Antje, Technical University of Denmark, Denmark  
CURI Claudia, Industrial Engineering from University of Roma, Italy  
DAOUIA Abdelaati, Université de Toulouse, France  
DARAIIO Cinzia, Università Pisa, Italy  
DAVISON Anthony, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland  
EILERS Paul, Leiden University, The Netherlands  
EVANS Stephen, LSHTM, London, UK  
FRIED Roland, Technische Universität Dortmund, Germany  
GATHER Ursula, Technische Universität Dortmund, Germany  
GENEST Christian, Université Laval, Québec, Canada  
GONZALEZ-MANTEIGA Wenceslao, Universidad de Santiago de Compostela, Spain  
GUIZZARDI Andrea, University of Bologna, Italy  
HENDERSON Daniel, State University New York, USA  
HJORT Nils, University of Oslo, Norway  
KNEIP Alois, University of Bonn, Germany  
KREISS Jens-Peter, Institut für Mathematische Stochastik, TU Braunschweig, Germany  
LINTON Oliver, London School of Economics, UK  
MA Yanyuan, Université de Neuchâtel, Switzerland  
MCKEAGUE Ian, Columbia University, New York, USA  
MEYER Mary, University of Georgia, USA  
MIKOSCH Thomas, University of Copenhagen, Denmark  
MIZON Graham, University of Southampton, Southampton, UK  
OJEDA Cabrera Jorge, Universidad de Zaragoza, Spain  
OPSOMER Jean, Iowa State University, Ames, USA  
PARDO-FERNANDEZ Juan Carlos, Universidade de Vigo, Spain  
PARK Byeong, Seoul National University, South Korea  
ROUEFF Christian, ENST, TSI, Paris, France  
SAN MARTIN Ernesto, PUC Santiago, Chile  
SCHWARZ Maik, University of Heidelberg, Germany  
SEOK-OH Jeong, Hankuk University of Foreign Studies, South Korea  
SNIJDERS Tom, University of Groningen, The Netherlands  
TADJUIDGE Joseph, University of Kaiserslautern, Germany  
WEGKAMP Marten, Florida State University, Tallahassee, USA  
WEIHS Clauss, University of Dortmund, Germany  
WILSON Paul, Clemson University, South Carolina, USA

### 3.7. Statistics seminars

February 5, 2007, joint seminar with IAP

Juan Carlos Pardo-Fernandez, University of Vigo, Vigo, Spain  
"Tests in nonparametric regression based on the error distribution"

February 5, 2007, joint seminar with MAPA and IAP

Wenceslao Gonzalez-Manteiga, University of Santiago de Compostela, Spain  
"Comparing spatial dependence structures"

February 15, 2007

Christian Genest, Université Laval, Québec, Montréal  
"Abraham De Moivre : Génie en exil"

February 23, 2007

Jérémie Bigot, Université Toulouse III, France  
"Homeomorphic smoothing splines : a new tool for monotonizing a non-constrained estimator in nonparametric regression"

Geurt Jongbloed, Vrije Universiteit Amsterdam, The Netherlands  
"Density estimation in the Current Status model"

March 23, 2007

Joseph Tadjuidje Kamgaing, University of Kaiserslautern, Germany  
"Geometric Ergodicity for a Class of Autoregressive Driven Processes with Exogenous Components"

Johan Segers, UCL, Belgium  
"Multivariate regular variation of heavy-tailed Markov chains"

April 20, 2007

Paul Eilers, Utrecht University, The Netherlands  
"Statistical Classification for Reliable High-volume Genetic Measurements"

Kristel Van Steen, Universiteit Gent, Belgium  
"Perspectives on genetic association screening"

May 11, 2007

Jean Opsomer, Iowa State University, USA  
"Endogenous Post-stratification in Surveys"

Céline Vial, ENSAI, Rennes, France  
"Local model selection in inverse problem"

September 28, 2007

François Roueff, ENST, TSI, Paris, France  
"Semi-parametric estimation of the long-memory parameter"

Gerda Claeskens, KULeuven, Belgium  
"Prediction Focused Model Selection for Autoregressive Models"

October 12, 2007

Andrea Krajina, Tilburg University, The Netherlands  
"A method of Moments Estimator of Tail Dependence"

Thomas Mikosch, University of Copenhagen, Denmark  
*"Regularly varying functions"*

October 26, 2007

Michel van de Velden, Erasmus University Rotterdam, The Netherlands  
*"Oblique rotation in correspondence analysis : a step forward in the search of the simplest interpretation"*

Patrick Groenen, Erasmus University Rotterdam, The Netherlands  
*"Minimization for support Vector Machines by Iterative Majorization"*

November 16, 2007, joint seminar with CORE

Peter Boswijk,, University of Amsterdam, The Netherlands  
*"Testing for Cointegration with Nonstationary Volatility"*

Hashem Pesaran and Alexander Chudik, University of Cambridge, UK  
*"Infinite Dimensional VARs"*

November 29, 2007, joint seminar with IAP

Anthony Davison, EPFL Lausanne, Switzerland  
*"Likelihood inference for a problem in particle physics"*

Ernesto San Martin, PUC Santiago, Chile  
*"On the statistical Meaning of the Parameters of Skew-Normal Experiments"*

November 30, 2007

Roland Fried, Universität Dortmund, Germany  
*"Robust and detail-preserving signal extraction from time series"*

December 14, 2007

Maria-Pia Victoria-Feser, Université de Genève, Switzerland  
*"Robust Estimation in SEM"*

Christophe Croux, KULeuven, Belgium  
*"Robust forecasting with Holt-Winters smoothing"*

## 3.8 Organization of Scientific Meetings

### Organization of a summer school on "Advanced Nonparametric Statistics" by the FNRS Graduate School in Statistics and Actuarial Sciences and the IAP Network in Statistics,

Institute of Statistics, Louvain-la-Neuve, Belgium, May 21-24, 2007

#### Speakers and program:

**Oliver Linton**, "Nonparametrics in Financial Econometrics: Discrete time volatility and Continuous time processes", London School of Economics, London, UK

**Jean Opsomer**, "Nonparametric regression methods in surveys", Iowa State University, Ames, USA

**Marten Wegkamp**, "Aggregation and sparsity oracle inequalities in nonparametric classification and regression", Florida State University, Tallahassee, USA

