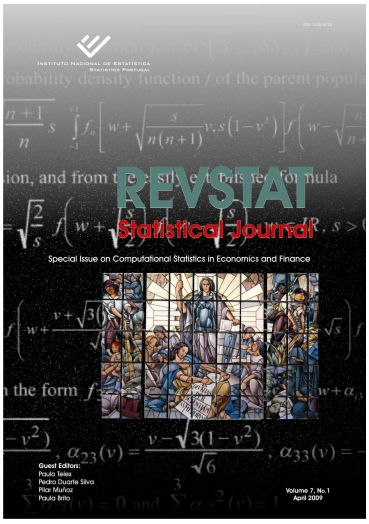
26 october, 2011

**Multithemes**

Volume 9, Number 2, June 2011

## REVSTAT- STATISTICAL JOURNAL

**REVSTAT-STATISTICAL JOURNAL, Volume 9, No. 2 – June 2011**




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** is proud to be chosen by the scientific community to publish their research results.

The **REVSTAT** is covered by the following abstracting/indexing services: *Current Index to Statistics, DOAJ, Google Scholar, Journal Citation Reports/Science Edition, Mathematical Reviews, Science Citation Index Expanded®, SCOPUS* and *Zentralblatt für Mathematic*.

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

This Volume of **REVSTAT: Volume 9, No. 2 - June 2011**, includes four articles. Their abstracts are presented below:

**A MEASURE OF DEPARTURE FROM AVERAGE MARGINAL HOMOGENEITY FOR SQUARE CONTINGENCY TABLES WITH ORDERED CATEGORIES**

Authors: *Kouji Yamamoto, Shuji Ando* and *Sadao Tomizawa*.

For the analysis of square contingency tables, Tomizawa, Miyamoto and Ashihara (2003) considered a measure to represent the degree of departure from marginal homogeneity. However, the maximum value of this measure cannot distinguish two kinds of marginal inhomogeneity. This paper proposes a measure which can distinguish two kinds of marginal inhomogeneity for square tables with ordered categories. The measure is constructed using the arc-cosine function of symmetric cumulative probabilities. Especially the proposed measure is useful for

representing the degree of departure from marginal homogeneity when the extended marginal homogeneity model holds. Examples are given.

**STATISTICS OF EXTREMES IN ATHLETICS**

Authors: *Lígia Henriques-Rodrigues, M. Ivette Gomes* and *Dinis Pestana.*

TV shows on any athletic event make clear that those who want *gold medals* cannot dispense *statistics*. And the statistics more appealing to champions and coachers are the *extreme order statistics*, and in particular *maximum* (or *minimum*) *values* and *records*. The models in *statistics of extremes* are usually semi-parametric or even non-parametric in nature, with the imposition of a few *regularity conditions* in the appropriate tail of the unknown model underlying the available data. The primordial parameter is the *extreme value index*, the shape parameter in the (unified) *extreme value* distribution. The estimation of the *extreme value index* is one of the basis for the estimation of other parameters of rare events, like the *right endpoint* of the model underlying the data, a *high quantile*, the *return period* and the *probability of exceedance* of a high level. In this paper, we are interested in an application of *statistics of extremes* to the best personal marks in a few athletic events. Due to the way data are collected, we begin with a parametric data analysis, but we pay special attention to the semi-parametric estimation of the extreme value index and the right endpoint whenever finite, the possible *world record*, given the actual conditions. In order to achieve a better decision we consider a few alternative semi-parametric estimators available in the literature, and heuristic rules for the choice of thresholds.

**A SPATIAL UNIT LEVEL MODEL FOR SMALL AREA ESTIMATION**

Authors: *Pedro S. Coelho* and *Luis N. Pereira*.

This paper approaches the problem of small area estimation in the framework of spatially correlated data. We propose a class of estimators allowing the integration of sample information of a spatial nature. Those estimators are based on linear models with spatially correlated small area effects where the neighbourhood structure is a function of the distance between small areas. Within a Monte Carlo simulation study we analyze the merits of the proposed estimators in comparison to several traditional estimators. We conclude that the proposed estimators can compete in precision with competitive estimators, while allowing significant reductions in bias. Their merits are particularly conspicuous when analyzing their conditional properties.

**GENERALIZED SUM PLOTS**

Authors: *J. Beirlant, E. Boniphace* and *G. Dierckx*.

Sousa and Michailidis (2004) developed the sum plot based on the Hill (1975) estimator as a diagnostic tool for selecting the optimal *k* when the distribution is heavy tailed. We generalize their method to any consistent estimator with any tail type (heavy, normal and light tail). We illustrate the method associated to the generalized Hill estimator and the moment estimator.

As an attempt to reduce the bias of the generalized Hill estimator, we propose new estimators based on the regression model which are based on the estimates of the generalized Hill estimator. Here weighted least squares and weighted trimmed least squares is proposed. The bias and the mean squared error (MSE) of the estimators is studied using a simulation study. A few practical examples are proposed.