

# CONFERENCE ON "QUANTITATIVE METHODS IN STATISTICS, BIOSTATISTICS AND ACTUARIAL SCIENCES"

## *Speakers - Abstract*

<b>ARTZNER Philippe</b> Multiperiod risk measurement : problems, examples, comments	2
<b>CARROLL Raymond J.</b> What percentage of children in the U.S. are eating a healthy diet ? A statistical approach.	3
<b>EECKHOUDT Louis</b> From risk aversion to other risk attitudes	4
<b>EILERS Paul</b> The beauty of tensor-product P-splines	5
<b>EMBRECHTS Paul</b> Risk, regulation and statistics	6
<b>GONZÁLEZ-MANTEIGA Wenceslao</b> Extending induced ROC methodology to the functional context	7
<b>JANSSEN Paul</b> Goodness-of-fit tests for the frailty distribution in proportional hazards models with shared frailty	8
<b>KNEIP Alois</b> Boundary estimation in the presence of measurement error with unknown variance	9
<b>LAEVEN Roger</b> Risk management. When financial institutions can get in trouble together	10
<b>LIEBSCHER Eckhard</b> Approximation of distributions using Anderson-Darling and Cramér-von Mises statistics	11
<b>LINTON Olivier</b> Estimation of risk measures under heavy tails	12
<b>PATILEA Valentin</b> Projection-based nonparametric goodness-of-fit testing with functional data	13
<b>RONDEAU Virginie</b> Cure frailty models for survival data: Application to recurrences for breast cancer and to hospital readmissions for colorectal cancer	14
<b>WELLNER Jon A.</b> Log-concavity - and why it matters	15

# *Multiperiod risk measurement : problems, examples, comments*

## **AUTHOR**

ARTZNER Philippe (Université de Strasbourg)

## **ABSTRACT**

Risk measurement for insurance supervisory purposes requires a first «market consistent» valuation of usually non hedgeable obligations cash-flows. It is now well agreed to use expectations (or «best estimates») with respect to a pricing measure of the traded securities. It is then possible to speak of «risk bearing capital» as the difference of market value of assets (less insurance obligations payments) and some market consistent value of remaining obligations. These values are adapted and respectively retrospective and prospective processes.

We present an axiomatic way to give a definition of solvency. One can indeed posit the existence of an operator from risk bearing capital processes to «free capital» processes, positivity of free capital meaning solvency.

It is natural to require from the free capital assessment to be «supervisory arbitrage free» to prevent manipulations of solvency standards by too clever market instruments. It is also natural to expect a «time consistency» of the free capital operator. Given an exogeneous obligations portfolio, asset portfolio choice is being used as control variable. The «supervisory provision» is the minimum asset value allowing some hedging of the obligations as well as solvency (or «acceptability») of the hedged obligations. The supervisory provision is a market consistent assessment of obligations.

If the current asset value is smaller than the supervisory provision the sine qua non condition is not fulfilled and the «exit value» or «technical provision» is required from the non acceptable company for a transfer of its obligations as well as of the technical provision to a new company. The technical provision, another market consistent assessment of obligations, equals the sum of the best estimate and of the «risk margin» while the new shareholders provide an «absolute capital requirement».

Finally solvency capital requirement is defined to deal with the case where shareholders / managers want a definite, non optimal asset portfolio. It adds an «hedgeable» term to the absolute capital requirement and fulfills the supervisory accounting equality.

Application of the axiomatic approach to the case of non hedgeable assets will be mentioned, in connection with supervisory treatment of reinsurance.

# *What percentage of children in the U.S. are eating a healthy diet? A statistical approach*

## **AUTHOR**

CARROLL Raymond J. (Texas A&M University)

## **ABSTRACT**

In the United States the preferred method of obtaining dietary intake data is the 24-hour dietary recall, yet the measure of most interest is usual or long-term average daily intake, which is impossible to measure. Thus, usual dietary intake is assessed with considerable measurement error. Also, diet represents numerous foods, nutrients and other components, each of which have distinctive attributes. Sometimes, it is useful to examine intake of these components separately, but increasingly nutritionists are interested in exploring them collectively to capture overall dietary patterns and their effect on various diseases. Consumption of these components varies widely: some are consumed daily by almost everyone on every day, while others are episodically consumed so that 24-hour recall data are zero-inflated. In addition, they are often correlated with each other. Finally, it is often preferable to analyze the amount of a dietary component relative to the amount of energy (calories) in a diet because dietary recommendations often vary with energy level.