### 3.9 Academic visits

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

#### January 2007

**Sébastien Van Bellegem**, "Inference in semi- and non-parametric regression with instrumental variables", Seminar of the Weierstrass Institute, Weierstrass Institute for Applied Analysis and Stochastics, Berlin, Germany

**Sébastien Van Bellegem**, "Regularization techniques in non and semi-parametric instrumental regression", Econometrics seminar of the Tinbergen Institute Amsterdam, Amsterdam, The Netherlands

#### March 2007

**Ingrid Van Keilegom**, "Change-point tests for the error distribution in nonparametric regression", Department of statistics, Paul Sabatier University, Toulouse, France

**Ingrid Van Keilegom**, "Change-point tests for the error distribution in nonparametric regression", Department of statistics, University of Pau, Pau, France

#### April 2007,

**Christian Hafner**, "Introduction to Statistics of Financial Markets", Haifa University, Israel

**Léopold Simar**, Department of Economics, Union College, Schenectady NY, USA

**Léopold Simar**, Department of Economics, Clemson University, Clemson SC, USA

**Ingrid Van Keilegom**, Department of economics, University of Bonn, Germany

#### May 2007

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

**Léopold Simar**, "Introduction to the Bootstrap with Applications in Econometrics", Department of Statistical Science, University of Bologna, Italy

**Léopold Simar**, "Stochastic Versions of FDH/DEA Envelopment Estimators", Statistics Seminar, GREMAQ, Toulouse I, France

#### June 2007

**Sébastien Van Bellegem**, "Local adaptive estimation for nonstationary time series", Statistics seminar, Ensaï, Rennes, France

**Ingrid Van Keilegom**, "A goodness-of-fit test for semiparametric models in multiresponse regression", Department of statistics, University of Vigo, Vigo, Spain

**Ingrid Van Keilegom**, "Survival analysis", doctoral course, University of Vigo, Vigo, Spain

**Ingrid Van Keilegom**, "Estimation of a semiparametric transformation mode", Department of statistics, University of Santiago de Compostela, Spain

#### July 2007

**Sébastien Van Bellegem**, "Local adaptive estimation for nonstationary time series", Kolloquium über Statistik, Institut für Angewandte Mathematik, Universität Heidelberg, Germany

**Ingrid Van Keilegom**, Department of statistics, University of Minnesota, USA

**Ingrid Van Keilegom**, Department of statistics, Iowa State University, USA

#### August 2007

**Sébastien Van Bellegem**, Brown University, Providence, R.I, USA

#### September 2007

**Léopold Simar**, Facoltà di Ingegneria, Università di Pisa, Italy

**Ingrid Van Keilegom**, "Flexible modeling based on copulas in nonparametric regression", Department of statistics, University of Mannheim, Germany

**Ingrid Van Keilegom**, Department of statistics, Paul Sabatier University, Toulouse, France

#### November 2007

**Ingrid Van Keilegom**, Department of statistics, University of Minnesota, USA

**Ingrid Van Keilegom**, "Flexible modeling based on copulas in nonparametric regression", Department of statistics, Iowa State University, USA

#### December 2007

**Ingrid Van Keilegom**, Department of statistics, University of Santiago de Compostela, Spain

### 3.10 Participation to conferences and scientific meetings

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

#### January-February 2007

**Léopold Simar, "Policies for Research and Innovation in the Move towards the European research area",** PRIME Annual Conference, Università di Pisa, Italy

#### March 2007

**Léopold Simar, "Semiparametric and Nonparametric Methods in Econometrics",** invited, Workshop on Semiparametric and Nonparametric Methods in Econometrics, March 19-23, 2007, Oberwolfach, Germany

**Ingrid Van Keilegom, "Estimation of a semiparametric transformation model",** invited, Workshop on Semiparametric and Nonparametric Methods in Econometrics, March 19-23, 2007, Oberwolfach, Germany

**Ingrid Van Keilegom, "Change-point tests for the error distribution in nonparametric regression",** invited, First Joint Statistical Meeting of the 'Deutsche Arbeitsgemeinschaft Statistik', March 27-30, 2007, Bielefeld, Germany

#### April 2007

**Léopold Simar, "Recent Advances in Nonparametric Estimation of Efficient Production Frontiers",** invited, Workshop at the Department of Business and Economics, Union College, Schenectady, New-York, USA

#### May 2007

**Dominique Deprins, "En désespoir de cause, la norme ?",** invited, Congrès International organisé par la Ligue Bruxelloise Francophone pour la Santé Mentale (LBFSM), Bruxelles, Belgium

**Dominique Deprins, "Jusqu'ici tout va bien... Mouvements en santé mentale : entre clinique, social et politique",** invited, Parc des Expositions de Bruxelles, Bruxelles, Belgium

#### June 2007

**Léopold Simar, "Robust Alternatives to Estimate Benchmark Frontiers",** invited, Workshop Productivity and Efficiency of European Banks, Dipartimento di Scienze Economiche, Università di Verona, Italy

**Léopold Simar,** as member of the Scientific Committee, 39<sup>o</sup> Journées de Statistique, Société Française de Statistique, Angers, France

**Léopold Simar,** as member of the Scientific Committee and invited discussant, 10<sup>th</sup> European Workshop on Efficiency and Productivity Analysis, Lille, France

**Ingrid Van Keilegom, "On the use of the bootstrap in nonparametric functional regression",** invited, Journées de Statistique Fonctionnelle et Opératoirelle, June 21-22, Lille, France

#### July 2007

**Johan Segers, "Regularly varying multivariate time series",** invited, INFORMS Applied Probability conference, July 11, Eindhoven, The Netherlands

**Rainer von Sachs, "A multiscale approach for statistical characterization of functional images"**, invited, 6<sup>th</sup> International Congress on Industrial and Applied Mathematics (ICIAM), July 16-20, Zürich, Switzerland

**Ingrid Van Keilegom, "A goodness-of-fit test for parametric and semiparametric models in multiresponse regression"**, invited, Joint Statistical Meetings, July 29 -August 2, Salt Lake City, UT, USA

**Sébastien Van Bellegem, "Local adaptive estimation for nonstationary time series"**, invited, Joint Statistical Meetings, July 29 -August 2, Salt Lake City, UT, USA

