Contents

1 INTRODUCTION

The following report briefly describes the results of the research activities concerning the year 2004.

Presentation

The Institut de statistique has been created in 1992. It coordinates and organises all the statistical activities at the Université catholique de Louvain: research, teaching and consulting.

Research

The research activities of the Institute cover a large number of fields. A major part of the activities are centered on mathematical statistics (semi- and non-parametric statistics, bayesians statistics, multivariate analysis, regression estimation, mixtures models, survival analysis, time series, ...) with important implications in various fields of application (econometrics, biostatistics, insurance, industrial statistics, transportation problems,...).

Since January 2002, the Institut de statistique has been the coordinator of an IAP network : "Statistical techniques and modeling for complex substantive questions with complex data".

This network includes 6 institutions : Université catholique de Louvain (Belgium), Katholieke Universiteit Leuven (Belgium), Limburgs Universitair Centrum (Diepenbeek, Belgium), Université Libre de Bruxelles (Belgium), Aachen Technical University (Germany) and Université Joseph Fourier (Grenoble, France).

Teaching

In the field of teaching, the activity of the Institute is as follow:

- basic education

- second cycle education

- third cycle education.

- continued education

Concerning the first and second cycle studies, the Institute is responsible for the organisation of probability and statistics courses offered at the university. Moreover, the Institute organises second and third cycle studies: the "master spécialisé en statistique" (Master), the "diplôme d'études approfondies en statistique" (DEA) with several orientations and the "doctorat en statistique". Since the academic year 2002-2003, the Ph. D. degree has been organised within the context of the Graduate School in Statistics.

The Institute is also organizing a "Certificat en statistique" in the framework of continued education.

More information concerning the programs in statistics is available on the webpage http://www.stat.ucl.ac.be/.

Consulting

The Institute offers consulting services to researchers using statistics in their profession. This service includes discussions on statistical problems encountered by researchers working at other departments or laboratories of the Université catholique de Louvain. Consulting services are also offered to business cooperations searching for support in dealing with statistical questions. Furthermore, service courses (continued education) are provided for interested groups.

Other events of the year 2004

Miscellany

Irène GIJBELS is an elected member of the Board of the Belgian Statistical Society.

Léopold SIMAR is member of the jury, selection committee, *Management Science Strategic Innovation Price*, European Asssociation of Operational Research Societies (EURO).

Ingrid VAN KEIGELOM has become an elected member of the International Statistical Institute.

Rainer von SACHS has become an elected Member of the International Statistical Institute since end of 2004.

2 PERSONNEL

Academic Members

Michel DENUIT

Dominique DEPRINS

Irène GIJBELS (until September, 30)

Bernadette GOVAERTS

Philippe LAMBERT (Academic Secretary)

Christian RITTER

Jean-Marie ROLIN (Chairman since September, 1st)

Léopold SIMAR (Chairman until August, 31)

Sébastien VAN BELLEGEM

Ingrid VAN KEILEGOM

Rainer von SACHS

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

Jean-Marie de KETELE (em.), Faculty of Psychology

Eric LE BOULENGE, Faculty of Sciences

Guy LORIES, Faculty of Psychology

Robert PEETERS (em.), Faculty of Economical, Social and Political Sciences

Annie ROBERT, Faculty of Medicine

Michel VERLEYSEN, Faculty of Applied Mathematics

Invited Academic Members

Yves BERGER

Libei CHEN

Gerda CLAESKENS, OSTC return grant (until September, 30)

Anne DE FRENNE

Isabelle DE MACQ (since September, 1st)

Anne-Cécile GODERNIAUX

Emeritus Professor

Michel MOUCHART (em. 2004)

José PARIS (em. 1998)

Researchers and Doctoral Students

Carlos ALMEIDA

Hilmar BÖHM

Jean-Philippe BOUCHER

Taoufik BOUEZMARNI

Natacha BROUHNS

Céline BUGLI

Cindy COURTOIS

Alexandra DASKOVSKA

Antoine DELWARDE

Isabelle DE MACQ (until December, 11)

Anouar EL GHOUCH

Nancy FRANCOIS

Géry GEENENS

Cédric HEUCHENNE

Astrid JULLION

Catherine KRIER (since August, 1st)

Alexandre LAMBERT

Céline LE BAILLY de TILLEGHEM

Xavier MARECHAL

Giovanni MOTTA

S. PITREBOIS

Oana PURCARU

Franck QUASHIE

Réjane ROUSSEAU

Bianca TEODORESCU

Marie VANDRESSE (until June, 30)

Postdoctoral Researchers

Jérémie BIGOT (until August, 31)

Seok-Oh JEONG

Abderrahim OULHAJ (until February, 28)

Valentin ZELENYUK (since September, 1st)

Associate Researchers

Maria KEY PRATO

Philippe VANDEN EECKAUT

Computer Scientists

Laurent BUSET

Eric LECOUTRE

Jean-Luc MARRION (Responsible)

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

Administrative Responsible

Dominique ANDRE (until April, 30)

Chritine DENAYER (since June, 14)

Administrative Staff

Anne BALFROID

Anne-Marie BELLEMANS

Monique DESCAMPS

Sophie MALALI

Short Term Visitors

Marc AERTS, Limburgs Universitair Centrum, Belgium

Anestis ANTONIADIS, University of Grenoble, France

Thomas BRUSS, Université Libre de Bruxelles, Belgium

Daniela COCCHI, University of Bologna, Italy

Pierre DARDENNE, Centre wallon de recherches agronomiques, Département qualité des productions agricoles, Gembloux, Belgium

Aurore DELAIGLE, UCSD, USA

Michel DELECROIX, ENSAI, Rennes, ENSAE, Paris, France

Miguel A. DELGADO, University of Carlos III, Madrid, Spain

Holger DETTE, University of Ruhr, Bochum, Germany

Walthere DEWE, Eli Lilly, Mont-Saint-Guibert, Belgium

Dimitri D'OR, Université catholique de Louvain, Belgium

Holger DREES, University of Hamburg, Germany

Paul EILERS, University of Leiden, The Netherlands

Gilles FAY, Université de Lille, France

Wenceslao GONZALEZ MANTEIGA, University of Santiago de Compostela, Spain

Peter HALL, Australian National University, Australia

Daniel J. HENDERSON, State University of New York, USA

Javier HIDALGO, London School of Economics, United Kingdom

Nils HJORT, University of Oslo, Norway

Joël HOROWITZ, University of Iowa, USA

Geurt JONGBLOED, Vrije Universiteit Amsterdam, The Netherlands

Alois KNEIP, University of Mainz, Germany

Maria LANZERATH, W. L. Gore, Putzbrunn, Germany

Young Kyung LEE, Seoul National University, South Korea

Eckhard LIEBSCHER, Technical University Ilmenau, Germany

Grayham E. MIZON, University of Southampton, United Kingdom

Nathalie NEUMEYER, University of Bochum, Germany

Mila NIKOLOVA, ENST, Paris, France

Abderrahim OULHAJ, Oxford University, United Kingdom

Byeong PARK, Seoul National University, South Korea

Mari PAZ MORAL, University of Bilbao, Spain

Pierre PESTIAUX, Total-Fina, France

Ioana POPESCU, INSEAD, Paris, France

Didier RENARD, Eli Lilly, Mont-Saint-Guibert, Belgium

Juan RODRIGUEZ POO, University of Cantabria, Santander, Spain

Daniele ROMANO, University of Cagliari, Italy

Peihua QIU, University of Minnesota, USA

Cesar SANCHEZ SELLERO, University of Santiago de Compostela, Spain

Ernesto SAN MARTIN, PUC, Santiago, Chile

Eliana SCHEIHING, University Austral of Chile, Valdivia, Chile

Wim SCHOUTERS, Katholieke Universiteit Leuven, Belgium

Tom SNIJDERS, University of Groningen , The Netherlands

Fabian TIBALDI, Limburgs Universitair Centrum, Belgium

Paul W. WILSON, University of Texas, Austin, USA

Valentin ZELENYUK, National University of Kiev, Ukranian

Doctoral Visitors

Cinzia DARAIO, Scuola Superiore S. Anna, Pisa, Italy

Maria Amalia JACOME PUMAR, University of La Coruña, Spain

J. C. PARDO FERNANDEZ, E. U. Enxeneria Tecnica Industrial, Vigo, Spain

3 THE INSTITUTE OF ACTUARIAL SCIENCE

The Institute of actuarial science has received an interdepartmental status since 2002. It is run jointly by the Institut d'Administration et de Gestion (IAG), the UCL business school, and the Institute of Statistics (STAT), combining so an unique expertise as well in deep mathematical background as in business aspects. It offers actuarial studies at master level. The two-year programme is recognized by the Belgian Actuarial Society (KVBA-ARAB), as well as by the Belgian regulatory authorities to be certified Appointed Actuary.

Working regularly with major actuarial departments of other European and North American universities (KULeuven, Lyon, Strasbourg, Barcelona, Amsterdam, London, Laval, etc.) the Institute develops highly respected degree courses, innovative research projects and offers numerous services to the actuarial community (CPD, colloquia, seminars, consulting, etc.). It benefits from several funding programs (PARC, FRIA and several research projects funded by the Région Wallonne) as well as from two chaires (by the Fortis and AXA groups).

In 2004, two books and more than 35 papers have been authored by the members of the Institute, who were granted several international prizes. Thirteen PhD students are supervised by the academic members of the Institute, and the actuarial research group continues to rank well in the international bibliometric studies.

For a detailed account of the activities, we refer the interested reader to the website actu.ucl.ac.be.

4 PUBLICATIONS AND EDITING ACTIVITIES

The Institute publishes a Discussion Papers series and a Reprint series. The papers in both series are the output from the statistical research activities. Many collaborations (national and international) are going on with researchers from abroad. The following Discussion Papers and Reprints were issued during the period concerned by this report.