We propose the first model appropriate for this type of data, and give the first workable solution to fit such a model. After describing the model, we use survey-weighted MCMC computations to fit the model, with uncertainty estimation coming from balanced repeated replication. The methodology is illustrated through an application to estimating the population distribution of the Healthy Eating Index-2005 (HEI-2005), a multi-component dietary quality index involving ratios of interrelated dietary components to energy, among children aged 2-8 in the United States. We pose a number of interesting questions about the HEI-2005, and relate it also to the risk of developing colorectal cancer.

# *From risk aversion to other risk attitudes*

## **AUTHOR**

EECKHOUDT Louis (Université catholique de Louvain and Ieseg, Lille)

## **ABSTRACT**

In this presentation I'll first indicate how economists have progressively looked at concepts more specific than that of risk aversion, such as e.g. prudence or temperance. Then by using the idea that decision makers are «correlation averse», I'll show that these concepts can be interpreted in a uniform manner. Finally attention will be paid to the behavior of risk lovers.

# *The beauty of tensor-product P-splines*

## **AUTHOR**

EILERS Paul (Erasmus Universiteit MC, Rotterdam)

## **ABSTRACT**

P-splines have become popular for smoothing. They are based on a simple principle: use many equally-spaced B-splines in a regression basis, and tune smoothness with a discrete difference penalty on the coefficients. P-splines have many applications and strong ties to mixed models and Bayesian estimation.

Generalization of the regression basis to two dimensions is relatively straightforward with tensor products of B-splines. The coefficients now form a matrix and penalties get more complicated. Instead of one penalty, as in the one-dimensional case, there will be a penalty for each row and each column of the matrix. Implementation of tensor-product P-splines has to be done with care, but it leads to a flexible tool for two-dimensional smoothing, interpolation and extrapolation. The data can be normal, counts, or binary observations.

When data have been collected on a rectangular grid, we can use GLAM, the generalized linear array model. The computations with tensor products can be organized in such a way that dramatic reductions of memory use and computation time can be obtained when the grid is large.

The presentation will be non-technical and it will be illustrated by real-life applications.

# *Risk, regulation and statistics*

## **AUTHOR**

EMBRECHTS Paul (ETH Zürich)

## **ABSTRACT**

Since the early nineties, regulators of financial institutions worldwide (the Basel Committee) have hard-wired the calculation of extreme risk measures into the national laws on banking and insurance regulation. A prominent example is the so-called Value-at-Risk (VaR), a quantile based risk measure to be calculated far in the loss tail of the Profit-&-Loss (P&L) distribution. Standard quantile levels are 95%, 99%, 99.9% even 99.97%, hence EXTREME quantiles. In this talk I will review some of the methods from extreme value theory used for the estimation of such risk measures and this for typical financial time series. Besides giving an overview of some of the underlying methodological and practical issues, I will also show how questions of the above type may lead to interesting research in statistics. The latter will be based, for instance, on the recent: Chavez-Demoulin, V., Embrechts, P. and Sardy, S. (2011) Extreme-quantile tracking for financial time series. *Journal of Econometrics*, to appear.

# *Extending induced ROC methodology to the functional context*

## **AUTHOR**

GONZÁLEZ-MANTEIGA Wenceslao (Universidad de Santiago de Compostela)

## **ABSTRACT**

The receiver operating characteristic (ROC) curve is the most widely used measure for evaluating the discriminatory performance of a continuous marker. Often, covariate information is also available and several regression methods have been proposed to incorporate covariate information in the ROC framework. Until now, these methods are only developed for the case where the covariate is univariate or multivariate. We extend ROC regression methodology for the case where the covariate is functional rather than univariate or multivariate. To this end, semiparametric- and nonparametric-induced ROC regression estimators are proposed. A simulation study is performed to assess the performance of the proposed estimators. The methods are applied to and motivated by a metabolic syndrome study in Galicia (NW Spain).