#### August 2007

**Géry Geenens, "Non- and semiparametric test for conditional independence in two-way contingency tables"**, invited, Advances in semiparametric methods and applications, Satellite meeting of ISI 2007, Lisbon, Portugal

**Géry Geenens, "Nonparametric test for conditional independence in two-way contingency tables"**, invited, 15th European Young Statisticians Meeting, September 10-14, Castro Urdiales, Spain

**Christian Hafner, "Locally Stationary Factor Models"**, invited, 61st European Meeting of the Econometric Society, Budapest, Hungary

**Johan Segers, "Nonparametric inference for bivariate extreme-value copulas when margins are unknown"**, invited, 56<sup>th</sup> Session of the International Statistical Institute, August 25, Lisbon, Portugal

**Léopold Simar, "Frontier and Boundary Estimation"**, (as member of the Scientific Committee) Advances in semiparametric methods and applications, Satellite meeting of ISI 2007, Lisbon, Portugal

**Sébastien Van Bellegem, "Local adaptive estimation for nonstationary time series"**, invited, Advances in semiparametric methods and applications, Satellite meeting of ISI 2007, Lisbon, Portugal

**Sébastien Van Bellegem, "Local adaptive estimation for nonstationary time series"**, invited, 61st European Meeting of the Econometric Society, Budapest, Hungary

**Rainer von Sachs, "Recent advances in multiscale lifting"**, invited, 56<sup>th</sup> Session of the International Statistical Institute, August 25, Lisbon, Portugal

#### September 2007

**Bernadette Govaerts, "Analytical method validation based on risk estimation and tolerance intervals "**, Conference ENBIS, Dortmund, Germany

**Léopold Simar, "Efficiency and Productivity in European Universities"**, invited, AQUAMETH meeting, PRIME Network of Excellence, 6th EC Framework Programme, University of Sussex, United Kingdom

#### October 2007

**Dominique Deprins, "Quel type d'homme se construit sous le paradigme de Précaution ?"**, invited, Présentation du nouvel Observatoire du Principe de Précaution sous la présidence de F. EWALD, CNAM, Paris, France

**Dominique Deprins, "En désespoir de cause, la norme ? D'une causalité probabiliste à une causalité psychique"**, invited, Colloque *Psychanalyse et Pratiques Sociales* organisé aux F.U.S.L. en hommage à J. FLORENCE, Bruxelles, Belgium

**Léopold Simar, "Nonparametric Stochastic Frontier Models"**, invited, CREST-ENSAI, Rennes, France

**Léopold Simar, "Stochastic Versions of FDH/DEA Envelopment Estimators"**, (keynote speaker), 15<sup>th</sup> Annual Meeting of the Belgian Statistical Society, University of Antwerp, Antwerp, Belgium

**Ingrid Van Keilegom, "Flexible modeling based on copulas in nonparametric regression"**, invited, Workshop on "Reassessing the Paradigms of Statistical Model Building", October 22-26, Oberwolfach, Germany

**November 2007**

**Léopold Simar, "Economies of scale in teaching and research: an European comparison at discipline level"**, invited, Italian Research in Innovation Systems (IRIS) Workshop, Politecnico di Torino, Italy

**December 2007**

**Rainer von Sachs, "Locally stationary factor models"**, invited, Séminaire Parisien de Statistique, December 10, 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 on the page <http://www.uclouvain.be/prog-2008-minstat.html>

### 4.3 New Master programs in Statistics

In September 2007, the Institute launched its two new Master programs: one in Statistics, and one in Biostatistics. The goal of the Master is to attract students holding a Bachelor diploma, into a 2 year program in statistics. The Master in Statistics is declined in two directions: the "finalité approfondie" is research-oriented, whereas the "finalité spécialisée" is business-oriented.

Both programs present a combination of courses, seminars, and research projects within the UCL or in collaboration with private organizations. They are tailored to attract every student interested by statistics, from any field. The full program can be found on: [www.uclouvain.be/prog-2008-stat2m.html](http://www.uclouvain.be/prog-2008-stat2m.html) and <http://www.uclouvain.be/prog-2008-bsta2m.html>.

### 4.4 Certificate

The Institute of Statistics also proposes a *Certificate in statistics* that is mostly devoted to adult students. This program aims at providing 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. The program is jointly organized with IUF (Institut universitaire de formation continue) and more details can be found on <http://www.uclouvain.be/formationcontinue-statistique.html>.

## 4.5 Doctoral School

### 4.5.1. Presentation

In 2007, the Institute of Statistics continued to play the leading role in the administration and organization of the FNRS Graduate School in Statistics and Actuarial Sciences, an interuniversity school with partners in 9 departments at 5 French-speaking universities (ULB, UCL, ULg, FUNDP, FUSAG).

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 2007, the Graduate School has organized an impressive number of short courses given by international visitors in various fields of methodological and applied statistics, some of which during a "Summer school on nonparametric statistics", held at Louvain-la-Neuve. 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 collaborate 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.

In order to share the information about our activities among the members of our Graduate School, we have further developed our web page. The Graduate School being part of the FNRS Graduate College in Science, this page is now to be found at the address:

[http://www.graduatecollegescience.be/edt\\_objectifs.php?lang=en&ecole=STATS](http://www.graduatecollegescience.be/edt_objectifs.php?lang=en&ecole=STATS).

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.

## 4.5.2 Doctoral dissertations and PhD thesis in progress

### Doctoral Dissertations

Almeida, Carlos

*Testing specifications in partial observability models : a Bayesian encompassing approach*, 04-10-2007,

Supervisor : M. Mouchart

El Ghouch, Anouar

*Nonparametric statistical inference for dependent censored data*, 05-10-2007,

Supervisor : I. Van Keilegom

### Ph.D. thesis in progress

Hilmar Böhm

*Shrinkage estimation in the frequency domain of 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 Belleghem

Rachida El Mehdi

*L'analyse d'efficience des domaines de développement au Maroc*

Supervisor : Christian Hafner

Nancy François

*Statistical analysis of time intensity curves in sensory analysis*

Supervisors : Bernadette Govaerts and Philippe Lambert

Bernard Francq

*Development of statistical tools to test the equivalence between analytical measurement methods*

Supervisor : Bernadette Govaerts

Jean-Marc Freyermuth

*Tree-structured wavelet thresholding with applications in nonparametric curve estimation*