4.1 Discussion Papers

0401. GIJBELS, I., LAMBERT, A. and P. QIU, Jump-preserving regression and smoothing using local linear fitting : a compromise.

This paper deals with nonparametric estimation of a regression curve, where the estimation method should preserve possible jumps in the curve. At each point x at which one wants to estimate the regression function, the method chooses in an adaptive way among three estimates: a local linear estimate using only datapoints to the left of x, a local linear estimate based on only datapoints to the right of x, and finally a local linear estimate using data in a two–sided neighbourhood around x. The choice among these three estimates is made by looking at differences of the weighted residual mean squares of the three fits. The resulting estimate preserves the jumps well and in addition gives smooth estimates of the continuity parts of the curve. This property of compromise between local smoothing and jump–preserving is what distinguishes our method from previously proposed methods, that mainly focused on local smoothing and consequently blurred possible jumps, or mainly focused on jump–preserving and hence led to rather noisy estimates in continuity regions of the underlying regression curve. Strong consistency of the estimator is established and its performance is tested via a simulation study. This study also compares the current method with some existing methods. The current method is illustrated in analyzing a real dataset.

0402. VANDENHENDE, F. and Ph. LAMBERT, Local dependence estimation using non-parametric archimedean copulas.

The selection of a dependence model is often arbitrary but rarely automated when modeling multivariate data. We define new non-parametric copulas that overcome the problem. We focus our attention to Archimedean copulas. Then, the non–parametric family is obtained using a continuous piecewise log-linear combination of existing Archimedean generators. We provide an efficient constrained least-squares method to estimate the involved coefficients. We show how these coefficients can be interpreted as local dependence measures. We also present a penalization algorithm for smoothing that dependence over the joint distribution. The usefulness of the non-parametric family is illustrated in a example: the lifetime analysis of Danish twins, where dependence increases with decreasing joint survival probability.

0403. BUGLI, C., LAMBERT, Ph., BOULANGER, B., LEDENT, E., PEREIRA, A. and P. NARDONE, Statistical analysis of electroencephalograms : independent component analysis of event-related potentials.

Electroencephalogram (EEG) is an important diagnostic tool in clinical neurophysiology. However, EEGs are not often used in clinical studies because of intrinsic problem like the huge quantity of data or artifacts. In this paper, we shall describe statistical tools to detect and quantify the effect of drugs on the brain by the analysis of EEGs. We first use Independent Component Analysis (ICA) to detect and remove automatically artifacts from EEGs. In the second step, ICA reduces the dimension of the problem. Using data from a clinical trial, we show that eight ICA components can reconstruct more than 80 percents of the data from the twenty-eight electrodes. Some of these eight ICA components can reconstruct an interesting characteristic of the signals (an event–related potential named P300). Finally, we shall show how the analysis of these two components allow to detect and quantify a treatment effect. Lorazepam decreases the P300 peak amplitude and increases the time of occurrence of the P300 peak.

0404. SANCHEZ SELLERO, C., GONZALEZ MANTEIGA, W. and I. VAN KEILEGOM, Uniform representation of product-limit integrals with applications.

Let X be a d-variate random vector that is completely observed, and let Y be a random variable that is subject to right censoring and left truncation. For arbitrary functions φ we consider expectations of the form $E[\varphi(X, Y)]$, which appear in many statistical problems, and we estimate these expectations by using a product–limit estimator for censored and truncated data, extended to the context where covariates are present. An almost sure representation for these estimators is obtained, with a remainder term that is of a certain negligible order, uniformly over a class of $\varphi-$ functions. This uniformity is important for the application to goodness-of-fit testing in regression and to inference for the regression depth, which we consider in more detail. A bootstrap approximation is proposed to approximate the distribution of the product–limit integrals.

0405. VAN KEILEGOM, I., GONZALEZ MANTEIGA, W. and C. SANCHEZ SELLERO, Goodness-of-fit tests in parametric regression based on the estimation of the error distribution.

Consider a heteroscedastic regression model $Y = m(X)+\sigma(X)\varepsilon$, where $m(X) = E(Y|X)$ and $\sigma^2(X) = Var(Y|X)$ are unknown and the error ε is independent of the covariate X. We propose a new type of test statistic for testing whether the regression curve $m(\cdot)$ belongs to some parametric family of regression functions. The proposed test statistic measures the distance between the empirical distribution function of the parametric and of the nonparametric residuals. The asymptotic theory of the proposed test is developed and

14

the proposed testing procedure is illustrated by means of a small simulation study and the analysis of a data set.

0406. BADIN, L. and L. SIMAR, A bias corrected nonparametric envelopment estimator of frontiers.

In productivity analysis, the efficiency scores economic producers are evaluated by measuring their distance toward a production frontier. This frontier is defined as the set of the most efficient alternatives among all possible combinations in input-output space, and it has estimated from a random sample of firms. The nonparametric envelopment estimators rely on the assumption that all the observations fall on the same side of the frontier. Therefore any deviation frontier is due only to inefficiency. The Free Disposal Hull (FDH) estimator of the attainable set, introduced by Deprins, Simar and Tulkens (1984), is the smallest free disposal set covering all the observations. However, by construction, the FDH estimator is an inward–biased estimator of its theoretical correspondent. In this paper we consider the univariate extreme values representation of the FDH estimators and we propose a bias corrected estimator for the efficient frontier or for the efficiency scores, based on order statistics and closely related to FDH. the presentation is based on a probabilistic formulation of the model where the input-output pairs are realizations of independent random variables drawn from a joint distribution whose support is production attainable set. The bias corrected estimator shares the asymptotic properties of the FDH estimator. In finite samples, Monte-Carlo experiments indicate that our corrected estimator reduces significantly not only the bias of the FDH estimator, but also its mean squared error, with computational cost.

0407. PITREBOIS, S., DENUIT, M. and J-F. WALHIN, Bonus-malus scales in segmented tariffs: Gilde & Sundt's work revisited

In a recent paper, the authors have shown how to compute analytically the Bayesian relativities for a Bonus-Malus scale superimposed on a segmented tariff. However, the percentages associated with the levels of such a scale may exhibit an abrupt rise or drop from one level to another. For commercial reasons, it is sometimes preferable that each level inflicts the same relative penalty. In this paper we will show how to obtain such linear relativities for a given Bonus-Malus System. The model allows for a priori ratemaking. It is applied to a real-life portfolio.

0408. DENUIT, M., GENEST, C. and M. MESFIOUI, Stop-loss bounds on functions of possibly dependent risks in the presence of partial information on their marginals

A general strategy is proposed for deriving upper and lower stop-loss bounds on directionally convex, increasing functions of possibly dependent, non–negative random variables when partial information is available about their joint copula or their marginal distributions. Illustrations are presented which involve looking for a stop-loss bound on a sum of insurance risks whose means, variances or range are finite and known. The key to these developments, which generalize recent findings of Genest, Marceau and Mesfioui (2002), is a result due to Muller and Scarsini (2001) on the stochastic comparison of random vectors having a common copula.

0409. MAGIS, C., DENUIT, M. and J-F. WALHIN, Une proposition de tables prospectives pour le marché belge des rentes

Nous analysons la problématique des tables de mortalité prospectives pour le marché des rentes viagères. Nous déduisons une proposition de mise à jour des tables de mortalité à retenir en vue de la tarification et du provisionnement rentes viagères.

0410. MAGIS, C., DENUIT, M. and J-F. WALHIN, La mortalité, un phénomène en pleine mutation: quelle solution pour le marché des rentes?

Les progrès de la médecine ainsi que l'amélioration du niveau de vie dans nos pays industrialisés provoquent une diminution de la mortalité. Dans un tel contexte, il est nécessaire de développer des tables de mortalité prospectives, prenant en compte cette constante évolution. Sans quoi, dans les prochaines années, de graves déficits budgétaires surviendront dans la plupart des compagnies d'assurance vie offrant des rentes viagères. Dans cet article, nous analysons l'établissement de telles tables.

0411. MAGIS, C., DENUIT, M. and J-F. WALHIN, La TPRV française: dépassée?

Les progrès de la médecine ainsi que l'amélioration du niveau de vie dans nos pays industrialisés provoquent une diminution de la mortalité. Dans un tel contexte, il est nécessaire de développer des tables de mortalité prospectives, prenant en compte cette constante évolution. Celles-ci deviennent rapidement lourdes à utiliser. En effet, les assureurs sont alors contraints de travailler avec des tables comportant deux entrées: l'âge x de l'assuré et sa génération g, ce qui implique une gestion assez fastidieuse. Les assureurs préfèrent dès lors disposer d'une base de calcul moins volumineuse que la collection de tables prospectives (une par génération et sexe). C'est pourquoi nous nous intéressons dans cet article au développement d'une table du type de la TPRV (Table Pour Rentes Viagères) qui est d'application actuellement en France.

0412. DARAIO, C. and L. SIMAR, A robust nonparametric approach to evaluate and explain the performance of mutual funds

The topic of the measurement of mutual funds' performance is receiving an increasing interest both from an applied and a theoretical perspective. Beside the traditional financial literature, a growing body of works has started to apply the tools of frontier analysis for benchmarking comparisons in portfolio analysis. Our paper contributes to this literature proposing a robust nonparametric approach for analysing mutual funds. It is based on the concept of order m frontier (Cazals, Florens and Simar, 2002) and on a probabilistic approach (Daraio and Simar, 2003) to find out the factors explaining mutual funds' performance. The usefulness of this approach is illustrated by using US mutual funds data, grouped for category by objective. Economies of scale, slacks and market risks are investigated. A comparison of traditional, nonparametric and robust performance measures is also offered.

