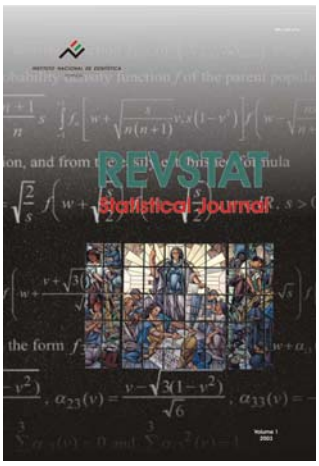
January 07, 2005

Estatísticas Gerais

Número 2 - 2004

REVSTAT-STATISTICAL JOURNAL

REVSTAT-STATISTICAL JOURNAL, Volume 2, No. 2 – November 2004



In 2003 the National Statistical Institute launched the scientific statistical journal **REVSTAT-STATISTICAL JOURNAL**, published in English two times a year, with a prestigious international Editorial Board, which came to substitute the *Revista de Estatística* [Statistical Review], published in Portuguese between 1996 and 2002.

The aim of the Editorial Board of **REVSTAT** is to publish articles of high scientific content, developing innovative statistical scientific methods and introducing original research, grounded in substantive problems, covering all branches of Probability and Statistics. Surveys of important areas of research in the field are also welcome.

REVSTAT hopes to become a place where scientists may feel proud of publishing their research results changing the character of the previous *Revista de Estatística* from a national to an international scientific journal.

For more information about **REVSTAT**, namely on-line articles, subscription of the publication, and submission of papers, please visit the link of the National Statistical Institute's website: http://www.ine.pt/revstat.html

This Volume of **REVSTAT: Volume 2, No. 2–November 2004**, which is now come out, publish four articles of which abstracts are presented down:

LOCAL FITTING WITH A POWER BASIS

Author: *Jochen Einbeck*

Local polynomial modeling can be seen as a local fit of the data against a polynomial basis.

In this paper we extend this method to the power basis, i.e. a basis which consists of the powers of an arbitrary function.

Using an extended Taylor theorem, we derive asymptotic expressions for bias and variance of this estimator.

We apply this method to a simulated data set for various basis functions and discuss situations where the fit can be improved by using a suitable basis.

Finally, some remarks about bandwidth selection are given and the method is applied to real data.

CENTRAL PARTITION FOR A PARTITION-DISTANCE AND STRONG PATTERN GRAPH

Authors: *Joaquim F. Pinto da Costa* e *P. R. Rao*

When several clustering algorithms are applied to a dataset *E* or the same algorithm with different parameters, we get several different partitions of the dataset.

In this paper we consider the problem of finding a consensus partition between the set of these partitions.

This consensus partition, called *central partition*, minimises the average number of disagreements between all of the partitions and has been considered for instance in [14, 5] in a different context from ours. We consider it in the context of partition-distance defined in [7].

We focus our attention in two particular distance functions between partitions and then do an experimental comparison between the two corresponding central partitions.

In addition, by using the concept of strong patterns [maximal subset of elements that are always clustered together in all partitions], we define a new graph where the nodes are the strong patterns.

This graph contains essentially the same information as the partition graph corresponding to the set *E* defined in [7], but is much simpler as the number of strong patterns is expected to be much smaller than the cardinal of *E*. Then, some properties of this new graph are proved.

EXTENSIONS OF KATZ-PANJER FAMILIES OF DISCRETE DISTRIBUTIONS

Authors: *Dinis D. Pestana* e *Sílvio F. Velosa*

Let $N_{\alpha,\beta,\gamma}$ be a discrete random variable whose probability atoms $\{p_n\}_{n\in N}$ satisfy $\frac{f(n+1)}{f(n)} = \alpha + \beta \frac{E(U_0^n)}{E(U_\gamma^n)}, n = 0,1,\ldots,$ for some $\alpha, \beta \in \Re$, where $U_\gamma \sim Uniform(\gamma,1), \gamma \in (-1,1]$.

When $\gamma \to 1, U_\gamma \to U_1$, the degenerate random variable with unit mass at 1, and the above iterative expression is $\frac{p_{n+1}}{p_n} = \alpha + \frac{\beta}{n+1}$ for $n = k, k+1, \ldots,$ used by Katz and by Panjer (*k =0*), by Sundt and Jewell and by Willmot (*k =1*) and, for general $k \in N$, by Hess, Lewald and Schmidt.

We investigate the case $U_\gamma \sim Uniform(\gamma,1)$ with $\gamma \in (-1,1)$ in detail for $\alpha = 0$.

We then construct classes C_γ of discrete infinitely divisible randomly stopped sums such that $N_{\alpha,\beta,\gamma} \in C_\gamma$. C_0 is the class of compound geometric random variables, C_1 is the class of compound Poissons, and $|\gamma_1| < \gamma_2 \le 1$ implies $C_{\gamma_1} \subset C_{\gamma_2} \subseteq C_1$.

www.ine.pt Informações adicionais: Serviço de Comunicação e Imagem [Tel: 21.842.61.00 _ Fax: 21.842.63.73 _ sc@ine.pt]

EXTREMAL BEHAVIOUR IN MODELS OF SUPERPOSITION OF RANDOM

Author: *Luísa Pereira*

Let $\mathbf{X}^{(i)} = \left\{X_{g_i(n)}\right\}_{n\geq 1}, i = 1,2,$ be sequences of random variables, where $\left\{g_i(n)\right\}_{n\geq 1}$ are disjoint and strictly increasing sequences of integer numbers such that $\left\{g_1(n)\right\}_{n\geq 1} \cup \left\{g_2(n)\right\}_{n\geq 1} = N$.

Using superposition of point processes, we study the extremal behaviour of a superposed sequence

$$\left\{X_n\right\}_{n\geq 1} = \left\{X_{g_1(n)}\right\}_{n\geq 1} \cup \left\{X_{g_2(n)}\right\}_{n\geq 1},$$

where we consider the proportion of variables superposed from each sequence asymptotically constant and $\left\{X_n\right\}_{n\geq 1}$ verifying some dependence conditions. We apply the obtained results in the computation of the bivariate extremal index.

www.ine.pt Informações adicionais: Serviço de Comunicação e Imagem [Tel: 21.842.61.00 _ Fax: 21.842.63.73 _ sc@ine.pt]