Supervisor : Rainer von Sachs

Gery Geenens

*Conditional independence tests for two-way contingency tables*

Supervisor : Léopold Simar

Julien Hunt

*Calcul stochastique en univers semi-markovien et applications financières*

Supervisor : Pierre Devolder

Astrid Jullion

*Statistical analysis of PET (Positron Emission Tomography) scan data*

Supervisor : Philippe Lambert

Thomas Meinguet

*Extreme value theory for stochastic processes with continuous time and/or spac coordinates*

Supervisor : Johan Segers

Giovanni Motta

*Factor modelling and nonparametric estimation for multivariate locally stationary time series*

Supervisor : Rainer von Sachs

Sandra Pitrebois

*Bonus-malus scales and segmented tariffs*

Supervisors : Michel Denuit and Jean-François Walhin

Olga Reznikova

*Adaptive modelling of the dependence in multivariate time series*

Supervisor : Christian Hafner

Réjane Rousseau

*Outils statistiques pour identification de biomarqueurs de toxicité métaboliques*

Supervisors : Bernadette Govaerts and Michel Verleysen

Bianca Teodorescu

*Generalized linear conditional models under left truncation and right censoring*

Supervisor : Ingrid Van Keilegom

Catherine Timmermans

*Analyse de données fonctionnelles appliquées à la physique solaire et à l'étude des relations terre-soleil*

Supervisors : Rainer von Sachs, Véronique Delouille (ORB)

### 4.5.3. Doctoral Seminars

February 15, 2007

Olga Reznikova, Institute of Statistics, UCL

*Adaptive copula estimation : sensitivity analysis and application*

March 8, 2007

Réjane Rousseau, Institute of Statistics, UCL

*Methods of pre-treatment for spectral metabonomics data of 1H-NMR spectroscopy*

March 15, 2007

Astrid Jullion, Institute of Statistics, UCL

*Spatially adaptive Bayesian P-splines models with copulas*

April 19, 2007

Julien Hunt, Institute of Statistics, UCL

*Option pricing : stochastic volatility models*

May 3, 2007

Gery Geenens, Institute of Statistics, UCL

*Nonparametric test for conditional independence in two-way contingency tables*

May 4, 2007

Bianca Teodorescu, Institute of Statistics, UCL

*Goodness-of-fit tests for conditional models with time-dependent coefficients under censoring and truncation - A case study*

May 10, 2007

Olga Reznikova, Institute of Statistics, UCL

*Adaptive copula estimation with applications dans le cadre de Workshop on Multivariate Time Series Modelling, CORE*

September 21, 2007 **Young researchers days**

Bianca Teodorescu, Institute of Statistics, UCL, Belgium

*Goodness-of-fit tests for conditional models with time-dependent coefficients under censoring and truncation*

Rachida El Mehdi, Institute of Statistics, UCL, Belgium

*Efficiency Analysis Application to financing of municipalities in Morocco*

Alexandra Réjane Rousseau, Institute of Statistics, UCL, Belgium

*Combination of Independent Component Analysis and modelisations for the identification of Metabonomic biomarker in 1H-NMR spectroscopy*

Gery Geenens, Institute of Statistics, UCL, Belgium

*Non- and semiparametric tests for conditional independence in two-way contingency tables*

Aurélien Miller, Institut interdépartemental des sciences actuarielles, UCL, Belgium

Julien Hunt, Institute of Statistics, UCL, Belgium

*Options and semi-Markov regime-switching*

Olga Reznikova, Institute of Statistics, UCL, Belgium

*Adaptive estimation procedure with applications*

#### 4.5.4 Short courses

**"Modelling dependence with copulas"** by Christian Genest, (Université Laval, Québec),  
January 29 - February 1, 2007

**"Business forecasting"**, by Professor Andrea Guizzardi, (Università di Bologna, Italy),  
March 26 - April 7, 2007

**"Advanced nonparametric statistics"**, Summer school organised by the Graduate School,  
May 21-24, 2007, (cfr page 25)

**"Statistics and genomics"**, by Amy Anderson, (Western Washington University, USA),  
November 22-24, 2007

## 5. Support for the University and internal Consulting Activities

---

### 5.1 Presentation

One fundamental mission of the Institute of Statistics 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 SMCS “Statistical methods and computing service” has been created in 2007 (<http://www.stat.ucl.ac.be/SMCS/>). It offers a more integrated, recognized and visible structure to three services already offered by the Institute of Statistics in the past:

The consulting service gives advice and collaboration to researchers and members of the university community about 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.

In addition, 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. These seminars welcome students in statistics but also other participants from university, business or industry.

### 5.2 SMCS activities

In 2007, the consulting service has offered numerous advices to members of the university community. Among them, more than 60 have been identified as “conseils” (advices of more than 2 hours of consulting work) and are listed on the consulting service web page : <http://www.stat.ucl.ac.be/SMCS>. 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 2007, the consulting service also organised 9 courses on the use of statistical softwares (SAS, SAS-Guide, SPSS, R) for the researchers and administrative members of UCL and Louvain Academy.

The results of some substantial consulting activities are summarized in formal consulting reports.

#### Consulting reports

**CR 0701 - Schémas, modèles, situations images: apprentissage de la biologie à l'université. Analyse statistique des résultats d'entretiens semi-structurés réalisés auprès d'étudiants de biologie de l'UCL.**

**Auteurs:** Benoit, A., Govaerts, B., Guillet, A. and C. Ritter

**CR 0702 - Study on the long-term a thermal biological effects of radar irradiation.**

**Auteur:** Rousseau, R.

**CR 0703- Implémentation d'un suivi qualité du processus de traitement du sang dans les sites de préparation des produits sanguins du Service du sang de la Croix-Rouge de Belgique.**

**Auteurs:** Benoit, A., Govaerts, B., Guillet, A. and A. Rapaille

**CR 0704 - Evaluation des programmes d'études: conception et réalisation de fiches résumés des évaluations de cours par année d'étude.**

**Auteurs:** Ourliac, A., Rakotondrabe, V., Govaerts, B., Guillet, A. and C. Ritter

**CR 0705 - Résolution de problèmes probabilistes et statistiques liés à l'évaluation de la diversité de collections de mutants et au titrage de suspensions de bactéries en biochimie.**