0413. WANG, L. and I. VAN KEILEGOM, Nonparametric test for the form of parametric regression with time series errors

We propose a new nonparametric method for testing the parametric form of a regression function in the presence of time series errors. The nonparametric test is motivated by recent advancement in the theory of ANOVA with large number of factor levels and also utilizes a new difference-based estimation method in non–parametric regression with time–series errors proposed by Hall and Van Keilegom (2003). The test statistic is asymptotically normal under the null and local alternative hypotheses. We also propose a bootstrap method to calculate the critical values and prove its consistency. In a Monte Carlo study, we demonstrate that this bootstrap procedure has good properties for moderate sample size.

0414. HJORT, N.L., McKEAGUE, I.W. and I. VAN KEILEGOM, Extending the scope of empirical likelihood

This paper extends the scope of empirical likelihood methodology in three directions: to allow for plug–in estimates of nuisance parameters in estimating equations, slower than $\sqrt{n}$-rates of convergence, and settings in which there are a relatively large number of estimating equations compared to the sample size. Calibrating empirical likelihood confidence regions with plug-in is sometimes intractable due to the complexity of the asymptotics, so we introduce a bootstrap approximation that can be used in such situations. We provide a range of examples from survival analysis and nonparametric statistics to illustrate the main results.

0415. BROUHNS, N., DENUIT, M. and I. VAN KEILEGOM, Bootstrapping the Poisson log–bilinear model for mortality forecasting

This paper proposes bootstrap procedures for expected remaining lifetimes and life annuity single premiums in a dynamic mortality environment. Assuming a further continuation of the stable pace of mortality decline, a Poisson log–bilinear projection model is applied to the forecasting of the gender– and age–specific mortality rates for Belgium on the basis of mortality statistics relating to the period 1950–2000. Bootstrap procedures are then used to obtain confidence intervals on various actuarial quantities.

0416. PARDO-FERNANDEZ, J.C., VAN KEILEGOM, I. and W. GONZALEZ-MANTEIGA, Comparison of regression curves based on the estimation of the error distribution

Assume that (X_1, Y_1) and (X_2, Y_2) are two independent random vectors, that satisfy the non–parametric regression models $Y_j = m_j(X_j) + \sigma_j(X_j)\varepsilon_j$, for $j = 1, 2$, where $m_j(X_j) = E(Y_j|X_j)$ and $\sigma_j^2(X_j) = Var(Y_j|X_j)$ are smooth but unknown regression and variance functions respectively, and the error variable ε_j is independent of X_j. Suppose that ε_1 and ε_2 have the same distribution. In this paper we introduce a procedure to test the hypothesis $H_0 : m_1(\Delta) = m_2(\Delta)$ of equality of the regression functions. The test is based on the comparison of the empirical distribution of the errors in each model. A Kolmogorov–Smirnov type statistic and a Cramér–von Mises type statistic are considered, and their asymptotic distribution is obtained. The proposed tests can detect local alternatives converging to the null hypothesis at the rate n studied. An application to a real data set is included.

0417. KUMBHAKAR, S.C., PARK, B.U., SIMAR, L. and E.G. TSIONAS, Non-parametric stochastic frontiers: a local maximum likelihood approach

This paper proposes a nonparametric approach for stochastic frontier (SF) models based on local maximum likelihood techniques. The SF model is presented as encompassing some anchorage parametric model in a nonparametric way. First, we derive asymptotic properties of the estimator for the general case (local linear approximations). Then the results are tailored to a SF model where the convoluted error term (efficiency plus noise) is the sum of an half normal and a normal random variable. The parametric anchorage model is a linear production function and an homoscedastic error term. The local approximation is thus local linear for the production function and local constant for the parameters of the error terms. The performance of our estimator is first established with a simulated data set and then with real data on milk production in Spanish dairy farm. The methods appear to be robust, numerically stable and particularly useful for investigating a production process and the derived efficiency scores.

0418. LE BAILLY DE TILLEGHEM, C., BECK, B., BOULANGER, B. and B. GOVAERTS, A fast exchange algorithm for designing focused libraries in lead optimisation

Combinatorial chemistry is widely used in drug discovery. Once a lead compound has been identified, a series of R–groups and reagents can be selected and combined to generate new potential drugs. The combinatorial nature of this problem leads to chemical libraries containing usually a very large number of virtual compounds, far too large to permit their chemical synthesis. Therefore, one often wants to select a subset of ”good” reagents for each R–group of reagents and synthesise all their possible combinations. In this research, one encounters some difficulties. First, the selection of reagents has to be done such that the compounds of the resulting sub–library simultaneously optimise a series of chemical properties. For each compound, we use a desirability index, a concept proposed by Harrington [20], to summarise those properties in one fitness value. Then a loss function is used as objective criteria to globally quantify the quality of a sub–library. Secondly, there are a huge number of possible sub–libraries and the solutions space has to be explored as fast as possible. The WEALD algorithm proposed in this paper starts with a random solution and iterates by applying exchanges, a simple method proposed by Federov [13] and often used in the generation of optimal designs. Those exchanges are guided by a weighting of the reagents adapted recursively as the solutions space is explored. The algorithm is applied on a real database and reveals to converge rapidly. It is compared to results given by two other algorithms presented in the combinatorial chemistry literature: the Piccolo algorithm of W. Zheng et al. [37] and the Ultrafast algorithm of D. Agrafiotis and V. Lobanov [4].

0419. DAOUIA, A. and L. SIMAR, Nonparametric efficiency analysis: a multivariate conditional quantile approach

This paper proposes a unifying probabilistic framework efficiency and productivity analysis in a complete multivariate setup (multiple inputs and multiple outputs). Properties of Farrell's efficiency scores are derived in terms of the characteristics of the probability distribution of data generating processs.This allows to introduce a notion of α–quantile efficiency scores related to a non–standard conditional α–quantile frontier. Nonparametric estimators are then naturally introduced providing estimators of the production frontier more robust to outliers and/or extreme values than traditional envelopment estimators (FDH/DEA). The paper defines a new concept efficiency measurement, analyzes its properties and proposes a nonparametric estimator with all its asymptotic properties. Numerical examples (simulated data mutual funds data) illustrate its practical use. This extends and generalizes previous works of Cazals, Florens and Simar (2002), Aragon, Daouia and Thomas–Agnan (2002) and Daraio and Simar (2003).

0420. DELAIGLE, A. and I. GIJBELS, Data-driven boundary estimation in deconvolution problems

This paper deals with estimation of the support of a density function when only a contaminated sample from the density is available. Delaigle and Gijbels (2003) introduced a kernel-based method and studied theoretical bias and variance of the estimator. the present paper deals with the practical implementation issues of this method, and is a necessary supplement ot the theoretical results to get to a data-driven method that is widely applicable. We propose two such practical data-driven procedures, and show via simulations that they perform well for a wide variety of densities (including quite difficult cases.) The methods can also be applied for error-free data and as such also present data-driven procedures for estimation of boundaries in the case of non-contaminated data. Moreover they can be applied for estimating discontinuities of a density, as is shown. We illustrate the use of the proposed data-driven boundary estimation procedures in frontier.

0421. MOUCHART, M. and M. VANDRESSE, Bargaining powers and market segmentation in freight transport

Freight transport market is characterized by heterogeneities of the provided services and of the firms. Moreover, agents on this market typically bargain simultaneously on a price–attributes basis rather than on a price–quantity basis. The presence of heterogeneity leads to develop an empirical model of market imperfection and the presence of a bargaining based simultaneously on the price and the attributes of the transport leads to develop a model without exogeneity assumption. As a consequence, the proposed model is based on the estimation of the joint distribution of the price and the attributes of the transport rather than on the expectation of the price conditional to the attributes. The model proposes an empirical measure of market imperfection and bargaining powers of the agents and examines the sensitivity of the results to the choice of the attributes and, in order to detect potential market segmentation, to the selection of the data.

0422. PURCARU, O., GUILLEN, M. and M. DENUIT, Linear credibility models based on time series for claim counts

In this paper, several models for longitudinal analysis of claim frequencies are presented. These models are fitted to a large panel data set relating to a major Spanish insurance company. Credibility theory is then used to update the individual claim frequencies. The results are compared and the consequencies of increasing the strength of dependence are clearly visible.

0423. DENUIT, M., PURCARU, O. and I. VAN KEILEGOM, Bivariate Archimedean copula modelling for loss-ALAE data in non-life insurance

This paper considers the bivariate loss–ALAE modelling problem in actuarial science, taking into account the particular form of the censorship present in

the data. Specifically, a graphical selection procedure for the generator of the underlying archimedean copula is proposed, based on a nonparametric estimation of the generator. The approach is in line with Genest & Rivest (1993) who considered complete data and Wang & Wells (2000b) who treated doubly censored data. A loss–ALAE data set provided by the US Insurance Services Office is used for the numerical illustrations, and comparisons with previous results appeared in the actuarial literature are performed.

0424. JEONG, S.-O. and B.U. PARK, Limit distribution of convex-hull estimators of boundaries

Given n independent and identically distributed observations in a set $G = (x, y) \in [0, 1]_p \times \mathbb{R} : 0 \leq y \leq g(x)$ with an unknown function g, called a boundary or frontier, it is desired to estimate g from the observations. The problem has several important applications including classification and cluster analysis, and is closely related to edge estimation in image reconstruction. It is particularly important in econometrics. The convex–hull estimator of a boundary or frontier is very popular in econometrics, where it is a cornerstone of a method known as 'data envelope analysis' or DEA. In this paper we give a large sample approximation of the distribution of the convex–hull estimator in the general case where $p \geq 1$. We discuss ways of using the large sample approximation to correct the bias of the convex–hull and the DEA estimators and to construct confidence intervals for the true function.

0425. JEONG, S.-O., Asymptotic distribution of DEA efficiency scores

Data envelopment analysis (DEA) estimators have been widely used in productivity analysis. The asymptotic distribution of DEA estimator derived by Kneip, Simar and Wilson (2003) is too complicated and abstract for analysts to use in practice, though it should be appreciated in its own right. This paper provides another way to express the limit distribution of the DEA estimator in a tractable way.

