National and Kapodistrian University of Athens
Department of Mathematics

BOOK OF ABSTRACTS

ONE DAY WORKSHOP ON PROBABILITY AND STATISTICS

Dedicated to Professors Emeriti

CHARALAMBOS A. CHARALAMBIDES AND OURANIA CHRYSSAFINOU

on their retirement

June 17, 2013

Department of Mathematics
National and Kapodistrian University of Athens

Sponsors

	University of Athens
	H.S.P.H.R.C.E.
	Blue Star Ferries
	Hellenic Mathematical Society

Local Organizing Committee

(from the Section of Statistics and Operations Research,
Department of Mathematics
National and Kapodistrian University of Athens)

- **A. Burnetas**
- **A. Economou**
- **N. Papadatos**
- **E. Vaggelatou**

Speakers

- M. Akritas (Penn. State University, U.S.A.)
- N. Balakrishnan (McMaster University, Canada)
- T. Cacoullos (University of Athens, Greece)
- T. Christofides (University of Cyprus, Cyprus)
- M.V. Koutras (University of Piraeus, Greece)
- N. Limnios (Univ. Techn. Compiègne, France)
- N. Papadatos (University of Athens, Greece)
- G. Touloumi (University of Athens, Greece)

ABSTRACTS

Test-Based Variable Selection in Nonparametric Models

Michael Akritas

Department of Statistics

Pennsylvania State University, USA

Abstract

Under the nonparametric model $Y = m(\mathbf{X}) + \sigma(\mathbf{X})\,\epsilon$, we develop an ANOVA-type test for the null hypothesis that a particular coordinate of $\mathbf{X}$ has no influence on the regression function. The asymptotic distribution of the test statistic, using residuals based on local polynomial regression, is established under the null hypothesis and local alternatives. Simulations suggest that the test outperforms existing procedures in heteroscedastic settings. Using *p*-values from this test, a variable selection method based on False Discovery Rate corrections is proposed, and proved to be consistent in estimating the set of indices corresponding to the significant covariates. Simulations suggest that, under a sparse model, dimension reduction techniques can help avoid the curse of dimensionality. A backward elimination version of this procedure, called BEAMS (Backward Elimination ANOVA-type Model Selection), performs competitively against well established procedures in linear regression settings, and outperforms them in nonparametric settings. A real data set is analyzed.

Signatures

N. Balakrishnan

McMaster University

Hamilton, Ontario, Canada

Abstract

In this talk, I will first introduce the concept of signatures of coherent systems and give some examples. Next, I will describe some of its properties such as stochastic orderings and comparison of different reliability systems. Then, I will present an extended signature of a system that would facilitate comparisons of systems of different sizes. Following this, I will introduce a new notion of dynamic signatures and describe some of its properties. Finally, I will consider a situation wherein there are two systems with some shared components and introduce the concept of bivariate signatures and establish some of its structural properties as well as associated stochastic orderings.

The distribution of ratios from all centered spherically or elliptically symmetric distributions are Cauchy

T. Cacoullos

Department of Mathematics

University of Athens, Greece

Abstract

A multivariate Cauchy distribution is shown to arise as the joint distribution of ratios of components of a random vector $\boldsymbol{X} = (X_1, \ldots, X_n)$ which follows a spherically or elliptically symmetric distribution centered at zero. This extends, for $n>2$, the results for the case of one ratio ($n=2$), which is known to follow a standard or general (one-dimensional) Cauchy distribution ([1], [2], [3]). Cauchy distributions are closely examined mainly as distributions of component ratios of vector uniformly distributed on spheres. Furthermore, the well-known distributions of Student's t-statistic and the correlation coefficient ρ (when $\rho = 0$) follow immediately from the distribution of the polar angle (colatitude) θ, namely, under spherical symmetry, $\sqrt{n-1}\cot(\theta)$ is t_{n-1}.

References

[1] B. C. Arnold, P. L. Brockett, On Distributions Whose Component Ratios Are Cauchy, *Amer. Statist.* **46** (1992), 25-26.

[2] M. C. Jones, Distributional Relationships Arising From Simple Trigonometric Formulas, *Amer. Statist.* **53** (1999), 99-102.

[3] M. C. Jones, The distribution of the ratio X/Y for all centered elliptically symmetric distributions, *J. Multivariate Anal.* **99** (2008), 572-573.