**Auteurs:** Govaerts, B. and J.-M. Rolin

## 5.3 Applied Statistics Workshops

### March 2, 2007

Claire Dujardin, Unité de géographie, UCL, Belgium  
*"Ségrégation résidentielle et chômage dans l'agglomération bruxelloise"*

### March 16, 2007

Antje Christensen, Novo Nordisk Pharma  
*"The Tale of the Wounded Rat Obtaining and Explaining Data for a "Go No Go" decision in a Pharmaceutical Development Project - A case study"*

Claudia Gajewski and Heiko Woith, Geophysical Institute Postdam, Germany  
*"A possible influence of the 29 October 2003 magnetic storm on radon cycles in the thermal spring Tiberias, Israël"*

### March 30, 2007

Anne de Frenne, MathX, Belgium  
*"Les SLP (Synthetic Load Profil) du marché belge de l'électricité et du gaz pour les clients résidentiels"*

### April 27, 2007

Manuel Piette, Fortis, Belgium  
*"The Logit Target Mean approach for binary classification"*

Philippe Vanden Eeckhaut, Université de Lille, France  
*"Statistical Inference and nonparametric Cost Frontier - An Illustration on Electricity Network in Spain"*

### May 4, 2007

Claire Beguin, Hôpitaux universitaires Saint-Luc, Belgium

Vincent Lorant, Ecole de Santé publique, UCL, Belgium  
*"Bias in measuring socio-economic inequalities in health : a comparison of a survey and a census"*

### October 5, 2007

Christian Ritter, Institute of Statistics, UCL, Belgium  
*"Presentation of Quantitative Information"*

Christian Ritter, Institute of Statistics, UCL, Belgium  
*"Combining Software Tools to Practice Statistics"*

### November 9, 2007

Anne Benoit, UCL - Alain Guillet, Institute of Statistics, Belgium  
*"Implémentation d'un suivi qualité du processus de traitement du sang pour la Croix-Rouge de Belgique"*

Walthère Déwé, GSK biologicals - Bernadette Govaerts, Institute of Statistics, UCL, Belgium  
*"Validation d'une méthode analytique en laboratoire de recherche et production industrielle : Principes généraux et méthodologie par estimation de risque et intervalles de tolérance"*

### November 23, 2007

Stephen Evans, London School of Hygiene and Tropical Medicine, UK  
*"Lessons from a trial : statistical methods for demonstrating misconduct and problems with journals"*

Marc Buyse, IDDI, Belgium  
*"SAFE : an automated system to detect errors and fraud in clinical trials"*

**December 7, 2007**

Stéphane Laurent, GSKBio, Rixensart, Belgium  
*"Analyse bayésienne de l'efficacité vaccinale"*

Jean-Marc Freyermuth, Institute of Statistics, UCL, Belgium  
*"Gestion bayésienne de stocks de pièces de rechange"*

## 6. Appendix

---

### 6.1. Abstracts of the discussion papers

#### **DP 0701 - Neural modelling of ranking data with an application to stated preference data**

Krier, C., Mouchart, M. and A. Oulhaj

Although neural networks are commonly encountered to solve classification problems, ranking data present specificities which require adapting the model. Based on a latent utility function defined on the characteristics of the objects to be ranked, the approach suggested in this paper leads to a perceptron-based algorithm for a highly non linear model. Data on stated preferences obtained through a survey by face-to-face interviews, in the field of freight transport, are used to illustrate the method. Numerical difficulties are pinpointed and a Pocket type algorithm is shown to provide an efficient heuristic to minimize the discrete error criterion. A substantial merit of this approach is to provide a workable estimation of contextually interpretable parameters along with a statistical evaluation of the goodness of fit.

#### **DP 0702 - Identifiability and estimability of parametric Rasch-type models**

San Martin, E. and J.-M. Rolin

It is widely accepted that the Rasch model with abilities distributed according to a normal distribution is identified provided that the mean is fixed at zero. Nevertheless, the following questions arise: why is this result valid? Is it still valid if the distribution of the random effect (namely, the abilities) changes? Is it still valid if both the link function defining the item characteristic curve (ICC) and the distribution of the random effect change? What about the asymptotic behavior of the posterior expectation of identified parameters? Finally, what can we learn about model construction from these eventual identification results? This paper is intended to answer these questions.

#### **DP 0703 - Multivariate regular variation of heavy-tailed Markov chains**

Segers, J.

The upper extremes of a Markov chain with regularly varying stationary marginal distribution are known to exhibit under general conditions a multiplicative random walk structure called the tail chain. More generally, if the Markov chain is allowed to switch from positive to negative extremes or vice versa, the distribution of the tail chain increment may depend on the sign of the tail chain on the previous step. But even then, the forward and backward tail chain mutually determine each other through a kind of adjoint relation. As a consequence, the finite-dimensional distributions of the Markov chain are multivariate regularly varying in a way determined by the back-and-forth tail chain. An application of the theory yields the asymptotic distribution of the past and the future of the solution to a stochastic difference equation conditionally on the present value being large in absolute value.

#### **DP0704 - Tails of random sums of a heavy-tailed number of light-tailed terms**

Robert, C. Y. and J. Segers

The tail of the distribution of a sum of a random number of independent and identically distributed nonnegative random variables depends on the tails of the number of terms and of the terms themselves. This situation is of interest in the collective risk model, where the total claim size in a portfolio is the sum of a random number of claims. If the tail of the claim number is heavier than the tail of the claim sizes, then under certain conditions the tail of the total claim size does not change asymptotically if the individual claim sizes are replaced by their expectations. The conditions allow the claim number distribution to be of consistent variation or to be in the domain of attraction of a Gumbel distribution with a mean excess function that grows to infinity sufficiently fast. Moreover, the claim number is not necessarily required to be independent of the claim sizes.

**DP 0705 - Bayesian consistency of density estimators**

Rolin, J.-M.

Under a very mild assumption, the consistency of posterior expectations of densities for an independent and identically distributed sample of observations is proved. A particular case of this general result consists in nonparametric densities.