# *Goodness-of-fit tests for the frailty distribution in proportional hazards models with shared frailty*

## **AUTHOR**

JANSSEN Paul (Université de Strasbourg)  
*Joint work with C. Geerdens, G. Claeskens*

## **ABSTRACT**

Frailty models account for the clustering present in grouped event time data. A proportional hazards model with shared frailties expresses the hazard for each subject. Often a one-parameter gamma distribution is assumed for the frailties. The choice of a particular frailty distribution is, most of the time, based on the availability of software, rather than on the way it fits the data. In this paper we construct formal goodness-of-fit tests to test for gamma frailties. We construct a new class of frailty models that extend the gamma frailty model by using certain polynomial expansions that are orthogonal with respect to the gamma density. For this extended family we obtain an explicit expression for the marginal likelihood of the data. The order selection test is based on finding the best fitting model in such a series of expanded models. A bootstrap is used to obtain p-values for the tests. Simulations and data examples illustrate the test's performance.

# *Boundary estimation in the presence of measurement error with unknown variance*

## **AUTHOR**

KNEIP Alois (Universität Bonn)

*Joint work with L. Simar, I. Van Keilegom*

## **ABSTRACT**

Boundary estimation appears naturally in economics in the context of productivity analysis. The performance of a firm is measured by the distance between its achieved output level (quantity of goods produced) and an optimal production frontier which is the locus of the maximal achievable output given the level of the inputs (labor, energy, capital, etc.). Frontier estimation becomes difficult if the outputs are measured with noise and most approaches rely on restrictive parametric assumptions. This paper contributes to the direction of nonparametric approaches. We consider a general setup with unknown frontier and unknown variance of a normally distributed error term, and we propose a nonparametric method which allows identifying and estimating both quantities simultaneously. The asymptotic consistency and the rate of convergence of our estimators are established, and simulations are carried out to verify the performance of the estimators for small samples. We also apply our method on a dataset concerning the production output of American electricity utility companies.

# *Risk management. When financial institutions can get in trouble together*

## **AUTHOR**

LA EVEN Roger (Tilburg University)

## **ABSTRACT**

We study a statistical loss model in which big financial losses tend to come in close succession, reflecting the features in the real data. We characterize in closed-form the tails of this model, both in the univariate and in the multivariate version of the model. We further develop analytic risk measure formulae for this model. The closed-form feature of our approach has some important advantages: it means, for example, that the effort in computing the risk measures is minimal, and that our risk measure formulae may readily be used for comparative statics, parameter calibration, capital requirement calculation, and the computation of a possible systemic risk charge.

## **KEYWORDS:**

Risk measures; Systemic risk; Jumps; Mutually exciting processes.

# *Approximation of distributions using Anderson-Darling and Cramér-von Mises statistics*

## **AUTHOR**

LIEBSCHER Eckhard (Hochschule Merseburg)

## **ABSTRACT**

In many applications the distribution of the data differs significantly from common distribution models. Especially in the situation of high-dimensional data, it is hard to find an appropriate distribution model at all. In this case it frequently happens that all hypotheses of known distribution models are rejected by usual significance tests. In this talk we consider the problem of approximation by a given parametric family of distributions. We search for the distribution from the parametric family which approximates best the distribution of the data. It should be pointed out that the corresponding best parameter depends heavily on the approximation measure. There is no «true parameter».

In the talk we consider two situations. The first part of the talk deals with approximations of the distribution by use of the Anderson-Darling statistic. The second part of the talk is devoted to the approximation of copulas. Here the focus is on approximations by a version of the Cramér-von Mises statistic. For both situations, we provide results on the asymptotic behaviour of the estimator for the parameter (consistency, asymptotic normality) and a test of goodness of approximation. Considering significance test of goodness of fit, one has to tackle the problem that the asymptotic distribution of usual statistics is rather complicated (not normal). In our context this problem does not exist. The techniques used for establishing the results work also in other cases where the test statistic can be represented as a U-statistic.