Demimartingale Inequalities and Related Results

Tasos C. Christofides

Department of Mathematics and Statistics

University of Cyprus, Cyprus

Abstract

Demimartingales and N-demimartingales generalize in a natural way the classes of positively and negatively associated random variables respectively and include, as a special case, the class of martingales with respect to the natural choice of σ-algebras. In this review talk we provide several examples of demimartingales (and N-demimartingales) and various results including probability inequalities and asymptotic theorems. Furthermore, we introduce the class of conditional demimartingales and make the connection between demimartingales and stochastic orders.

Exact and asymptotic results for pattern-related statistics

Markos V. Koutras

Department of Statistics and Insurance Science

University of Piraeus, Greece

Abstract

The study of statistics related to patterns has a long history and the examination of their properties has been a topic of continuous research interest till today. This is chiefly due to the numerous applications they have been used for in many fields, including biological sequence analysis, randomness testing, quality control, educational psychology etc.

Besides the traditional combinatorial analysis (which is not easy to generalize and, when dealing with multi-state trials, becomes intractable), two alternative methods that have been suggested are the finite Markov chain imbedding technique (FMCI) and techniques exploiting conditional generating functions. Both of them can be profitably used for developing exact and asymptotic results for pattern enumeration and pattern waiting time problems.

In this talk we shall first present a review of the methods used to study pattern-related problems (exact and asymptotic). We shall then illustrate how the FMCI method can be exploited to derive a recursive scheme for the evaluation of the tail probabilities for the waiting time for the first and r-th occurrence of a pattern. Besides that, we shall present a class of asymptotic results which, instead of taking the classical approach that exploits the Perron-Frobenius eigenvalue of the transition probability matrix associated to the pattern of interest, makes use of the trace of it. These results cover a quite wide class of waiting time problems and, in most cases, perform much better than the ones using the Perron-Frobenius eigenvalue.

On the approximation of Poisson and Binomial distributions

Nickos Papadatos

Department of Mathematics

University of Athens, Greece

Abstract

The order of the classical approximation of Binomial distribution $\mathcal{B}\mathrm{in}(\lambda\,;\lambda/n)$ (with fixed $\lambda>0$ and $n\to+\infty$) by the Poisson law $\mathcal{P}\mathrm{o}(\lambda)$ is not, in general, best possible. In the present talk we consider order n^{-1} perturbations of the Poisson mean θ of the form

$$\theta = \theta_n(\lambda; y) := \lambda + \frac{y}{n}\lambda, \quad y \in R,$$

and our purpose is to calculate the minimal value of the constant $C=C(\lambda;y)>0$ in the limit

$$\lim_{n\to\infty}\left\{n\mathrm{d}_{TV}\left[\mathcal{B}\mathrm{in}\left(n;\frac{\lambda}{n}\right),\mathcal{P}\mathrm{o}\left(\lambda+\frac{y}{n}\lambda\right)\right]\right\} =: C(\lambda; y),$$

where d_{TV} denotes the total variation distance. Clearly, the practical interest is to obtain the value(s) of $y \in R$ as a function of any fixed $\lambda>0$. Thus, by definition, $y=y(\lambda)$ denotes the value(s) of y that minimizes the constant $C(\lambda;y)$, appearing in the above limit. In the present talk we present a simple closed form for any integer value of λ, i.e., for the sequence $\{y(k)\}_{k=1}^{\infty}$. In particular, the following, somewhat surprising result, is true:

$$y(k+k^2) = 0 \text{ and } y(k^2) = \tfrac{1}{2} \quad (k = 1,2,\dots).$$

As a by-product it turns out that a squared integer value of $\lambda>0$ requires the maximal (among the set of integers) perturbation. In contrast, the perturbation is completely unnecessary if and only if the integer $\lambda>0$ is "far from squares", that is, $\lambda = k + k^2$ for some integer $k>0$.

Methods for competing risks data: dealing with missing cause of failure

Giota Touloumi and Giorgos Bakoyannis

Department of Hygiene, Epidemiology and Medical Statistics

Athens University Medical School, Greece

Abstract

Competing risks data arise naturally in medical research, when subjects under study are at risk of more than one mutually exclusive event such as death from different causes. The competing risks framework also includes settings where different possible events are not mutually exclusive but the interest lies on the first occurring event. When competing risks data arise, information on actual cause of failure for some subjects might be missing. Possible strategies for analyzing such data include the complete case (CC) analysis as well as an analysis where the missing causes are classified as an additional failure type (AC). Another recently proposed method is the multiple imputation (MI) method. In this presentation I will introduce competing risks data and I will critically review the widely used statistical methods for estimation and modeling. I will mainly focus on the increasingly popular proportional subdistribution model (Fine and Gray model). Finally, results from simulation experiments under various scenarios evaluating the comparative performance of the CC, AC and MI method to analyze competing risk data with missing cause of failure, will be presented.