**DP 0706 - Shrinkage estimation in the frequency domain of multivariate time series**

Böhm, H. and R. von Sachs

In this paper on developing shrinkage for spectral analysis of multivariate time series of high dimensionality, we propose a new non-parametric estimator of the spectral matrix with two appealing properties: compared to the traditional smoothed periodogram our shrinkage estimator has a smaller  $L_2$  risk and is numerically more stable due to a smaller condition number. We use the concept of "Kolmogorov" asymptotics where simultaneously the sample size and the dimensionality tend to infinity, to show that the smoothed periodogram is not consistent and to derive the asymptotic properties of our regularized estimator. This estimator is shown to have asymptotically minimal risk among all linear combinations of the identity and the averaged periodogram matrix. Compared to existing work on shrinkage in the time domain, our results show that in the frequency domain it is necessary to take the size of the smoothing span as "effective sample size" into account. Furthermore, we perform extensive Monte Carlo studies showing the overwhelming gain in terms of lower  $L_2$  risk of our shrinkage estimator, even in situations of oversmoothing the periodogram by using a large smoothing span.

**DP 0707 - Testing the normality of the latent variables in the polychoric correlation model**

Almeida, C. and M. Mouchart

This paper develops a specification of the model defining the polychoric correlations, where the manifest ordinal variables are assumedly generated as a discretization of latent continuous ones. Taking advantage of the fact that in this model, the marginal distributions of the latent variables are not identified, we use the tool of the copula concept; some identification issues are analysed. In the second part, we develop a Bayesian encompassing specification test for testing the Gaussianity of the underlying copula and consider the discretization model as a case of partial observability.

**DP 0708 - Gamma kernels estimator for density and hazard rate of right censored data**

Bouezmarni, T., El Ghouch, A. and M. Mesfioui

In this paper we consider the nonparametric estimation for the density and hazard rate functions for right censored data using the kernel smoothing techniques. The "classical" fixed symmetric kernel type estimator of these functions performs well in the interior region, but it suffers from the problem of bias in the boundary region. Here, in order to correct the boundary bias, we propose the gamma kernel estimator for both the density and the hazard rate functions. The mean integrated squared error, the asymptotic normality and the law of iterated logarithm are investigated. We show some simulations results which demonstrate the performance of the proposed estimators.

**DP 0709 - Causality, structural modelling and exogeneity**

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

This paper deals with causal analysis in the social sciences. We first present a conceptual framework according to which causal analysis is based on a rationale of variation and invariance, and not only on regularity. We then develop a formal framework for causal analysis by means of structural modelling. Within this framework we approach causality in terms of exogeneity in a structural conditional model based on (i) model fit, (ii) invariance under a large variety of environmental changes, and (iii) congruence with background knowledge. We also tackle the issue of confounding and show how latent confounders can play havoc with exogeneity. This framework avoids making untestable metaphysical claims about causal relations and yet remains useful for cognitive and action-oriented goals.

**DP 0710 - Development of a new predictive modelling technique to find with confidence equivalence zone and design space of chromatographic analytical methods**

Lebrun, P., Govaerts, B., Debrus, B., Ceccato, A., Caliaro, G., Hubert, P. and B. Boulanger

A new method for modelling chromatographic responses is presented as a critical piece for the achievement of automated development of analytical methods. This methodology is based on four parts. First, we propose to use a very little set of statistical equations to create predictive models for retention time based responses as the apex, the width and the asymmetry of peaks. Second, an experimental design is set up to realize experiments. Third, using grid search over the domain, multi criteria decision is taken with respect to different local or global optimization criteria, used as desirability functions. This allows finding an optimal chromatogram. Fourth, we advice to investigate how the predictive error of the models propagates around optimal solution. This allows to give confidence in the optimal solution, in finding a set of zones that presumably will give an acceptable solution. Design spaces can be derived with a similar technique. The approach is exemplified with a real case and predictions of models at optimal analytical conditions are validated through new experiments. Flexibility is left over all the presented methodology.

**DP 0711 - Comparison of some chemometric tools for metabonomics biomarker identification**

Rousseau, R., Govaerts, B., Verleysen, M. and B. Boulanger

NMR-based metabonomics discovery approaches require statistical methods to extract, from large and complex spectral databases, biomarkers or biologically significant variables that best represent defined biological conditions. This paper explores the respective effectiveness of six multivariate methods: multiple hypotheses testing, supervised extensions of principal (PCA) and independent components analysis (ICA), discriminant partial least squares, linear logistic regression and classification trees. Each method has been adapted in order to provide a biomarker score for each zone of the spectrum. These scores aim at giving to the biologist indications on which metabolites of the analyzed biofluid are potentially affected by a stressor factor of interest (e.g. toxicity of a drug, presence of a given disease or therapeutic effect of a drug). The applications of the six methods to samples of 60 and 200 spectra issued from a semi-artificial database allowed to evaluate their respective properties. In particular, their sensitivities and false discovery rates (FDR) are illustrated through receiver operating characteristics curves (ROC) and the resulting identifications are used to show their specificities and relative advantages. The paper recommends to discard two methods for biomarker identification: the PCA showing a general low efficiency and the CART which is very sensitive to noise. The other 4 methods give promising results, each having its own specificities.

**DP 0712 - Local linear quantile regression with dependent censored data**

El Ghouch, A. and I. Van Keilegom

It is known from the literature that the least squares regression estimator is optimal and is equivalent to the maximum likelihood estimator when the errors follow a normal distribution. However, in many non-Gaussian situations they are far from being optimal. This is especially the case when the data involve asymmetric distributions as is the case, in general, with survival times. Another known drawback of the least squares method is its extreme sensitivity to modest amounts of outlier contamination. An attractive alternative to the classical regression approach based on the quadratic loss function is the use of the absolute error criterion, which leads to the well known median regression function, or more generally, to the quantile regression function. In this paper, we consider the problem of nonparametrically estimating the conditional quantile function from censored dependent data. The method proposed here is based on a local linear fit using the check function approach. The asymptotic properties of the proposed estimator are established. Since the estimator is defined as a solution of a minimization problem, we also propose a numerical algorithm. We investigate the performance of the estimator for small samples through a simulation study, and we also discuss the optimal choice of the bandwidth parameters.

**DP 0713 - Empirical likelihood based testing for regression**

Van Keilegom, I., Sánchez-Sellero, C. and W. González-Manteiga

Consider a random vector  $(X, Y)$  and let  $m(x) = E(Y | X = x)$ . We are interested in testing  $H_0 : m \in \mathcal{M}_{\Theta, \mathcal{G}} = \{\gamma(\cdot, \theta, g) : \theta \in \Theta, g \in \mathcal{G}\}$  for some known function  $\gamma$ , some compact set  $\Theta \subset \mathbb{R}^p$  and some function set  $\mathcal{G}$  of real valued functions.