0426. BONACCORSI, A., DARAIO, C. and L. SIMAR, Advanced indicators of productivity of universities. An application of robust nonparametric methods to italian data

In recent years the pressure on public budgets in almost all industrialised countries has lead governments to pursue efficiency in the allocation and management of resources trying to apply to scientific research and higher education some fundamental ideas of the economic analysis such as the concepts of economies of scale and scope. This paper explores scales, scope and trade-off effects in scientific research and education. Advanced productivity methods are used to analyse the Etalian system of universities. In particular, robust methods based on the concept of order-m frontiers (Cazals,Florens and Simar, 2002)are really useful in this framework for their properties of not being influenced by extremes and noise in the data. Furthermore, in the field

21

of science and education, external factors and enviromental conditions may because of heterogeneity and influence dramatically the performance of universities. Hence, we apply the Daraio and Simar (2003) full nonparametric methodology (that overcomes most limitations of previous one or two-stage approaches) to robustly take into account external-environmental factors. From a preliminary investigaton on Italina data we find that economies of scale and scope are not significant factors in explaining research and education productivity. We do not find any evidence of the trade-off research vs teaching: on the contrary, increasing scientific quality improves educational efficiency;on the other hand, a good educational efficiency does not deteriorate research efficiency. About the trade-off academic publications vs industry oriented research, local effects of a complementarity/rivalry relation seem to emerge: it seems that initially, collaboration with industry may improve productivity, but beyond a certain level the compliance with industry expectations may be too demanding and deterirate the publication profile. Nevertheless, the existence of an inverted U-shaped relation should be confirmed by more evidence. Advanced robust methods in efficiency analysis are shown as useful tools for measuring and explaining the performance of a public research system of universities. Further developments of the analysis are outlined.

0427. CAO, R. and I. VAN KEILEGOM, Empirical likelihood tests for two-sample problems via nonparametric density estimation

In this paper we study the problem of testing whether two populations have the same law, by comparing kernel estimators of the two density functions. The proposed test statistic is based on a local empirical likelihood approach. We obtain the asymptotic distribution of the test statistic and propose a bootstrap approximation to calibrate the test. A simulation study is carried out, in which the proposed method is compared with two competitors, and a procedure to select the bandwidth parameter is studied. The proposed test can be extended to more than two samples and to multivariate distributions.

0428. BUGLI, C. and Ph. LAMBERT, Functional ANOVA with random functional effects: an application to event-related potentials modelling for electroencephalograms analysis

The differential effects of basic visual auditory stimuli on electroencephalograms (EEG), named event related potentials (ERPs), are evaluate impact treatments brain performances. In the present paper, we propose a P–splines based model that can be used to evaluate treatment effect on the timing and the amplitude of some peaks of the ERPs curves. Functional ANOVA is an adaptation linear model or analysis of variance to analyse functional observations. The changes functional interest effects are generally described using smoothing splines. Eilers and Marx (1996) proposed to work with P–splines, a combination of B–splines and difference penalties on coefficients. We define a P–splines model for ERPs curves combined with random effects. In

particular, we show that it is useful alternative to classical strategies requiring the visual and usually imprecise localization of specific ERP peaks from curves with a low signal–to–noise ratio.

0429. ZELENYUK, V. and V. ZHEKA, Corporate governance and firm's efficiency: the case of a transitional country, Ukraine

In this study, we lool for empirical support for the hypothesis that there is a positive relationship between the levels of corporate governance quality across firms and the relative efficiency levels of these firms. This hypothesis is related to Liebenstein's idea of X-efficiency. We use the data envelopment analysis (DEA) estimator to obtain proxies for X-(in)efficiency of firms in our sample and then analyze them with respect to different ownership structures by comparing distributions and aggregate efficiencies across different groups. We also use truncated regression with bootstrap, following Simar and Wilson (2003), to infer on relationship of inefficiency to various indicators of quality of corporate governance, ownership variables, as well as industry and year specific dummies. The data is coming from 7 industries in Ukraine.

0430. FÄRE, R. and V. ZELENYUK, On aggregation of scale elasticities across firms

In this paper we propose new aggregate or 'group' primal and dual scale elasticity measures of an economic system (e.g.,industry). The new aggregate measures are the weighted averages of individual scale elasticities. Remarkably, the aggregation function and the aggregation weights are not ad hoc but derived from economic theory-duality theory of Shephard (1953) and the aggregation theory of Koopmans (1957), Fare, Grosskopf

0431. HENDERSON, D. J. and V. ZELENYUK, Testing for catching-up: statistical analysis of DEA efficiency estimates

We use advances in DEA techniques to examine efficiency scores and investigate the issue of convergence/divergence in a cross-country analysis. Specifically, we take use of bootstrapping techniques to examine a data set of 52 developed and developing countries. We find that when using the standard DEA model the results are sometimes less than desirable. Further, we break the sample into the two groups to examine the two-club convergence phehomenon. We find that efficiency scores are significantly different between the two groups and that there is some evidence of convergence of efficiency scores within each group.

0432. LAMBERT, Ph. and P. H. C. EILERS, Bayesian survival models with smooth time-varying coefficients using penalized Poisson regression

It is rather well known that one can approach survival problems without covariates in an actuarial way. The time axis is divided into intervals (named bins), and in each bin the number of people at risk is counted as well as the

number of events. the relationship between time and probability of an event can then be estimated with a parametric or semi-parametric model. Here, we consider a subdivision of the time scale into a large number of bins. The number of events observed in each bin is described using a Poison distribution with the log mean specified using a flexible penalized B-splines model with knots located at the bins limits. Regression on pertinent covariates can easily be performed using the same log-linear model, leading to the classical proportional hazard model. We propose to extend that model by allowing the regression coefficients to vary in a smooth way with time. Penalized B-splines models will be proposed for each of these coefficients. We show how the regression parameters and the penalty weights can be estimated efficiently using Bayesian inference tools based on the Metropolis-adjusted Lagevin algorithm.

0433. NGUTI, R., CLAESKENS, G. and P. JANSSEN, One-sided tests in shared frailty models

Tests for the presence of heterogeneity in frailty models use an alternative hypothesis in which the heterogeneity parameter is subject to an inequaltiy constraint. As a result the classical likelihood ratio asymptotic chi-squared distributed random variables. The results are shown to hold when the data might be subject to right censoring and when covariate information is present. A data example on a diabetic retinopathy study illustrates the tests.

0434. SIMAR, L. and V. ZELENYUK, On testing equality of distributions of technical efficiency scores

The challenge of the econometric problem in production efficiency analysis is that the very efficiency scores to be analyzed are unobserved. Recently, statistical properties have been discovered for a class of estimators popular in the literature, known as data envelopment analysis (DEA) approach. This opens a wide range of possibilities for a well-grounded statistical inference about the true efficiency scores from their DEA-estimates. In this paper we investigate possibility of using existing tests for equality of two distributions for such a context. Considering statistical complications pertinent to our context, we consider several approaches to adapt the Li (1996) test to the context and explore their performance in terms of the size and the power of the test in various Monte Carlo experiments. One of these approaches showed good performance both in the size and in the power, thus encouraging for its wide use in empirical studies.

4.2   Consulting Reports

CR0401 GOVAERTS, Bernadette, "Recherche de conditions optimales de brasage Ni-Au de stators en inconel 718".

CR0402 MOUCHART, Michel and Jeroen ROMBOUTS, "Evaluating and updating econometric models for nowcasts on R& D variables: 2004 results".

CR0405 HOEFFELMAN, J., DECAT, G., LILIEN, J.L., DELAIGLE, A. and B. GOVAERTS, "Assessment of electric and magnetic field levels in the vicinity of the HV overhead power lines in Belgium".

4.3   Published Papers

214. DELAIGLE, A. and I. GIJBELS. Practical bandwidth selection in deconvolution kernel density estimation. *Computational Statistics and Data Analysis*, 45, 249-267, 2004.

219. DELAIGLE, A. and I. GIJBELS. Bootstrap bandwidth selection in kernel density estimation from a contaminated sample. *Ann. Inst. Statist. Math.*, 56(1), 19-47, 2004.

220. GIJBELS, I. and A.-C. GODERNIAUX. Data-driven discontinuity detection in derivatives of a regression function. *Communication in Statistics*, 33, 4, 851-871, 2004.

221. GIJBELS, I., HALL, P. and A. KNEIP. Interval and band estimation for curves with jumps. *Journal of Applied Probability, Special Volume 41A, "Papers in honour of Chris Heyde"*, 65-79, 2004.

222. BEGUIN, C. and L. SIMAR. Analysis of the expenses linked to hospital stays: how to detect outliers. *Health Care Management Science*, 7, 89-96, 2004.

223. GIJBELS, I. and A.-C. GODERNIAUX. Bandwidth selection for change-point estimation in nonparametric regression. *Technometrics*, 46, 1, 76-86, 2004.

224. SIMAR, L. and P.W. WILSON. Performance of the bootstrap for DEA estimators and iterating the principle. *International Series in Operations research & Management Science: Handbook On Data Envelopment Analysis edited by W.W. Cooper, L.M. Seiford and J. Zhu*, 10, 265-298, 2004.

225. GIJBELS, I. and N. HECKMAN. Nonparametric testing for a monotone hazard function via normalized spacings. *Journal of Nonparametric Statistics*, 16(3-4), 463-477, 2004.

226. GIJBELS, I. and A.-C. GODERNIAUX. Bootstrap test for change-points in nonparametric regression. *Journal of Nonparametric Statistics*, 16(3-4), 591-611, 2004.