# *Estimation of risk measures under heavy tails*

## **AUTHOR**

LINTON Olivier (University of Cambridge)

## **ABSTRACT**

We study estimation and inference of Expected Shortfall (ES) for time series with Infinite variance. The rate of convergence is determined by the tail thickness parameter and the limiting distribution is in the stable class with parameters depending on the tail thickness parameter of the time series and on the dependence structure, which makes inference complicated. A subsampling procedure is proposed to carry out statistical inference. We also analyze a nonparametric estimator of conditional expected shortfall. We provide an application to the Russian Rouble.

# *Projection-based nonparametric goodness-of-fit testing with functional data*

## **AUTHOR**

PATILEA Valentin (CREST - Ensai, Rennes)

*Joint work with C. Sánchez-Sellero, M. Saumard*

## **ABSTRACT**

The problem of nonparametric testing for the effect of a random functional covariate on a real-valued or functional error term is studied. The functional variables take values in  $L^2[0,1]$ , the Hilbert space of the square-integrable real-valued functions on the unit interval. The error term could be directly observed as a response or estimated from a functional parametric model, like for instance the functional linear regression. Our test is based on the remark that checking the no-effect of the functional covariate is equivalent to checking the nullity of the conditional expectation of the error term given a sufficiently rich set of projections of the covariate. Such projections could be on elements of norm 1 from finite-dimension subspaces of  $L^2[0,1]$ . Next, the idea is to search a finite-dimension element of norm 1 that is, in some sense, the least favorable for the null hypothesis. Finally, it remains to perform a nonparametric check of the nullity of the conditional expectation of the error term given the scalar product between the covariate and the selected least favorable direction. For such finite-dimension search and nonparametric check we use a kernel-based approach. As a result, our test statistic is a quadratic form based on univariate kernel smoothing and the asymptotic critical values are given by the standard normal law. The test is able to detect nonparametric alternatives, including the polynomial ones. The error term could present heteroscedasticity of unknown form. We do not require the law of the covariate  $X$  to be known. The test could be implemented quite easily and performs well in simulations and real data applications. We illustrate the performance of our test for checking the functional linear regression model.

*Cure frailty models for survival data:  
Application to recurrences for breast cancer and  
to hospital readmissions for colorectal cancer*

**AUTHOR**

RONDEAU Virginie (Université Victor Segalen, Bordeaux)

**ABSTRACT**

Owing to the natural evolution of a disease, several events often arise after a first treatment for the same subject. For example, patients with a primary invasive breast cancer and treated with breast conserving surgery may experience breast cancer recurrences, metastases or death. A certain proportion of subjects in the population who are not expected to experience the events of interest are considered to be 'cured' or non-susceptible. To model correlated failure time data incorporating a surviving fraction, we compare several forms of cure rate frailty models. In the first model already proposed non-susceptible patients are those who are not expected to experience the event of interest over a sufficiently long period of time. The other proposed models account for the possibility of cure after each event. We illustrate the cure frailty models with two data sets. First to analyze time-dependent prognostic factors associated with breast cancer recurrences, metastases, new primary malignancy and death. Second to analyze successive rehospitalizations of patients diagnosed with colorectal cancer. Estimates were obtained by maximization of likelihood using SAS proc NLMIXED for a piecewise constant hazards model. As opposed to the simple frailty model, the proposed methods demonstrate great potential in modelling multivariate survival data with long-term survivors ('cured' individuals).

# *Log-concavity - and why it matters*

## **AUTHOR**

WELLNER Jon A. (University of Washington)

## **ABSTRACT**

Log-concavity plays an interesting and important role in statistical theory. I will review definitions of log-concavity and strong log-concavity, and sketch a few of the many consequences for statistics and probability. I will also briefly review some progress on proving log-concavity for distributions appearing in statistical theory, and very briefly sketch recent progress concerning nonparametric estimation of log-concave densities and related families in  $R^1$  and  $R^d$ .