Specific examples of this general hypothesis include testing for a parametric regression model, a generalized linear model, a partial linear model, a single index model, but also the selection of explanatory variables can be considered as a special case of this hypothesis. To test this null hypothesis, we make use of the so-called marked empirical process introduced by Diebolt (1995) and studied by Stute (1997) for the particular case of parametric regression, in combination with the modern technique of empirical likelihood theory in order to obtain a powerful testing procedure. The asymptotic validity of the proposed test is established, and its finite sample performance is compared with other existing tests by means of a simulation study.

**DP 0714 - Semiparametric multivariate density estimation for positive data using copulas**

Bouezmarni, T. and J. Rombouts

In this paper we estimate density functions for positive multivariate data. We propose a semiparametric approach. The estimator combines gamma kernels or local linear kernels, also called boundary kernels, for the estimation of the marginal densities with semiparametric copulas to model the dependence. This semiparametric approach is robust both to the well known boundary bias problem and the curse of dimensionality problem. We derive the mean integrated squared error properties, including the rate of convergence, the uniform strong consistency and the asymptotic normality. A simulation study investigates the finite sample performance of the estimator. We find that univariate least squares cross validation, to choose the bandwidth for the estimation of the marginal densities, works well and that the estimator we propose performs very well also for data with unbounded support. Applications in the field of finance are provided.

**DP 0715 - A non-parametric bayesian method to smooth PET time-activity-curves**

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

Positron Emission Tomography (PET) is an imaging technique in which a radionuclide is introduced into a molecule of potential biological relevance (to form what is called a tracer) and administered to a patient. The regional evolution of the uptake of the tracer over time is called a Time-Activity-Curve (TAC) and is used to derive some clinical measures that give information about the process under study. One of these measures is the Distribution Volume (DV) which can be estimated by several methods, notably the Graphical Analysis Method (GA). It has been shown that using GA method on noisy TAC leads to a systematic underestimation of the Distribution Volume (Hsu et al., 1997; Slifstein and Laruelle, 1999). We propose a method that allows to smooth the Time-Activity-Curves in a non-parametric way by using Bayesian P-splines. This method may be used in all the cases, whatever the compartmental model that might underly the observed data. Simulations have shown that this method gives an unbiased estimation of the true TAC, whatever the level of noise. We show that smoothing TAC with the non-parametric method before computing the Distribution Volume allows to reduce considerably the bias. Logan et al. (2001) proposes to smooth data before computing the Distribution Volume by using the Generalized Linear Least Squares (GLLS) method if a one-tissue compartment model is considered and by applying the GLLS to the data in two parts for a two-tissues compartment model : one set of parameters is estimated from times 0 to T1 and a second set from T1 to the end time where T1 has to be chosen from data. This method provides good results but, if we are not sure about the compartmental model suitable for the data or, if we want to avoid the choice of T1, using the Bayesian nonparametric model is a good alternative.

#### DP 0716 - Nonparametric density estimation for multivariate bounded data

Bouezmarni, T. and J. Rombouts

We propose a new nonparametric estimator for the density function of multivariate bounded data. As frequently observed in practice, the variables may be partially bounded (e.g., nonnegative) or completely bounded (e.g., in the unit interval). In addition, the variables may have a point mass. We reduce the conditions on the underlying density to a minimum by proposing a nonparametric approach. By using a gamma, a beta, or a local linear kernel (also called boundary kernels), in a product kernel, the suggested estimator becomes simple in implementation and robust to the well known boundary bias problem. We investigate the mean integrated squared error properties, including the rate of convergence, uniform strong consistency and asymptotic normality. We establish consistency of the least squares cross-validation method to select optimal bandwidth parameters. A detailed simulation study investigates the performance of the estimators. Applications using lottery and corporate finance data are provided.

#### DP 0717 - Regularly varying multivariate time series

Basrak, B. and J. Segers

Extreme values of a stationary, multivariate time series may exhibit dependence across coordinates and over time. Aim of this paper is to offer a new and potentially useful tool called tail process to describe and model such extremes. The key property is the following fact: existence of the tail process is equivalent to multivariate regular variation of finite cuts of the original process. Certain remarkable properties of the tail process are exploited to shed new light on known results on certain point processes of extremes. The theory is shown to be applicable with great ease to stationary solutions of stochastic autoregressive processes with random coefficient matrices, an interesting special case being a recently proposed factor GARCH model. In this class of models, the distribution of the tail process is calculated by a combination of analytical methods and a novel sampling algorithm.

#### DP 0718 - Rank-based inference for bivariate extreme-value copulas

Genest, C. and J. Segers

Several nonparametric estimators are available for the Pickands dependence function of a bivariate extreme-value copula. All of them, however, require knowledge of the univariate margins. In this paper, rank-based versions of some of these estimators are proposed for the case where the margins are unknown. In particular, the asymptotic distributions of rank-based versions of the estimators of Pickands [*Bull. Inst. Internat. Statist.* 49 (1981) 859–878] and Capéreaù, Fougères and Genest [*Biometrika* 84 (1997) 567–577] are found using results on the limit behavior of a class of weighted bivariate empirical processes. Small- and large-sample comparisons indicate that even when the margins are known, a more efficient estimator arises if the information about the margins is ignored and ranks are used instead.

#### DP 0719 - A review of quantitative structure-activity relationship (QSAR) models

Le Bailly de Tillegem, C. and B. Govaerts

No abstract available.

#### DP 0720 - Goodness-of-fit tests for multiplicative models with dependent data

Dette, H., Pardo-Fernández, J. C. and I. Van Keilegom

Several classical time series models can be written as a regression model of the form  $Y_t = m(X_t) + \sigma(X_t) \varepsilon_t$ , where  $(X_t, Y_t)$ ,  $t = 0, \pm 1, \pm 2, \dots$  is a bivariate strictly stationary process. Some of those models, such as ARCH or GARCH models, share the property of proportionality of the regression function,  $m$ , and the scale function,  $\sigma$ . In this article, we present a procedure to test for this feature in a nonparametric context, which is a preliminary step to identify certain time series models. The test is based on the difference between two non parametric estimators of the distribution of the regression the error. Asymptotic results are proved and some simulations are shown in the paper in order to illustrate the finite sample properties of the procedure.