227. LAMBERT, Ph., COLLETT, D., KIMBER, A. and R. JOHNSON. Parametric accelerated failure time models with random effects and an application to kidney transplant survival. *Statistics in Medicine*, 23, 3177-3192, 2004.

228. BOUEZMARNI, T. and J.-M. ROLIN. Consistency of the beta kernel density function estimator. *The Canadian Journal of Statistics*, 31, 1, 89-98, 2003.

229. VAN KEILEGOM, I. A note on the nonparametric estimation of the bivariate distribution under dependent censoring. *Nonparametric Statistics*, 16, 3-4, 659-670, 2004.

230. FLORENS, J.-P. and L. SIMAR. Parametric approximations of nonparametric frontiers. *Journal of Econometrics*, 124, 91-116, 2005.

231. VAN BELLEGEM, S. and R. von SACHS. On adaptive estimation for locally stationary wavelet processes and its applications. *International Journal of Wavelets, Multiresolution and Information Processing*, 2, 4, 545-565, 2004.

232. VAN BELLEGEM, S. and R. von SACHS. Forecasting economic time series with unconditional time-varying variance. *International Journal of Forecasting*, 20, 611-627, 2004.

233. FRANCOIS, N., GOVAERTS, B. and B. BOULANGER. Optimal designs for inverse prediction in univariate nonlinear calibration models. *Chemometrics and Intelligent Laboratory Systems*, 74, 283-292, 2004.

4.4   Books published by members of the Institute

FLORENS, J.P., MOUCHART, M. and J.M. ROLIN. *Elements of Bayesian Statistics*, 544 pp, Marcel Dekker: New York, 1990.

HÄRDLE, W. and L. SIMAR (editors). *Computer Intensive Methods in Statistics*, 175 pp, *Statistics and Computing, I*, Physica-Verlag: Berlin, 1993.

HÄRDLE, W., KLINKE, S. and B.A. TURLACH. *XploRe: An Interactive Statistical Computing Environment*, 387 pp, *Statistics and Computing*, Springer-Verlag: New York, 1995.

FAN, J. and I. GIJBELS. *Local Polynomial Modelling and its Applications*, 341 pp, Chapman and Hall: London, 1996.

KAAS, R., GOOVAERTS, M.J., DHAENE, J., and M. DENUIT. *Modern Actuarial Risk Theory*, Kluwer Academic Publishers: Dordrecht, 2001.

WUNSCH, G., MOUCHART, M. and J. DUCHÊNE (editors). *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. *Applied Multivariate Statistical Analysis*, 486 pp., Springer Verlag: Berlin, 2003.

4.5 Editing activities

Michel DENUIT

> Proceedings Editor for *Insurance: Mathematics and Economics*
> Editor of *Belgian Actuarial Bulletin*
> Associate Editor Australian and New-Zeeland Journal of Statistics
> Member of the Advisory Board of the Wiley Encyclopedia of Actuarial
> Science.

Irène GIJBELS

> Associate Editor of *The Annals of Statistics* since January 2004.
> Associate Editor of *Journal of Computational and Graphical Statistics*
> since October 2000.
> Associate Editor of *Statistica Sinica* since August 2002.

Philippe LAMBERT

> Co-editor of *B-Stat News.*

Léopold SIMAR

> Associate Editor of *Journal of Productivity Analysis*

Ingrid VAN KEILEGOM

> Associate Editor of *Scandinavian Journal of Statistics*
> Associate Editor of *Journal of the Royal Statistical Society - Series B*
> Extramural fellow of CentER, Faculty of Economics and Business Administration, University of Tilburg, The Netherlands.

Rainer von SACHS

> Associate Editor of *Journal of the Royal Statistical Society - Series B*

5 SEMINARS, WORKSHOPS AND SHORT COURSES

A statistics seminar is organised each week. A diversity of subjects is presented at this seminar. Speakers are mainly coming from outside the University and visitors of the Institute are among the contributors.

From time to time, a joint statistics and econometrics seminar, organised in collaboration with CORE, takes place. At those occasions statisticians and econometricians meet and have extra opportunities to discuss on common research interests and elaborate joint research.

Further, an applied statistics workshop is organised by the Institute on a regular basis. At this applied statistics workshop, emphasis is on talks in which an applied statistical problem is presented.

There is also the doctoral seminar which is an extra stimulant for Ph.D students and other young researchers.

5.1 Statistics Seminars

1. January 30, 2004, Gerda CLAESKENS, Institut de statistique, UCL, Belgium.
   Bayesian-motivated tests of function fit and their asymptotic frequentist properties

2. February 6, 2004, Javier HIDALGO, London School of Economics, United Kingdom.
   Bootstrap test for breaks of a regression model with dependent data

3. February 13, 2004, F. Thomas BRUSS, Université Libre de Bruxelles, Belgium.
   The odds-theroem of optimal stopping and its descendants

4. February 13, 2004, Geurt JONGBLOED, Vrije Universiteit, Amsterdam, The Netherlands.
   Maximum likelihood estimation in deconvolution problems

5. February 27, 2004, Pierre DEVOLDER, Université catholique de Louvain, Belgium.
   Stochastic optimal control of insurance liabilities

6. March 19, 2004, Wim SCHOUTENS, Katholieke Universiteit Leuven, Belgium.
   Model risk for Exotic options

7. March 26, 2004, Fabian TIBALDI, Limburgs Universitair Centrum, Diepenbeek, Belgium.

Pseudo-likelihood estimation and inferences for a marginal multivariate survival model

8. April 2, 2004, Gilles FAY, Université des Sciences et Technologies de Lille, France.
Développements d'Edgeworth et processus linéaires

9. May 7, 2004, Grayham E. MIZON, University of Southampton, United Kingdom.
Congruence and Encompassing

10. September 24, 2004, Holger DREES, University of Hamburg, Germany.
Bivariate tail estimation: dependence in asymptotic independence

11. October 1st, 2004, Marc AERTS, Limburgs Universitair Centrum, Belgium.
Model selection based on incomplete and nonrandom samples

12. October 15, 2004, Eckhard LIEBSCHER, Technical University Ilmenau, Germany.
Mixing conditions and nonparametric smoothing estimators

13. October 22, 2004, Daniele ROMANO, University of Cagliari, Italy.
Computer experiments and robust design for advanced engineering design

14. November 5, 2004, Léopold SIMAR, Institut de statistique, UCL, Belgium.
Nonparametric efficiency analysis: a multivariate conditional quantile approach

15. November 19, 2004, Cesar SANCHEZ SELLERO, University of Santiago de Compostela, Spain.
The two-sample problem with truncated and censored data

16. November 26, 2004, Ioana POPESCU, INSEAD.
Robust mean-covariance solutions to stochastic optimization problems

17. December 3, 2004, Marco SAERENS, Université catholique de Louvain, Belgium.
The principal components analysis of a graph: application to a collaborative recommendation task

18. December 10, 2004, Ingrid VAN KEILEGOM, Institut de statistique, UCL, Belgium.
Empirical likelihood tests for two-sample problems via nonparametric density estimation

5.2 Joint Statistics and Econometrics Seminars

1. September 21, 2004, Werner HILDENBRAND, Univeristy of Bonn, Germany and Alois KNEIP, University of Mainz, Germany.
   Aggregate behavior and microdata

2. November 12, 2004, Michel MOUCHART, Institut de statistique, UCL, Belgium.
   Bargaining powers and market segmentation in freight transport

5.3 Joint Statistics Seminars and Institut de Mathématique Pure et Appliquée

1. February 20, 2004, Holger DETTE, Ruhr University, Germany.
   On a test for a parametric form of volatility in continuous time financial models

2. March 5, 2004, Paul WILSON, University of Texas, USA.
   Robust nonparametric estimation of efficiency and technical change in U. S. commercial banking

5.4 Joint Statistics Seminars and IAP network

1. October 29, 2004, Seok-Oh JEONG, Institut de statistique, UCL, Belgium.
   Linearly interpolated FDH estimator for nonconvex frontiers

2. December 17, 2004, Sébastien VAN BELLEGEM, Institut de statistique, UCL, Belgium.
   Semiparametric estimation by model selection for locally stationary processes

3. December 17, 2004, Tom SNIJDERS, University of Groningen, The Netherlands.
   Frequentist MCMC estimation methods for multilevel logistic regression

5.5 Applied Statistics Workshops

1. February 6, 2004, Eric LE BOULENGE, Unité d'écologie et de biogéographie, UCL, Belgium.
   Analysis of biological data on beetle populations structure that violate most "classical" requirements: Ignore the warnings and go ahead ?

2. March 5, 2004, Pierre DARDENNE, Centre wallon de recherches agronomiques, Département qualité des productions agricoles, Gembloux, Belgium.
   Application de SVM (Support Vector Machine) pour la discrimination des particules végétales et animales en imagerie spectrale proche infrarouge

3. March 12, 2004, Maria LANZERATH, W. L. Gore, Putzbrunn, Germany.
   How to stick the membrane to the fabric ? A case for fractional factorial experiment in manufacturing

4. March 26, 2004, Daniela COCCHI, University of Bologna, Italy.
   A hierarchical model for the analysis of local census undercount in Italy

5. April 23, 2004, Didier RENARD, Eli Lilly & Company, Mont-Saint-Guibert, Belgium.
   Pairwise likelihood estimation in multivlevel probit models

6. September 30, 2004, Christian RITTER, Institut de statistique, UCL, Belgium.
   The golden triangle of statistical analysis: combining the powers of a data base, a statistical language and a spreadsheet

7. October 7, 2004, Christian RITTER, Institut de statistique, UCL, Belgium.
   On the presentation of data

8. October 28, 2004, Pierre PESTIAUX, Centre de Recherche de Gonfreville, Groupe Mathématiques Appliquées et Statistiques, Total France.
   Tests inter-laboratoires: problématiaue générale et traitement de valeurs aberrantes

9. November 25, 2004, Dimitri D'OR, FSS Internation r& d.
   Géostatistiques et sites pollués

10. December 9, 2004, Walthere DEWE, Ely Lilly, Mont-Saint-Guibert, Belgium.
    Towards a better use of statistics in method validation and method transfer

5.6   Doctoral Seminars

1. January 30, 2004, Astrid JULLION, Institut de statistique, UCL, Belgium.
   Two non-parametric methods to quantify uncertainty in binding potential and receptor occupancy estimation

2. February 20, 2004, Alexandre LAMBERT, Institut de statistique, UCL, Belgium.
   Jump-preserving regression and smoothing using local linear fitting: a compromise

3. February 27, 2004, Nancy FRANCOIS, Institut de statistique, UCL, Belgium.
   Time-intensity curves in sensory analysis: Statistical analysis and surroundings !

4. March 12, 2004, Juan Carlos PARDO-FERNANDEZ, University of Vigo, Spain.
   Comparison of regression curves based on the errors distribution

5. March 19, 2004, Céline BUGLI, Institut de statistique, UCL, Belgium.
   Flexible modeling of event-related potentials: functional ANOVA with P-splines

6. April 2, 2004, Maria Amalia JACOME PUMAR, University of Vigo, Spain.
   Presmoothed estimation for censored data

7. April 23, 2004, Antoine DELWARDE, Institut de statistique, UCL, Belgium.
   Mortality projection: Lee-Carter model and extensions

8. October 14, 2004, Hilmar BÖHM, Institut de statistique, UCL, Belgium.
   Model selection in high dimensional nonstationary time series

9. October 21, 2004, Juan Carlos PARDO FERNANDEZ, University of Vigo, Spain.
   Comparison of regression curves with censored responses

10. November 18, 2004, Gery GEENENS, Institut de statistique, UCL, Belgium.
    Index coefficient in single-index models: the genralized maximum rank correlation estimator

11. November 25, 2004, Bianca TEODORESCU, Institut de statistique, UCL, Belgium.
    Generalized linear conditional models under left truncation and right censoring

12. December 16, 2004, Giovanni MOTTA, Institut de statistique, UCL, Belgium.
    Nonparametric estimation and hypothesis testing for the time-varying covariance matrix of multivariate nonstationary time series

13. December 23, 2004, Anouar EL GHOUCH, Institut de statistique, UCL, Belgium.
    Confidence intervals for Kaplan-Meier integrals under dependent data

5.7 Short courses

During the spring 2004, short courses were given by invited professors (see section 2. of this report) within the activities of the Graduate School in Statistics :

- Holger DETTE, Ruhr-Universität Bochum, Germany (Frebruary, March 2004).
  "Design of experiments"

- Peter HALL, Australian National University, Canberra (December 2004).
  "Advanced bootstrap methods"

6 DOCTORATES

6.1 Doctors honoris causa

Peter HALL, Australian National University, Canberra (1997)

Luc DEVROYE, McGill University, Montréal, Canada (2002)

6.2 Doctoral dissertations

Taoufik BOUEZMARNI (November 19, 2004)

> "Smoothed histograms and asymmetric kernel estimation for density functions"
> Promoter: Jean-Marie ROLIN

Isabelle DE MACQ (May 7, 2004)

> "Classification trees and methods based on projections"
> Promoter: Léopold SIMAR

6.3 Ph.D thesis in progress

Carlos ALMEIDA

> "Structural equation modeling with categorical ordered variables"
> Promoter: Michel MOUCHART

Hilmar BÖHM

> "Model selection in principal component analysis for multivariate time series"
> Promoter: Rainer von SACHS

Jean-Philippe BOUCHER

> "Risk classification in nonlife insurance"
> Promoter: Michel DENUIT

Natacha BROUHNS

"An integrated ratemaking tool for life and non life insurance"
Promoter: Michel DENUIT

Céline BUGLI

"Statistical analysis of evoked potentials in electroencephalograms"
Promoter: Philippe LAMBERT

Cindy COURTOIS

"Risk theory under partial information and dependence"
Promoter: Michel DENUIT

Alexandra DASKOVSKA

"Dynamical analysis of productivity index"
Promoter:Léopold SIMAR and Sébastien VAN BELLEGEM

Antoine DELWARDE

"Modern statistical methods for insurance pricing"
Promoter: Michel DENUIT

Anouar EL GHOUCH

"Empirical likelihood with incomplete data"
Promoter: Ingrid VAN KEILEGOM

Nancy FRANCOIS

"Statistical analysis of time intensity curves in sensory analysis"
Co-Promoters: Bernadette GOVAERTS and Philippe LAMBERT

Gery GEENENS

"Testing for conditional independence in contingency tables"
Promoter: Léopold SIMAR

Cédric HEUCHENNE

"Mean preservation in censored regression using preliminary nonparametric smoothing"
Promoter: Ingrid VAN KEILEGOM

Astrid JULLION

"Statistical analysis of pet scan data"
Promoter: Philippe LAMBERT

Maria KEY PRATO

"Detection and quantification of treatment effect on blood pressure profile curves"
Promoter: Philippe LAMBERT

Alexandre LAMBERT

"Nonparametric estimation of discontinuous curves and surfaces"
Promoter: Irène GIJBELS

Céline LE BAILLY DE TILLEGHEM

"Multiobjective optimization by computer simulations"
Co-Promoters: Bernadette GOVAERTS and Léopold SIMAR

Giovanni MOTTA

"Nonparametric estimation and hypothesis testing for the time-varying covariance matrix of multivariate nonstationary time series"
Promoter: Rainer von SACHS

S. PITREBOIS

"Experience rating in nonlife insurance"
Promoter: Michel DENUIT and Jean-François WALHIN

Oana PURCARU

“Modeling dependence in actuarial science"
Promoter: Michel DENUIT

Franck QUASHIE

“Life insurance with stochastic forces of mortality"
Promoter: Michel DENUIT

Réjane ROUSSEAU

“Outils statistiquees pour identification de biomarqueurs de toxicité
métabonomiques"
Promoter: Bernadette GOVAERTS

Bianca TEODORESCU

“Generalized linear conditional models under left truncation and right
censoring"
Promoter: Ingrid VAN KEILEGOM

7 CONTACT DAYS AND CONFERENCES

7.1 FNRS contact day : B-ENBIS Workshop on : Statistical tools for sensory analysis: A methodological and applied tour

Louvain-la-Neuve, Belgium, May 14, 2004

Group of Statistical Analysis

Organisers : Bernadette GOVAERTS and Nancy FRANCOIS

PROGRAM

"Sensory analysis and sensometrics: what is this ?"
Per Bruun BROCKHOFF, The Royal Veterinary and Agricultural University
of Copenhagen

"Experimental designs for sensory analysis"
Joachim KUNERT, University of Dortmund, Germany

"The statistical analysis of descriptive sensory data"
Per Bruun BROCKHOFF, The Royal Veterinary and Agricultural University
of Copenhagen

"Relating instrumental and sensory data: the use of multivariate data analysis"
Claire BOUCON, Unilever, The Netherlands

"Analysing time intensity curves: a case study in Beer tasting"
Nancy FRANCOIS, Université catholique de Louvain, Belgium

7.2 Second Young Researchers Day (YRD)

Louvain-la-Neuve, Belgium, April 30, 2004

Organisers : Carlos ALMEIDA, Taoufik BOUEZMARNI, Céline BUGLI, Anouar EL GHOUCH, Nancy FRANÇOIS, Gery GEENENS, Cédric HEUCHENNE, Astrid JULLION, Maria KEY PRATO, Alexandre LAMBERT, Céline LE BAILLY DE TILLEGHEM, Bianca TEODORESCU, Oana PURCARU, Sébastien VAN BELLEGEM, Marie VANDRESSE

Supported by the Graduate School of Statistics of the Institut de statistique and by SAS. It was also a FNRS contact day, Group of Statistical Analysis.

PROGRAM

SEMIPARAMETRIC SESSION

Elizabeth WILLIAMSON, London School of Hygiene & Tropical Medicine, United Kingdom
"Using propensity scores in practice: a comparison of methods for binary outcomes and exposures."

Michel DELECROIX, CREST, France
"Index Regression modeling : An Overview"

Vincent LEFIEUX, CREST, France
"Forecasting with MAVE: an empirical comparison between kernel estimator and moving average variance estimation"

Domenico GIANNONE, Université Libre de Bruxelles, Belgium
"A Maximum Likelihood Approach for Large Approximate Dynamic Factor Models"

NONPARAMETRIC SESSION

Natalie NEUMEYER, Ruhr-Universität Bochum, Germany
"Nonparametric comparison of regression curves - an empirical process approach"

Angelika ROHDE, Universität Heidelberg, Germany
"Approximate confidence sets for high dimensional signal centered at shrinkage estimators"

Karlien VANDEN BRANDEN, Katholieke Universiteit Leuven, Belgium
"Robust principal component analysis in high dimensions and applications"

Miguel DELGADO, Universidad Carlos III, Spain
"Significance testing in nonparametric regression based on the bootstrap"

8 ACADEMIC VISITS

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

January 2004

Michel DENUIT, "La longévité dans les produits de rente commercialisés par les sociétés belges d'assurance et les fonds de pension", Séminaire de l'Institut de Démographie, UCL, Belgium.

Michel DENUIT, "Lee-Carter type models for mortality projections", Technische Universität München, Zentrum Mathematik, Münich, Germany.

Ingrid VAN KEILEGOM, "Goodness-of-fit tests in nonparametric regression", Carlos III University, Madrid, Spain.

Ingrid VAN KEILEGOM, "Goodness-of-fit tests in nonparametric regression", University of La Coruna, Spain.

Ingrid VAN KEILEGOM, "Empirical likelihood in some non-standard settings", University of Santiago de Compostela, Spain.