**DP 0721 - Identification and estimation by penalization in nonparametric instrumental regression**

Florens, J.-P., Johannes, J. and S. Van Bellegem

The nonparametric estimation of a regression function  $\varphi$  from conditional moment restrictions involving instrumental variables is considered. The rate of convergence of penalized estimators is studied in the case where  $\varphi$  is not identified from the conditional moment restriction. We also study the gain of modifying the penalty in the estimation, considering for instance a Sobolev-type of penalty. We analyze the effect of this modification on the rate of convergence of the estimator and on the identification of the regression function  $\varphi$ .

**DP 0722 - A method of moments estimator of tail dependence**

Einmahl, J. H. J., Krajina, A. and J. Segers

In the world of multivariate extremes, estimation of the dependence structure still presents a challenge and an interesting problem. A procedure for the bivariate case is presented that opens the road to a similar way of handling the problem in a truly multivariate setting. We consider a semiparametric model in which the stable tail dependence function is parametrically modeled. Given a random sample from a bivariate distribution function, the problem is to estimate the unknown parameter. A method of moments estimator is proposed where a certain integral of a nonparametric, rank-based estimator of the stable tail dependence function is matched with the corresponding parametric version. Under very weak conditions, the estimator is shown to be consistent and asymptotically normal. Moreover, a comparison between the parametric and nonparametric estimators leads to a goodness-of-fit test for the semiparametric model. The performance of the estimator is illustrated for a discrete spectral measure that arises in a factor-type model and for which likelihood-based methods break down.

A second example is that of a family of stable tail dependence functions of certain meta-elliptical distributions.

**DP 0723 - On concordance measures for discrete data and dependence properties of Poisson model**

Bouezmarni, T., Mesfioui, M. and A. Tajar

The purpose of this paper is twofold. First, we study Kendall's tau and Spearman's rho concordance measures for discrete variables. We mainly propose their best bounds using positive dependence properties. Second, we provide useful dependence properties of the bivariate Poisson distribution and show the relationship between parameters of the Poisson distribution and both tau and rho.

**DP 0724 - Flexible modeling based on copulas in nonparametric regression**

Braekers, R. and I. Van Keilegom

Consider the model  $Y = m(X) + \varepsilon^x$ , where  $m(\cdot) = \text{med}(Y|\cdot)$  is unknown but smooth. It is often assumed that  $\varepsilon$  and  $X$  are independent. However, in practice this assumption is in many cases violated. In this paper we propose to model the dependence between  $\varepsilon$  and  $X$  by means of a copula model, i.e.  $(\varepsilon, X) \sim \mathcal{C}_\theta(F_\varepsilon(\cdot), F_X(\cdot))$ , where  $\mathcal{C}_\theta$  is a copula function depending on an unknown parameter  $\theta$ , and  $F_\varepsilon$  and  $F_X$  are the marginals of  $\varepsilon$  and  $X$ . Since many parametric copula families contain the independent copula as a special case, the so-obtained regression model is more flexible than the 'classical' regression model.

We estimate the parameter  $\theta$  via a pseudo-likelihood method and prove the asymptotic normality of the estimator, based on delicate empirical process theory. We also study the estimation of the conditional distribution of  $Y$  given  $X$ . The procedure is illustrated by means of a simulation study, and the method is applied to data on food expenditures in households.

**DP 0725 - Analysing total factor productivity through a bootstrap-based nonparametric approach: the case of Italian regions**

Bonaccorsi, A., Curi, C., Daraio, C., Mancuso, P. and L. Simar

This paper analyzes the Total Factor Productivity (TFP) growth of the Italian regions over the period 1980-2001. Malmquist Productivity Index (MPI) and its components (namely Efficiency

Change –EffCh- and Technical Change -TechCh) as well as confidence intervals have been estimated by applying the procedure proposed by Simar and Wilson (1999). Including human capital among inputs, we estimated an overall bias-corrected productivity gain of 2.1 percent, an efficiency gain of 0.5 and a technical gain of 1.6 percent. The bootstrap analysis revealed that for most Italian regions efficiency and technical changes did not show a statistically significant change.

## 6.2. Abstracts of the consulting reports

### **CR0702 – Study on the long-term a thermal biological effects of radar irradiation.**

Rousseau, R. (2007)

Bien que l'exposition aux micro-ondes soit devenue fréquente (gsm, four), son impact sur la santé est encore inconnu. Une collaboration entre l'armée et l'UCL souhaite explorer les effets biologiques à long terme de l'exposition aux micro-ondes de faible chaleur. Des données ont été récoltées au cours d'une expérience exposant des rats à différents types de radiations pendant un an (30 rats soumis à des ondes continues de 1GHz, 30 exposés à des ondes modulées, 30 recevant des ondes continues de 10Ghz et un groupe contrôle de 30 rats non exposés). La concentration sanguine de 15 paramètres cellulaires et hormonaux a été mesurée pour chaque rat, lors de 5 périodes différentes.

L'Institute of Statistics de l'UCL a été contactée afin d'analyser chaque type de données sanguines dans les nombreux aspects envisageables liés au temps et/ou type d'exposition ainsi que les éventuels liens entre évolution des différents paramètres.

Les questions suivantes sont abordées :

- La sélection des paramètres sanguins les plus altérés par une exposition (tout type d'exposition et toute période d'évaluation confondues) au moyen d'une méthode fondée sur le false discovery rate.
- La visualisation des corrélations entre les différents paramètres sanguins au moyen d'analyses en composantes principales des 15 paramètres pour une période donnée.
- La visualisation des évolutions temporelles des différents paramètres sanguins au moyen d'une analyse en composante principale. Les projections des 75 paramètres issus des 5 périodes a permis de tracer des trajectoires représentant l'évolution de chaque paramètre.
- Pour chaque paramètre et à chaque période, comparaison des moyennes entre les 4 groupes au moyen d'ANOVA.
- Successivement, pour chacun des 3 groupes exposés, pour chaque paramètre et à chaque période, comparaison de la moyenne du groupe à celle du groupe contrôle.