February 2004

Ingrid VAN KEILEGOM, "Goodness-of-fit tests in nonparametric regression", University of Vigo, Spain

March 2004

Sébastien VAN BELLEGEM, "Adaptive estimation in locally stationary wavelet processes", Université Libre de Bruxelles, Belgium.

Sébastien VAN BELLEGEM, Weierstrass Institute for Applied Analysis and Stochastics, Berlin, Germany.

Ingrid VAN KEILEGOM, "Extending the scope of empirical likelihood", Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland.

April 2004

Ingrid VAN KEILEGOM, University of La Coruna, Spain.

Ingrid VAN KEILEGOM, University of Santiago de Compostela, Spain.

May 2004

Léopold SIMAR, Université des Sciences Sociales, Toulouse I, France.

Sébastien VAN BELLEGEM, "Adaptive estimation in locally stationary wavelet processes", University of Heidelberg, Germany.

June 2004

Sébastien VAN BELLEGEM, University of Heidelberg, Germany.

July 2004

Léopold SIMAR, Universität of Mainz, Germany.

Sébastien VAN BELLEGEM, University of Heidelberg, Germany.

Ingrid VAN KEILEGOM, "Empirical likelihood in some non-standard settings", University of Dortmund, Germany.

September 2004

Rainer von SACHS, University of Heidelberg, Germany.

October 2004

Léopold SIMAR, Santa-Anna School of Advanced Studies, Pisa, Italy.

Sébastien VAN BELLEGEM, London School of Economics, United Kingdom.

Sébastien VAN BELLEGEM, "Adaptive estimation in locally stationary wavelet processes", Imperial College, London, United Kingdom.

Ingrid VAN KEILEGOM, "Empirical likelihood tests for two-sample problems via nonparametric density estimation", Université de Rennes 1, France.

Rainer von SACHS, "SLEX principal components analysis of multivariate non-stationary time series", University of Heidelberg, Germany.

Rainer von SACHS, Bristol University, United Kingdom.

Rainer von SACHS, "SLEX principal components analysis of multivariate non-stationary time series", London School of Economics, United Kingdom.

November 2004

Sébastien VAN BELLEGEM, "Semiparametric estimation by model selection for locally stationary processes", London School of Economics, United Kingdom.

Rainer von SACHS, Bristol University, United Kingdom.

Rainer von SACHS, Université Joseph Fourier, Grenoble, France.

Rainer von SACHS, HEC, Genève, Switzerland.

December 2004

Irène GIJBELS, University of Tromsø, Norway.

Léopold SIMAR, Universidad del Pais Vasco, Bilbao, Spain.

Sébastien VAN BELLEGEM, London School of Economics, United Kingdom.

Sébastien VAN BELLEGEM, "Forecasting economic processes with locally stationary time series", Université Libre de Bruxelles, Brussels, Belgium.

Ingrid VAN KEILEGOM, "Empirical likelihood tests for two-sample problems via nonparametric density estimation", Tilburg University, The Netherlands.

Rainer von SACHS, HEC, Genève, Switzerland.

9 CONFERENCES AND MEETINGS

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

February 2004

Michel DENUIT, *2nd Actuarial and Financial Mathematics Day, Royal Flemish Academy of Belgium for Science and the Arts*, "How to deal with correlated risks in actuarial science? A case study with loss-alae data", Brussels, Belgium.

Rainer von SACHS, *Workshop on Multivariate Time Series*, "SLEX analysis of multivariate non-stationary time series", invited speaker, University of Heidelberg, Germany.

March 2004

Bernadette GOVAERTS, *Workshop on "news developments in experimental design"*, "Analysing experimental design results when the response is a curve : a case study in the polymer industry", invited speaker, Dortmund, Germany.

April 2004

Céline BUGLI, *DIA 4th International Workshop on Statistical Methodology in Non-Clinical R&D*, "Flexible modelling of event-related potentials: functional ANOVA with P-splines", contributed speaker, Dublin, Ireland.

Bernadette GOVAERTS, *DIA 4th International Workshop on Statistical Methodology in Non-Clinical R&D*, Dublin, Ireland.

May 2004

Michel DENUIT, *Workshop on "Dependence Modelling: Statistical Theory and Applications to Finance and Insurance"*, "Dependence in actuarial science: Stochastic orderings and dependence models over one period", Québec, Canada.

Anouar EL GHOUCH, *SFdS - XXXVIèmes Journées de Statistique*, "Empirical likelihood for censored dependent data", oral presentation, Montepellier, France.

Irène GIJBELS, *Second Erich L. Lehmann Symposium*, "Nonparametric testing for monotonicity of a hazard rate", invited speaker in the session on "Semiparametric and Nonparametric Testing", Rice University, Houston, USA.

Philippe LAMBERT, *Workshop on "Dependence Modelling: Statistical Theory and Applications to Finance and Insurance"*, "Bayesian estimation of Archimedean copulas", Québec, Canada.

Bianca TEODORESCU, *SFdS - XXXVIèmes Journées de Statistique*, "Generalized Conditional Linear Models under Left-Truncation and Right-Censoring", oral presentation, Montpellier, France.

Marie VANDRESSE, *21ème Journées de Microéconomie appliquée*, Lille, France.

June 2004

Hilmar BÖHM, *International Workshop on "Recent advances in time series"*, Poster, Protaras, Cyprus.

Michel DENUIT, *8th International Conference on Insurance: Mathematics & Economics*, "Actuarial Theory for Dependent Risks", Roma, Italy.

Léopold SIMAR, *North American Productivity Workshop III*, "Nonparametric Efficiency Analysis: a multivariate conditional quantile approach", invited discussant and member of the scientific committee, University of Toronto. Canada.

Rainer von SACHS, *Workshop "Recent Advances in Time Series Analysis"*, organiser, Protaras, Cyprus.

July 2004

Nancy FRANCOIS, *7th Sensometrics*, "A statistical global view of a Time-Intensity case study", Poster presentation, Davis CA, USA.

Irène GIJBELS, *International Forecasting Symposium*, "Semi-and Nonparametric methods of forecasting", organizer of invited paper session, Sydney, Australia.

Cédric HEUCHENNE, *67th Annual IMS meeting and 6th World Congress Bernoulli Society*, "Nonparametric censored regression with preliminary kernel smoothing", contributed paper, Barcelone, Spain.

Alexandre LAMBERT, *67th Annual IMS meeting and 6th World Congress Bernoulli Society*, "Jump preserving regression and smoothing using local linear fitting: a compromise", contributed paper, Barcelona, Spain.

Léopold SIMAR, "Nonparametric Efficiency Analysis: a multivariate conditional quantile approach", invited talk, University of Mannheim, Germany.

Marie VANDRESSE, *10th World Congress on Transport Research*, Istanbul, Turkey.

Sébastien VAN BELLEGEM, *24th International Symposium on Forecasting*, "Forecasting inhomogeneous time processes using a wavelet-based model", invited speaker, Sydney, Australia.

Ingrid VAN KEILEGOM, *6th World Congress of the Bernoulli Society*, "Empirical likelihood based goodness-of-fit testing for parametric regression" and "Goodness-of-fit tests", chairman session, Barcelona, Spain.

August 2004

Céline BUGLI, *25th annual conference of the International Society for Clinical Biostatistics*, "Flexible modelling of event-related potentials: functional ANOVA with P-splines", contributed speaker, Leiden, The Netherlands.

Philippe LAMBERT, *25th Annual Conference of The International Society for Clinical Biostatistics*, "Bayesian survival models with smooth time-varying coefficients: a penalized Poisson regression approach" and "Flexible modeling of event-related potentials: functional ANOVA with p-splines", Leiden, The Netherlands.

Ingrid VAN KEILEGOM, *Joint Statistical Meetings*, "Comparison of density functions via empirical likelihood", invited speaker, Toronto, Canada.

September 2004

Michel DENUIT, *Workshop on Lee-Carter methods*, "Lee-Carter model with Poisson random structure, and applications in insurance", Edinburgh, United Kingdom.

Irène GIJBELS, *Workshop "Journées Fonctionnelles"*, "Jump-preserving regression and smoothing discontinuous regression surfaces and image analysis", invited speaker, Rennes, France.

Bernadette GOVAERTS, *ENBIS conference*, "Analysing experimental design results when the response is a curve : methodology and industrial application", oral presentation, Copenhagen, Denmark.

Léopold SIMAR, *4th International Symposium of DEA*, "Estimation and inference in two-stage, semi-parametric models of production processes", invited keynote speaker, discussant and panelist, Aston University, Birmingham, United Kingdom.

October 2004

Céline BUGLI, *12th annual meeting of the Belgian Statistical Society*, "Registration of event-related potentials with fractional polynomials", poster, Vielsalm, Belgium.

Nancy FRANCOIS, *12th annual meeting of the Belgian Statistical Society* "A statistical global view of a Time-Intensity case study", Poster presentation, Vielsalm, Belgium.

Irène GIJBELS, *12th annual meeting of the Belgian Statistical Society*, chairman of a session, Vielsalm, Belgium.

Bernadette GOVAERTS, *SAS Forum*, Louvain-la-Neuve, Belgium.

Cédric HEUCHENNE, *12th annual meeting of the Belgian Statistical Society*, "Estimation in nonparametric location-scale regression models with censored data", contributed paper, Vielsalm, Belgium.

Alexandre LAMBERT, *12th annual meeting of the Belgian Statistical Society*, "Edge preserving and denoising digital images and regression surface estimation", contributed paper, Vielsalm, Belgium.

Léopold SIMAR, *12th annual meeting of the Belgian Statistical Society*, chairman of a session, Vielsalm, Belgium.

Léopold SIMAR, *Working group on "Analysis of Productivity of European Universities", AQUAMETHPSR project under the PRIME Network of Excellence supported by the European Commission, 6th Framework Programme*, Lugano, Swiss.

November 2004

Philippe LAMBERT, *IAP meeting (Interuniversity Attraction Pole, Phase V, 2002-2006*, "Flexible proportional hazard model with time-varying coefficients: a Poisson approach", invited speaker, Leuven, Belgium.

December 2004

Bernadette GOVAERTS, *Chimiometry Congress 2004*, "Analyse des résultats d'un plan d'expérience quand la réponse est une courbe : méthodologie et application en chimie", oral presentation, Paris, France.

Léopold SIMAR, "Robust Nonparametric Estimation of Frontiers", invited talk, Universidad del Pais Vasco, Bilbao, Spain.

Marie VANDRESSE, *Euroconference Series in Quantitative Economics and Econometricsn :(EC)2*, Marseille, France.

10 RESEARCH PROJECTS UNDER CONTRACTS AND COOPERATION PROJECTS

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

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

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

Coordinators : I. GIJBELS (until September 30, 2004) and L. SIMAR

STAT researchers : J. BIGOT, G. GEENENS, S. O. JEONG, A. LAMBERT and B. TEODORESCU

Partners Institutions :

- Katholieke Universiteit Leuven, Belgium
- Limburgs Universitair Centrum, Diepenbeek, Belgium
- Université Libre de Bruxelles, Belgium
- Aachen Technical University, Germany
- Université Joseph Fourier, Grenoble, France

The point of departure of the network activities is that of a broad range of complex substantive data sets and questions arising in various disciplines (including psychology, biomedical sciences, economics, and climatology). The overall aim of our project then is to develop appropriate statistical models and techniques to deal with these data and questions.

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

Financing : Fonds spéciaux de recherche (FSR)

Promoter : R. von SACHS

Researchers : H. BÖHM

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

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

11 ACTIVITIES IN APPLIED STATISTICS

The Institut de statistique is developing many contacts within the Université catholique de Louvain and with several companies in the field of applied statistics. In addition to the seminars organised weekly (see point 5), the members of the Institute participate to research contracts in applied statistics and offer consulting services to other departments and institutions of the University. They also offer some courses of continued education at the University and in companies.

The major activities are described below.

11.1 Services for the Université catholique de Louvain

The Institut de statistique offers a consulting service to the University community. Researchers of the others faculties receive advice concerning appropriate methodologies and suitable statistical packages for their specific problems. Hence, scientific collaborations between different disciplines are often created.
More information is available on the web page:
http://www.stat.ucl.ac.be/ISconsultation

The scientific members and the computer scientists of the Institute are developing their knowledge about the evolution of many statistical software and they often give advice in this context. The Institute hosts the "Statistical application server" used daily by UCL researchers and students for their statistical computations.
More information concerning this server is available on the web page:
http://wwwsas.stat.ucl.ac.be

11.2 Applied research contracts

"Automatic statistical analysis in the time-frequency domain" (2000-2003)

Financing : "The National Institute of Mental Health" (USA)

Promoter : R. von SACHS

In the collaborative project with biostatisticians (H. OMBAO, Department of Statistics, University of Illinois, Champaign, USA and W. GUO, School of Public Health, UPENN, Philadelphia) and neurologist (B. LITT, UPENN, Philadelphia) we try to improve the modelling and understanding of the evolution of epileptic seizure in the brain. The final goal of this research is to try to localize the centre of the seizure more precisely in order to improve the surgical treatment. In the first phase of the project (2/2000-1/2001) we have developed a new algorithm for

bivariate spectral analyis of EEG data which is more localized but still relatively close to the methods medical doctors are used to. In the second phase (2/2001-1/2003) we have generalised our method in order to treat multivariate EEG signals of higher dimension.

"Assessment of quality differences between freight transport modes" (2001-2003)

Financing : Belgian Science Policy, Brussels, Belgium (Second multiannual scientific support plan for a sustainable development policy)

Promoter : M. MOUCHART

Researchers : C. KRIER, A. OULHAJ, G. SANTAMARIA and M. VANDRESSE

The objective of this contract is to analyze the qualitative differences between means of long-distance freight transport: reliability, security, flexibility, punctuality, information, damages, etc. We will measure the impact of these qualitative factors concerning the decision of a transport mode and, if possible, we will evaluate the corresponding advantages (or costs).

"Analysis and interpretation of electroencephalograms (EEG) in drug discovery" (2001-2005)

Financing : Eli Lilly (October 2001 till August 2003)
 FSR (September 2003)
 FRIA (October 2003 - 2005)

Promoter : Ph. LAMBERT

Researcher : C. BUGLI

Electroencephalograms (EEG) can be used to improve the understanding of the effects of experimental drugs on the body. Such signals can for example be generated during a drug development process using human volunteers exposed to varying drug concentrations. EEG signals typically take the form of longitudinal continuous responses measured at high frequency during at most 10 minutes at various locations at the surface of the skull. Currently, very simple descriptive analyses of such data are undertaken. Typically, each signal (generated at one electrode) is subdivided in consecutive periods of (say) 2 seconds. It is assumed stationary in these small epochs and separately analyzed using Fourier transforms. The goal of the project is

- to evaluate the potential of independent component analysis.

- to develop functional data analysis models to analyse event related potentials.

- to illustrate the potential of these techniques to quantify the effect of a treatment on the brain.

EEG signals generated during cross-over trials where patients alternatively received a placebo and a drug reported to have a large and well understood effect on the brain are potential elements to use in that illustration.

"Datawarehouse of civil servants : analysis of external needs" (2003-2004)

Financing : Belgian Science Policy, Brussels, Belgium

Promoter : M. DENUIT

Researcher : N. BROUHNS

The study aims to design a socio-economic datawarehouse relating to the pensions paid to civil servants. This information tool will be used by the Ministry of Finance to monitor financial aspects of these pensions.

"Mortality of civil servants : actuarial and demographic analysis" (2003-2004)

Financing : Belgian Science Policy, Brussels, Belgium

Promoter : M. DENUIT

Researcher : A. DELWARDE

The study aims to provide to the Ministry of Finance an actuarial and demographic support to the software Exsyspen used to forecast financial aspects of the pensions paid to civil servants.

"A model for trend estimation adapted to monitoring data" (2003-2004)

Financing : EUROCHLOR, Belgium

Promoter : Ph. VANDEN EECKAUT

Researcher : S. VAN BELLEGEM

This project is the last part of the second project between Euro Chlor and the Institute of Statistics. During the first collaboration between the Institute and Euro Chlor, we proposed an indicator for the estimation of slopes with monitoring data based on Sen's indicator. This indicator was illustrated on the WRC database. The conclusions of this work were mixed.
On the positive side, the indicator is easy to understand and to compute, and it

offers an adequate flexibility. On the negative side, we have problem to give a statistical interpretation of the results. The objective of our new proposal is to provide a statistical interpretation of the trend while maintaining a model which is as easy to interpret than the Sen's indicator. We use a completely different way to discover the trend which, we believe, is well adapted to the objective.

"Statistical analysis of PET scan data" (2003-2007)

Financing : Eli Lilly (October 2003 till September 2005)

Promoter : Ph. LAMBERT

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.

The final tool will allow to visualize where the drug is acting on the brain and how that action is changing over time.

Statistical evaluation of the medal distribution methodology in international wine tasting competitions : Application to the "Concours Mondial de Bruxelles" (2004-2005)

Financing : Vinopress, Brussels, Belgium

Promoters : B. GOVAERTS and C. RITTER

The goal of the project is to assess the medal allocation methodology in the Brussels Wine tasting international competition. It will provide a wide review of the literature, a statistical analysis of data available from former competitions and suggestions to improve reliability of future ones.

"Evaluating and updating econometric models for nowcasts on R & D variables: 2004 results" (2004)

Financing : CAMIRE

Promoters : M. MOUCHART

Ressearcher : J. ROMBOUTS

The main object of this report is to evaluate, on the basis of actual realizations, and to update the nowcast of R & D variables elaborated in Mouchart and Rombouts

(2003). The data set underlying this report updates the data set previously used and has been provided by CAMIRE. A secondary objective is to put on trial a new clustering strategy, based on a statistically stronger basis than the ones used up to now.

"Robust models for the analysis of clustered survival data" (2001-2005)

Financing : Fonds propres (patrimoine), Conventions

Promoter : Ph. LAMBERT

Researchers : D. COLLET and A. KIMBER (University of Reading)

This is a joint project with Prof. Dave Collett and Alan Kimber (University of Reading). We are extending the accelerated failure time model to deal with clustered survival data. These tools are robust alternatives to the frailty proportional hazard model which is usually proposed to deal with heterogeneity in survival data. They are used to analyse the survival times of grafted patients coming from one of several transplant centres in the UK. The data are provided by UK transplant (Bristol, UK).

11.3   Consulting for firms and public organisms

The principal consulting contracts treated in 2004 are the following:

Client : Media Conseil, Brussels, Belgium
Subject : Probability analysis of the expected gain of a new lotery game.
STAT participant : J.M. ROLIN

Client : Eurocontrol, Brussels, Belgium
Subject : Efficiency Analysis of Air Navigation Services Provision
STAT participants : M. MOUCHART and L. SIMAR

Client : Eurochlor, Brussels, Belgium
Subject : A model for trend estimation adapted to monitoring data
STAT participants : S. VAN BELLEGEM and Ph. VANDEN EECKHAUT

Client : IWEPS, Jambes, Belgium
Subjet : Nonparametric estimation of stochastic kernels
STAT participant : S. VAN BELLEGEM

Client : asbl La Rue
Subject : Hability for habitation: study of the healthiness of accomodations in Molenbeek
STAT participants: A. ELGHOUCH, B. GOVAERTS and E. LECOUTRE

Client : asbl La Rue
Subject : Suivi du code SPSS et de la base Filemaker utilisés pour les analyses statistiques
STAT participant: E. LECOUTRE